



Universidade Estadual de Maringá  
Centro de Ciências Exatas  
Departamento de Física

Dissertação de Mestrado

**Propriedades estatísticas das diferenças de tamanho entre plantas vizinhas em fileiras de milho (*Zea mays* L.)**

Acadêmico: Paulo Henrique Gonsalves

Orientador: Prof. Dr. Renio dos Santos Mendes

Co-Orientador: Prof. Dr. Sergio Picoli

Maringá, 30 de agosto de 2023



Universidade Estadual de Maringá  
Centro de Ciências Exatas  
Departamento de Física

Dissertação de Mestrado

**Propriedades estatísticas das diferenças de tamanho entre plantas vizinhas em fileiras de milho (*Zea mays* L.)**

Dissertação apresentada ao Programa de Pós-Graduação em Física da Universidade Estadual de Maringá como requisito parcial para obtenção do título de Mestre em Física.

Acadêmico: Paulo Henrique Gonsalves

Orientador: Prof. Dr. Renio dos Santos Mendes

Co-Orientador: Prof. Dr. Sergio Picoli

Maringá, 30 de agosto de 2023

Dados Internacionais de Catalogação-na-Publicação (CIP)  
(Biblioteca Central - UEM, Maringá - PR, Brasil)

G639p

Gonsalves, Paulo Henrique

Propriedades estatísticas das diferenças de tamanho entre plantas vizinhas em fileiras de milho (*Zea mays* L.) / Paulo Henrique Gonsalves. -- Maringá, PR, 2023.  
68 f.: il. color., figs., tabs.

Orientador: Prof. Dr. Renio dos Santos Mendes.

Coorientador: Prof. Dr. Sergio Picoli.

Dissertação (Mestrado) - Universidade Estadual de Maringá, Centro de Ciências Exatas, Departamento de Física, Programa de Pós-Graduação em Física, 2023.

1. Milho (*Zea Mays* L.). 2. Milho (*Zea Mays* L.) - Monocultura. 3. Estatística. 4. Física. I. Mendes, Renio dos Santos, orient. II. Picoli, Sergio, coorient. III. Universidade Estadual de Maringá. Centro de Ciências Exatas. Departamento de Física. Programa de Pós-Graduação em Física. IV. Título.

CDD 23.ed. 530.01

PAULO HENRIQUE GONSALVES

**PROPRIEDADES ESTATÍSTICAS DAS DIFERENÇAS  
DE TAMANHO ENTRE PLANTAS VIZINHAS EM  
FILEIRAS DE MILHO (*Zea mays* L.)**

Dissertação apresentada à Universidade  
Estadual de Maringá, como requisito parcial  
para a obtenção do título de mestre.

Aprovado em: Maringá, 30 de agosto de 2023.

**BANCA EXAMINADORA**

---

Prof. Dr. Renio dos Santos Mendes  
Universidade Estadual de Maringá – UEM

---

Prof. Dr. Sergio de Picoli Junior  
Universidade Estadual de Maringá – UEM

---

Prof. Dr. Luciano Rodrigues da Silva  
Universidade Federal do Rio Grande do Norte – UFRN

---

Dr. Diego Domingues Lopes  
Programa de Pós-Graduação em Física – PFI/UEM

# Sumário

Agradecimentos	v
Resumo	vi
Abstract	vii
Introdução	1
1 Variabilidade de tamanho em monoculturas de plantas: uma breve revisão	5
2 Métodos	14
2.1 Base de dados . . . . .	14
2.2 Caracterização da variabilidade do tamanho de plantas em uma fileira .	19
2.2.1 Assimetria, coeficiente de variação, índice de Gini e coeficiente de autocorrelação . . . . .	19
2.2.2 Variação local, <i>burstiness</i> e memória . . . . .	21
2.3 Correlações entre pares de índices . . . . .	23
3 Análise dos dados e resultados	24
3.1 Distribuições, correlações e outras propriedades estatísticas . . . . .	24
3.2 Análise das correlações entre pares de índices . . . . .	33
3.3 Variações do valor limite $h$ das diferenças de tamanho . . . . .	37
4 Discussão e conclusões	42
A Momentos das distribuições e $k$ -estatística de Fisher	47
B Índice de Gini e curva de Lorenz	48
C Coeficiente de correlação de Pearson	52
D Processos de Bernoulli e Poisson	54
Referências bibliográficas	56

# Lista de Figuras

1	Ilustração qualitativa de diferenças típicas entre sistemas complexos e sistemas simples ordenados ou desordenados . . . . .	2
1.1	Distribuições das circunferências dos caules de populações de árvores coníferas . . . . .	6
1.2	Distribuições dos pesos secos, dos logaritmos dos pesos secos e das alturas para quatro tipos de plantas anuais . . . . .	7
1.3	nome . . . . .	8
1.4	Distribuição do rendimento (peso dos grãos por planta) em culturas de milho . . . . .	8
1.5	Distribuição do peso seco calculada para diferentes amostras de <i>Plantago rugelli</i> L. (tanchagem) . . . . .	9
1.6	Distribuição do peso seco (biomassa) calculada para diferentes amostras de <i>Pisum sativum</i> L. (ervilha) . . . . .	9
1.7	Distribuição do tamanho (altura) de uma população de <i>Chenopodium album</i> L. (ansarina branca) . . . . .	11
1.8	Evolução temporal do valor médio, do coeficiente de variação, da assimetria e da curtose relativos à biomassa (peso seco) de populações de soja, girassol e milho . . . . .	12
1.9	Distribuições de tamanho de populações de soja, girassol e milho para populações com diferentes densidade . . . . .	13
2.1	Local de coleta das medidas e realização das medidas de tamanho de plantas em monocultura de milho . . . . .	15
2.2	Sequência dos tamanhos de plantas sucessivas . . . . .	17
2.3	Sequência dos incrementos de tamanho entre plantas sucessivas . . . . .	17
2.4	Sequência das diferenças de tamanho normalizadas . . . . .	18
2.5	Representação da ocorrência de eventos ao longo de cada amostra . . . . .	18
3.1	Distribuição do tamanho de plantas $P(s)$ . . . . .	25
3.2	Distribuição dos incrementos normalizados $P(\Delta s')$ . . . . .	26
3.3	Distribuição das diferenças de tamanho normalizadas $P(d')$ . . . . .	26
3.4	Distribuição $P(g)$ . . . . .	27
3.5	Distribuição acumulada dos intervalos de retorno normalizados $P_c(\lambda')$ . . . . .	27
3.6	Coefficiente de variação ( $CV_s$ ), índice de Gini ( $G_s$ ), assimetria ( $Sk_s$ ) e coeficiente de autocorrelação ( $r_s$ ) calculados para a sequência dos tamanhos das plantas . . . . .	28
3.7	Assimetria ( $Sk_{\Delta s}$ ) e curtose ( $K_{\Delta s}$ ) calculadas para as diferenças de tamanho entre plantas vizinhas . . . . .	28

3.8	Coeficiente de variação ( $CV_v$ ), índice de Gini ( $G_v$ ), assimetria ( $Sk_v$ ) e autocorrelação ( $r_v$ ) para as "volatilidades"normalizadas ( $d'_n$ ) . . . . .	29
3.9	Varição local ( $LV_\lambda$ ), <i>burstiness</i> ( $B_\lambda$ ), <i>burstiness</i> corrigido ( $BC'_\lambda$ ) e memória ( $M_\lambda$ ) nos intervalos de retorno $\lambda$ . . . . .	30
3.10	Diagrama <i>MB</i> . . . . .	32
3.11	Coeficiente de correlação de Pearson ( $r$ ) e $p$ - valores correspondentes obtidos por meio do teste $t$ de <i>student</i> para todos os pares de índices . . . . .	35
3.12	Gráficos de dispersão dos 13 pares de índices cujos valores do coeficiente de correlação de Pearson ( $r$ ) são estatisticamente significativos . . . . .	36
3.13	Coeficiente de correlação de Pearson ( $r$ ) para todas as combinações de pares dentro de um grupo de seis índices selecionados. . . . .	37
3.14	Varição de $LV_\lambda$ em função de $h$ no intervalo de 1.50 a 2.50 . . . . .	39
3.15	Varição de $B_\lambda$ em função de $h$ no intervalo de 1.50 a 2.50 . . . . .	39
3.16	Varição de $BC_\lambda$ em função de $h$ no intervalo de 1.50 a 2.50 . . . . .	40
3.17	Varição de $M_\lambda$ em função de $h$ no intervalo de 1.50 a 2.50 . . . . .	40
3.18	Coeficiente de correlação de Pearson ( $r$ ) para um grupo selecionado de pares de índices, calculados para diferentes valores do valor limite $h$ . . . . .	41
B.1	Proporção acumulada da população, $p_n$ versus proporção acumulada da variável $x$ . . . . .	49
B.2	Curvas de Lorenz referentes ao tamanho das plantas . . . . .	51
B.3	Curvas de Lorenz referentes às diferenças de tamanho entre plantas vizinhas . . . . .	51

# Lista de Tabelas

2.1	Descrição das amostras de híbridos de milho . . . . .	16
3.1	Breve descrição do conjunto de índices utilizados no presente trabalho para cada amostra . . . . .	33

# Agradecimentos

Agradeço à minha família pelo apoio, principalmente a meus pais. Mesmo estando longe eles foram minha base. Agradeço também ao meu irmão que sempre esteve comigo nessa jornada.

Aos professores Renio do Santos Mendes e Sergio Picoli pela orientação, paciência e conhecimento nesses anos de mestrado.

Aos meus amigos, que mesmo longe, fizeram-se presentes.

Aos professores e alunos (agora também amigos) membros do ComplexLab. Obrigado por me acolherem e tornarem os dias de trabalho mais divertidos.

Aos meus "divertidamente", que seguraram essa barra que é fazer mestrado em física.

# Resumo

Nas últimas décadas, efeitos da competição na estrutura de tamanho de populações de plantas têm sido objeto de intenso estudo em biologia, ecologia, botânica e agronomia. A variabilidade de tamanho de uma população de plantas tem sido caracterizada principalmente por meio da distribuição dos tamanhos e de índices relacionados a essa distribuição, como o coeficiente de variação, a assimetria e o índice de Gini. No presente trabalho, exploramos uma maneira alternativa de caracterizar a variabilidade do tamanho de plantas em monoculturas. Para ilustrar o método, utilizamos uma base de dados contendo a altura de plantas sucessivas em fileiras aproximadamente lineares em plantações de milho (*Zea mays* L.). Com base nesses dados, investigamos propriedades estatísticas i) dos tamanhos das plantas; ii) dos incrementos de tamanho entre plantas sucessivas; iii) das diferenças de tamanho entre plantas sucessivas (valor absoluto dos incrementos); e iv) dos intervalos de retorno entre diferenças de tamanho que excedem um certo limite. Além dos índices padrão, utilizamos índices que têm sido aplicados recentemente no estudo de uma grande variedade de sistemas complexos, como a variação local, o *burstiness* e a memória. Por fim, utilizamos uma análise das correlações entre pares de índices para identificar possíveis redundâncias.

**Palavras chave:** Monoculturas, diferenças de tamanho entre plantas vizinhas, intervalos de retorno, variação local, *burstiness*, memória.

# Abstract

In the last decades, competition effects in the size structure of plant populations have been the subject of intense research in biology, ecology, botany and agronomy. The size variability of a plant population has been characterized mainly through size distributions and related measures like coefficient of variation, skewness and Gini index. In the present study, we explore an alternative way to characterize the size variability of plant populations. We investigate a database containing the size of successive plants along nearly linear rows of maize (*Zea mays* L.). Specifically, we focused on statistical properties of i) plant sizes; ii) size increments between successive plants; iii) size differences between neighboring plants (the absolute value of size increments); and iv) return intervals between size differences that exceed a limit value. In addition to standard measures, we use indices that have been applied recently to study a variety of complex systems, such as local variation, burstiness and memory. Finally, we performed a pairwise correlation analysis to identify possible redundancies.

**Keywords:** Monocultures, size differences between neighboring plants, return intervals, local variation, burstiness, memory.

# Introdução

Sistemas complexos são usualmente compostos por um grande número de partes que interagem entre si. Em parte devido a essas interações, que são tipicamente não lineares, emergem propriedades que não ocorreriam se cada parte fosse analisada isoladamente. Outras características dos sistemas complexos incluem comportamento irregular (estocástico ou caótico) e auto-organização [1–3]. A figura 1 ilustra de maneira geral e qualitativa algumas diferenças que costumam ser observadas entre sistemas complexos e sistemas simples (ordenados ou desordenados). Em sistemas simples ordenados é usual observar padrões temporais, espaciais ou estruturais periódicos ou regulares. Em sistemas simples desordenados, costuma-se observar padrões temporais, espaciais ou estruturais aleatórios e não correlacionados. Por outro lado, sistemas complexos tendem a ocupar um estado intermediário entre esses dois extremos — entre o totalmente aleatório e o completamente regular ou periódico. Por exemplo, sistemas complexos podem exibir padrões temporais, espaciais ou estruturais não triviais tais como correlações de longo alcance, distribuições do tipo lei de potência e redes livre de escala. Plantas são sistemas vivos altamente organizados, compostos por elementos que interagem entre si. As interações ocorrem desde o nível molecular em um indivíduo até o nível macroscópico, com interações entre indivíduos em uma população de plantas. Essas interações tipicamente levam a diversos tipos de comportamento emergente [4–6].

Populações de plantas tipicamente exibem variabilidade de tamanho entre os indivíduos que a compõem. A palavra ‘tamanho’ aqui pode se referir à altura, ao peso seco (biomassa), à produção de sementes, ao diâmetro do caule ou a outra característica estrutural da planta. De modo geral, a variabilidade de tamanho dos indivíduos pode estar relacionada a fatores tais como diferenças de idade, diferenças genéticas, heterogeneidade do meio, efeitos de herbívoros, parasitas ou patógenos e competição. De modo geral, a estrutura de tamanho dos indivíduos em uma população de plantas será o resultado da interação entre fatores como esses. Em particular, monoculturas ou plantações agrícolas costumam ser caracterizadas por indivíduos de mesma idade e com genética semelhante. Além disso, a terra é cultivada — de modo que toda a população tenha os nutrientes necessários para seu desenvolvimento — e empregam-se

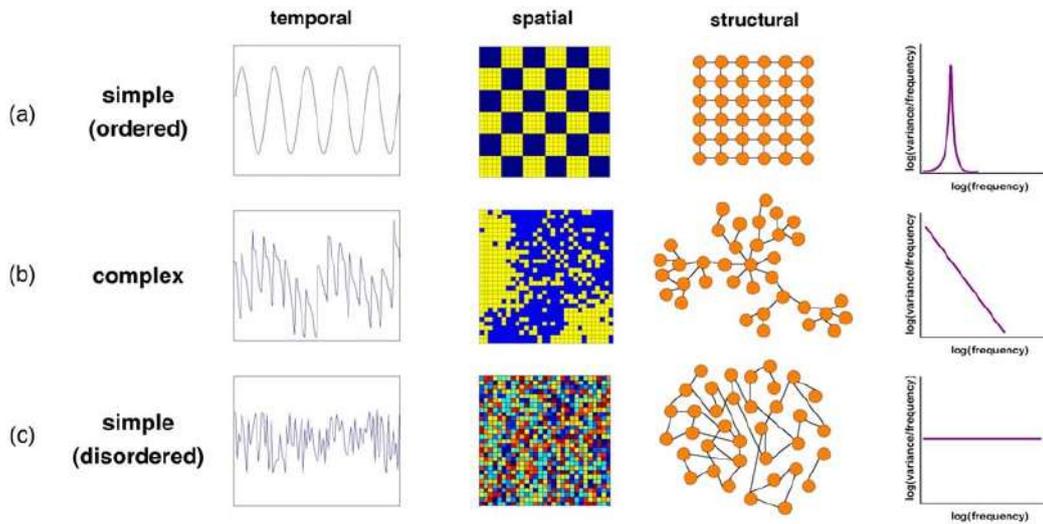


Figura 1: Ilustração qualitativa de diferenças típicas entre sistemas complexos e sistemas simples ordenados ou desordenados. Fonte: Referência [6] (M. Anand, A. Gonzalez, F. Guichard, J. Kolasa and L. Parrott, 2010).

técnicas para o controle de pragas. Desse modo, a competição entre plantas é um fator que se destaca no ambiente das monoculturas. Nessa direção, a estrutura de tamanho de plantas em monoculturas e sua relação com efeitos de competição tem sido objeto de vários estudos em biologia, ecologia, botânica e agronomia [7–26]. A variabilidade de tamanho pode ter implicações para as monoculturas de plantas. Por exemplo, há evidências de que o tamanho pode afetar a mortalidade e a produtividade de plantas individuais bem como a mortalidade e a produtividade de seus vizinhos. Tipicamente, plantas maiores são mais produtivas e têm maior probabilidade de sobreviver do que plantas menores [27–29].

Uma questão fundamental no estudo dos efeitos da competição em monoculturas tem a ver com a caracterização da variabilidade de tamanho da população. Nas últimas décadas, uma maneira bastante utilizada para descrever a variabilidade de tamanho é por meio da distribuição dos tamanhos das plantas e seus momentos estatísticos. Três índices bastante utilizados no estudo de monoculturas são o coeficiente de variação (razão entre o desvio padrão e o valor médio), a assimetria e a curtose da distribuição dos tamanhos [8, 9, 17]. Mais recentemente, a desigualdade de tamanho das plantas em monoculturas tem sido quantificada por meio do índice de Gini [15–17, 22]. O índice de Gini é uma medida bastante utilizada em economia, por exemplo, para quantificar a desigualdade de renda em uma população. Uma comparação entre o coeficiente de variação, a assimetria e o índice de Gini para caracterizar a estrutura do tamanho de plantas em monoculturas pode ser encontrada na ref. [30]. O coeficiente de variação, a assimetria, a curtose e o índice de Gini podem ser calculados sem levar em conta a dis-

tribuição espacial dos indivíduos da população. Por exemplo, esses índices costumam ser aplicados a dados coletados em uma área da plantação ou em uma linha de plantio ou fileira aproximadamente linear. Sequências espaciais relacionadas a uma fileira de plantas apresentam uma certa analogia com séries temporais. Essa analogia é ainda mais forte quando a distância entre plantas sucessivas ao longo da fileira é aproximadamente constante (como ocorre em plantações agrícolas). No contexto dos sistemas complexos e da física estatística, sequências espaciais têm sido investigadas usando-se técnicas comumente utilizadas no estudo de séries temporais. Exemplos incluem o estudo da sequência de nucleotídeos em cadeias de DNA [31–33] e da sequência de letras ou outras estruturas ao longo de textos escritos [34, 35]. Outro exemplo recente tem a ver justamente com a variabilidade de tamanho de plantas ao longo de fileiras em plantações agrícolas [36].

Uma maneira de investigar a dinâmica (ou a estrutura espacial) de sistemas complexos é por meio da análise dos tempos (ou distâncias) entre eventos [37–47]. Isso inclui a análise dos intervalos de retorno em sequências temporais e/ou espaciais [48–51]. Mais especificamente, os intervalos de retorno medem o tempo ou a distância entre eventos consecutivos, em que um evento ocorre sempre que a variável em estudo excede um certo valor limite. Um exemplo recente refere-se ao estudo da volatilidade de preços no mercado financeiro. Nesse contexto, a volatilidade pode ser definida como a diferença logarítmica entre preços consecutivos. Em trabalhos anteriores sobre o tema, o foco da investigação foi a distribuição da volatilidade e a presença de correlações temporais na série das volatilidades [52–54]. Mais recentemente, o estudo da volatilidade de preços no mercado financeiro foi revisitado. Dessa vez o foco foi a análise dos intervalos de retorno entre eventos que excedem um certo valor limite [55–57]. De uma maneira geral, índices calculados a partir de tempos entre eventos ou intervalos de retorno têm sido utilizados para investigar padrões em uma variedade de sistemas complexos. Vamos citar aqui três desses índices: a variação local [58, 59], o coeficiente *burstiness* e a memória [60]. Aplicações recentes desses índices incluem o estudo de terremotos, séries meteorológicas, atividades de comunicação (como e-mails, ligações telefônicas, mensagens de texto, redes sociais), atividade física e movimentos posturais [61–70].

No presente trabalho, utilizamos índices tradicionais, tais como o coeficiente de variação, a assimetria e o índice de Gini, para analisar uma base de dados contendo o tamanho de plantas sucessivas ao longo de fileiras em plantações de milho. Investigamos também a organização sequencial dos tamanhos das plantas ao longo das fileiras utilizando o coeficiente de autocorrelação. Assim como na ref. [36], analisamos propriedades estatísticas dos incrementos de tamanho ao longo das fileiras de milho. Em contraste com esse trabalho anterior, utilizamos a assimetria e a curtose da distribuição dos incrementos de tamanho. Além disso, analisamos propriedades estatísticas dos

valores absolutos dos incrementos de tamanho ou diferenças de tamanho entre plantas vizinhas. Para isso utilizamos o coeficiente de variação, a assimetria, o índice de Gini e o coeficiente de autocorrelação. Estudamos também propriedades estatísticas dos intervalos de retorno entre diferenças de tamanho que excedem um certo limite. Nessa análise utilizamos a variação local, o coeficiente *burstiness* e o coeficiente memória. Por fim, investigamos a presença de possíveis correlações entre diferentes pares de índices. O objetivo dessa análise é indicar possíveis redundâncias (índices fortemente correlacionados) e ajudar na seleção de índices apropriados para caracterizar diferentes aspectos da variabilidade de tamanho de plantas em monoculturas. No capítulo 1, apresentamos uma breve revisão sobre resultados empíricos no estudo da estrutura de tamanho de populações de plantas em monoculturas. No capítulo 2, descrevemos e visualizamos a base de dados utilizada — a sequência dos tamanhos (alturas) das plantas em fileiras de milho. Além disso, definimos a sequência dos incrementos de tamanho e a sequência dos valores absolutos dos incrementos de tamanho (as diferenças de tamanho entre plantas vizinhas). Definimos também os índices de variabilidade utilizados nesse trabalho. No capítulo 3, apresentamos a análise dos dados e os resultados obtidos. Isso inclui a apresentação dos valores empíricos dos índices e uma análise das correlações entre pares de índices. Por fim, apresentamos as discussões e conclusões.

# Capítulo 1

## Variabilidade de tamanho em monoculturas de plantas: uma breve revisão

Nesse capítulo, fazemos uma breve revisão da literatura sobre a análise empírica da variabilidade de tamanho de plantas em monoculturas. Com base em alguns trabalhos representativos, vamos ver exemplos de aplicações de índices tradicionais no estudo de monoculturas. Nesse contexto, índices tradicionais costumam ser calculados para populações com diferentes densidades de plantas. Além disso, índices tradicionais usualmente são calculados em função do tempo (idade da população de plantas). Esse tipo de análise tem ajudado a elucidar o papel da competição na variabilidade de tamanho das plantas em uma monocultura.

Na ref. [7] (E. D. Ford, 1975), uma das bases de dados investigada se refere às circunferências do caule de populações de árvores coníferas (*Picea sitchensis* L.) localizadas em uma região dos Estados Unidos. Na figura (1.1), vemos as distribuições das circunferências de diferentes amostras de floresta, com diferentes densidades. Os resultados sugerem que as distribuições tendem a ser mais assimétricas para as amostras com maior densidade de plantas. O índice utilizado para medir esse efeito foi a assimetria (representada por  $g_1$  na figura). Ainda na ref. [7], outras bases de dados analisadas se referem a espécies de plantas de ciclo anual, incluindo *Lycopersicon esculentum* var. (tomate), *Tagetes patula* L. (cravo-de-defunto), *Tagetes erecta* L. (cravo) e *Sinapis alba* L. (mostarda-branca). Nesse experimento, as plantas foram colhidas e sua altura e peso seco (biomassa) foram medidos. Isso foi feito para diferentes amostras de plantas, com diferentes densidades. Conforme ilustrado na figura 1.2, em alguns casos as distribuições de peso e de altura exibem comportamento não gaussiano. O índice utilizado aqui foi a curtose (representada por  $g_2$  na figura). Em alguns casos, a distribuição do logaritmo dos pesos é aproximadamente normal indicando que a distribuição

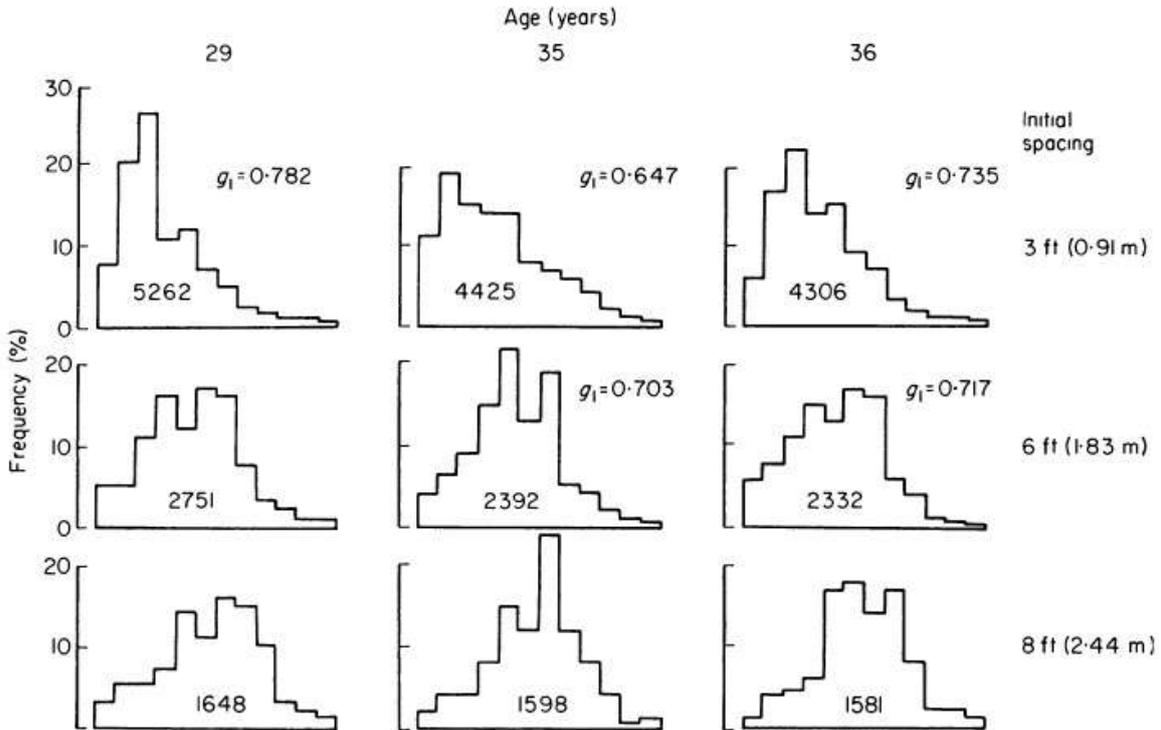


Figura 1.1: Distribuições das circunferências dos caules de populações de árvores coníferas. Os dados se referem a três diferentes amostras com diferentes densidades: as densidades (em árvores por hectare) estão indicadas dentro de cada histograma. As medições foram repetidas para três idades da população: 29, 35 e 36 anos. Fonte: Referência [7] (E. D. Ford, 1975).

dos pesos pode ser aproximada por uma log-normal. Além disso, verificou-se que as distribuições de tamanho e peso seco tendem a se afastar da distribuição normal com o aumento na densidade de plantas.

Na ref. [9] (G. O. Edmeades and T. B. Daynard, 1979), os dados se referem a plantações de milho (*Zea mays* L.) em uma estação de pesquisa na Inglaterra. O tamanho médio da área cultivada para cada amostra foi de 6m por 8m. As densidades de plantas nessas amostras foram de 50 mil, 100 mil, 150 mil e 200 mil plantas por hectare. Na figura (1.3), vemos a evolução temporal do valor médio, do desvio padrão e do coeficiente de variação do peso seco das plantas para cada uma das quatro amostras. Na figura, vemos uma tendência de crescimento do valor médio e do desvio padrão com o tempo para todas as densidade analisadas. Por outro lado, o coeficiente de variação permanece mais estável ao longo do tempo. Por fim, após um tempo suficiente, o coeficiente de variação tende a assumir um valor maior para as amostras com maior densidade de plantas. Na figura (1.4), vemos a distribuição do rendimento (peso dos grãos produzidos por planta) para as mesmas quatro amostras de milho. Em baixas densidades a distribuição é praticamente simétrica e aproximadamente gaussiana. À

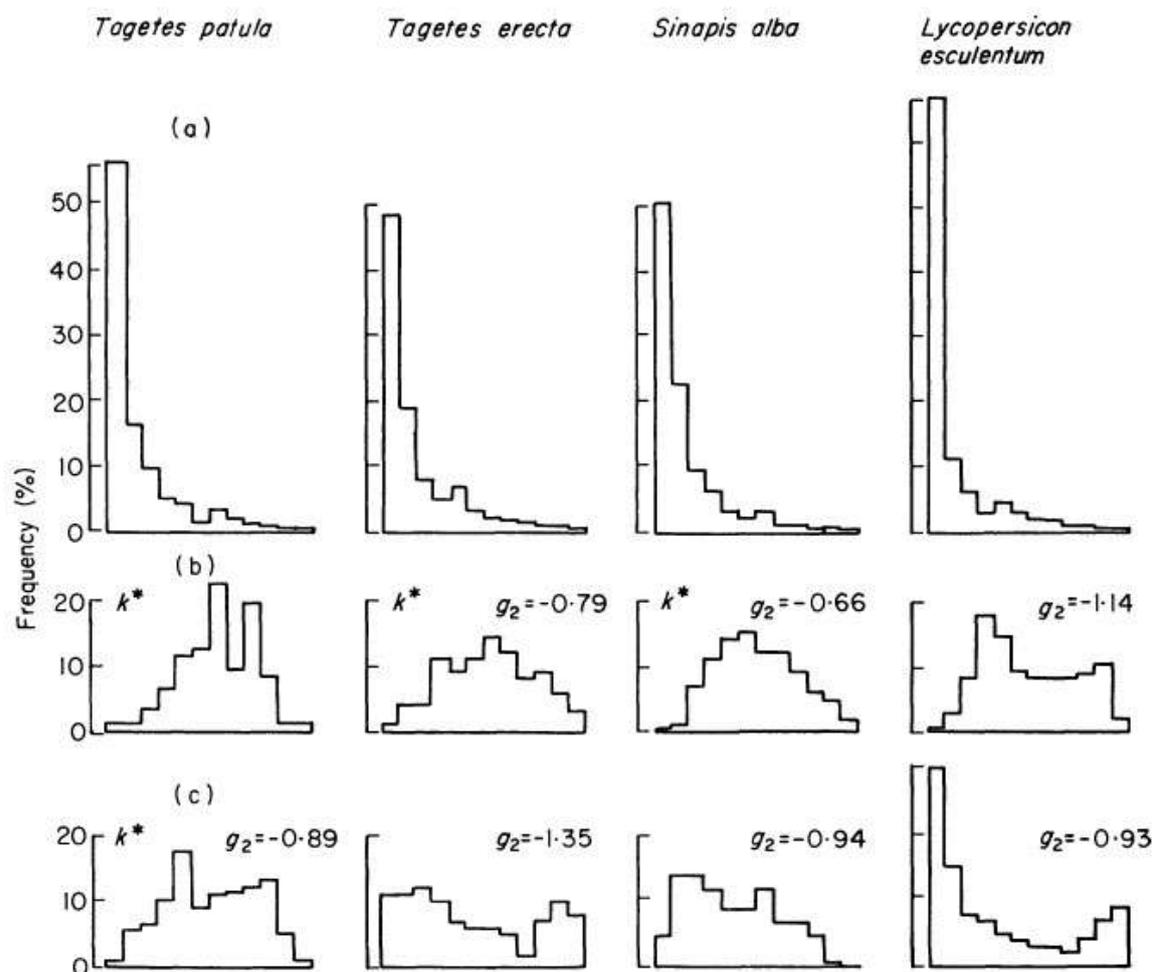


Figura 1.2: Distribuições dos pesos secos (em cima), dos logaritmos dos pesos secos (intermediário) e das alturas (em baixo) para quatro tipos de plantas anuais. Os valores da curtose ( $g_2$ ) são mostrados para as distribuições dos logaritmos do peso e das alturas.  $k^*$  indica que as distribuições são aproximadamente normais. Fonte: Referência [7] (E. D. Ford, 1975).

medida que a densidade aumenta, a distribuição se torna cada vez mais bimodal. À medida que a densidade da população aumenta, aumenta também a proporção de plantas com baixo rendimento. Esse resultado sugere que o rendimento médio tende a diminuir com o aumento da densidade.

A ref. [18] (J. Weiner and S. C. Thomas, 1986) apresenta uma breve revisão de resultados empíricos obtidos em trabalhos anteriores. Ao todo são 16 experimentos reportados. Em cada um desses experimentos, os dados são obtidos a partir de populações com diferentes densidades de plantas. Fornecem-se gráficos da distribuição dos tamanhos e medidas de variabilidade como o coeficiente de variação. Para cada caso, os autores calculam (ou pedem para os autores originais calcularem) o índice de Gini de cada população. As figuras 1.5 e 1.6 mostram as distribuições do peso seco e

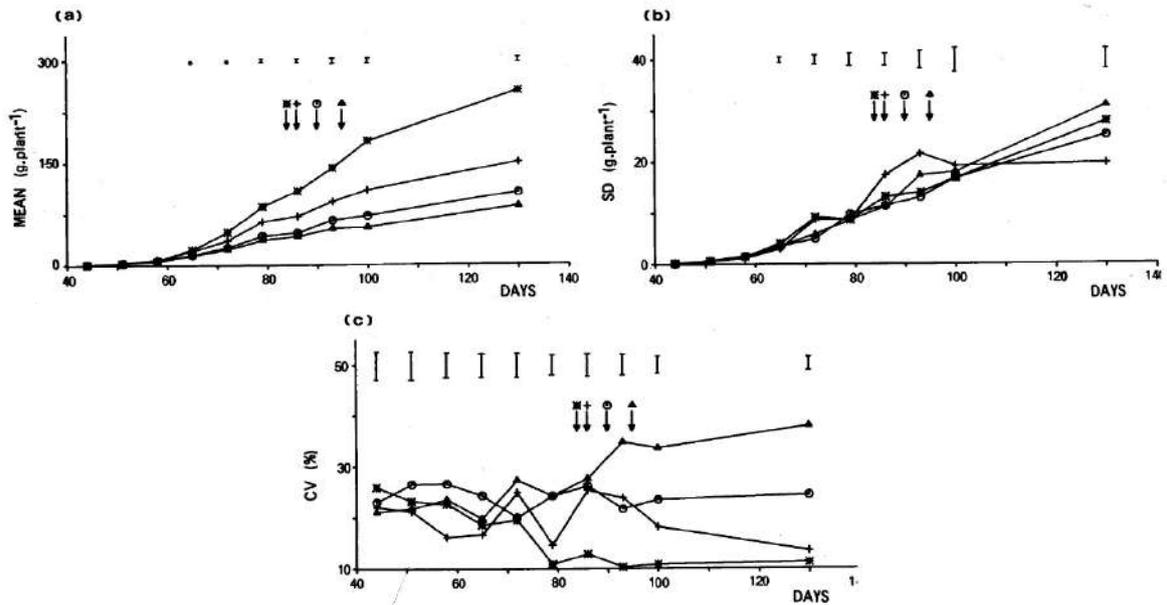


Figura 1.3: a) Valor médio; b) desvio padrão; e c) coeficiente de variação multiplicado por 100 (em porcentagem) do peso seco de plantas de milho em função do tempo (dias a partir da germinação). Os dados se referem a 4 amostras com diferentes densidades de plantas: 50 mil (asteriscos), 100 mil (cruz), 150 mil (círculos) e 200 mil (triângulos) plantas por hectare. Fonte: Referência [9] (G. O. Edmeades and T. B. Daynard, 1979).

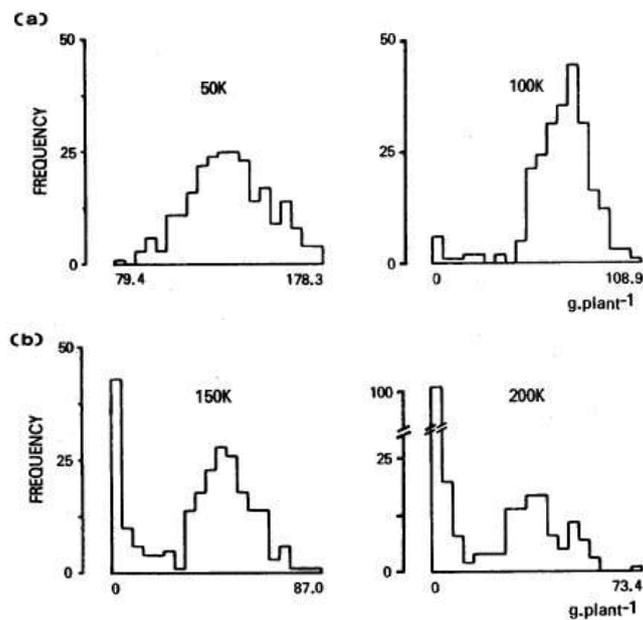


Figura 1.4: Distribuição do rendimento (peso dos grãos por planta) em culturas de milho. Os dados se referem a 4 amostras com diferentes densidades de plantas: a) 50 mil (50k), b) 100 mil (100k), c) 150 mil (150k) e d) 200 mil (200k) plantas por hectare. Fonte: Referência [9] (G. O. Edmeades and T. B. Daynard, 1979).

o índice de Gini calculados em dois desses experimentos: o primeiro para populações de *Plantago rugelli* L. (tanchagem) e o segundo para populações de *Pisum sativum*

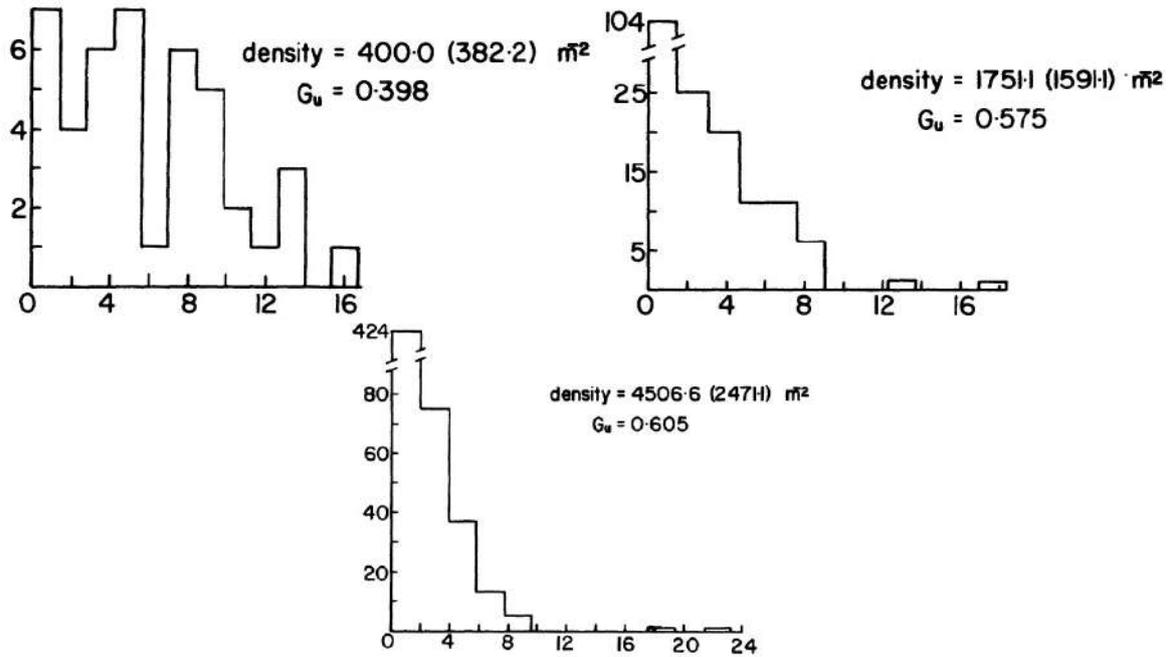


Figura 1.5: Distribuição do peso seco calculada para diferentes amostras de *Plantago rugelli* L. (tanchagem). As densidades são, respectivamente, 400, 1751 e 4506 plantas por metro quadrado. Os valores do índice de Gini ( $G_u$ ) são, respectivamente, 0,398, 0,575 e 0,605. Fonte: Referência [13] (W. R. Hawthorn and P. B. Cavers, 1982) e referência [18] (J. Weiner and S. C. Thomas, 1986).

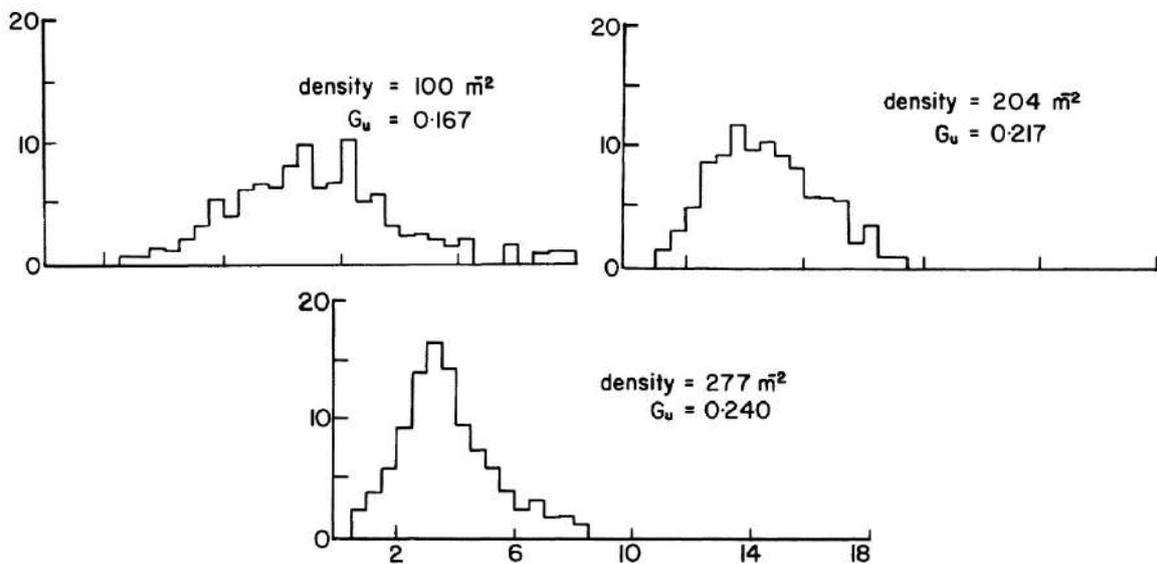


Figura 1.6: Distribuição do peso seco (biomassa) calculada para diferentes amostras de *Pisum sativum* L. (ervilha). As densidades são, respectivamente, 100, 204 e 277 plantas por metro quadrado. Os valores do índice de Gini ( $G_u$ ) são, respectivamente, 0,167, 0,217 e 0,240. Fonte: Referência [14] (C. L. Hedley, M. J. Ambrose e K. A. Pyke, 1983) e referência [18] (J. Weiner and S. C. Thomas, 1986).

L. (ervilha). Nos dois casos, o coeficiente Gini aumenta com o aumento da densidade de plantas. Resultados similares foram obtidos para a maioria dos 16 experimentos analisados.

Na ref. [20] (H. Nagashima, I. Terashima and S. Katoh, 1995), os dados analisados se referem a populações de *Chenopodium album* L. (ansarina branca). Os dados foram obtidos de populações com diferentes densidades de plantas (400, 800 e 3600 plantas por metro quadrado). Para cada densidade, foram obtidas a distribuição dos tamanhos (alturas) das plantas, a assimetria, a curtose e o índice de Gini. Isso foi feito para diferentes idades da população (representadas por  $T_1, T_2, \dots, T_9$ ), com a idade variando de poucos dias após a germinação ( $T_1$ ) até a maturação dos frutos ( $T_9$ ). Na figura 1.7, vemos os resultados obtidos para populações com densidade de 3600 plantas por metro quadrado. No estágio inicial  $T_1$ , a distribuição das alturas das plantas é aproximadamente gaussiana, com assimetria e curtose próximas de zero. Pouco tempo depois, no estágio  $T_2$ , a distribuição já é assimétrica (assimetria positiva) e o índice de Gini apresenta um aumento considerável em relação ao estágio anterior. A partir do estágio  $T_3$ , os valores da assimetria e da curtose são positivos e continuam assim até o estágio de maturação. O índice de Gini alcança valores próximos a 0.4 em  $T_3$  e se mantém assim até  $T_9$ . Outro resultado apontado no trabalho é que a forma final da distribuição dos tamanhos nas populações estudadas depende da densidade de plantas. Por exemplo, na densidade de 400 plantas por metro quadrado a distribuição em  $T_9$  é bimodal enquanto que para uma densidade de 3600 a distribuição em  $T_9$  é assimétrica para a direita.

Na ref. [23] (C. R. C. Vega and V. O. Sadras, 2003), os dados foram obtidos a partir de plantações agrícolas de soja (*Glycine max* L.), girassol (*Helianthus annuus* L.) e milho (*Zea mays* L.) localizadas na Argentina. Para cada espécie, plantações com duas densidades de plantas foram analisadas: padrão e alta densidade. Para a soja, as densidades são 30 e 60 plantas por metro quadrado; Para o girassol, 6 e 12; e para o milho, 8, 5 e 17. Na figura 1.8, vemos a evolução temporal do valor médio, do coeficiente de variação, da assimetria e da curtose do peso seco da soja, do girassol e do milho, calculados para densidade padrão e alta densidade. Para as três espécies consideradas, o coeficiente de variação e a assimetria assumem valores tipicamente maiores para as populações com densidade maior. Esse resultado parece ser robusto: se manteve praticamente durante todo o processo de crescimento das plantas. Comparando as diferentes espécies, vemos que o coeficiente de variação do peso seco é maior na soja e menor no milho. Além disso, enquanto a soja é caracterizada por assimetria positiva, o milho apresenta assimetria negativa. O girassol exibe uma transição de assimetria negativa para baixas densidades para uma assimetria positiva para altas densidades. Na figura 1.9, vemos a distribuição dos tamanhos (alturas) calculadas para as mesmas

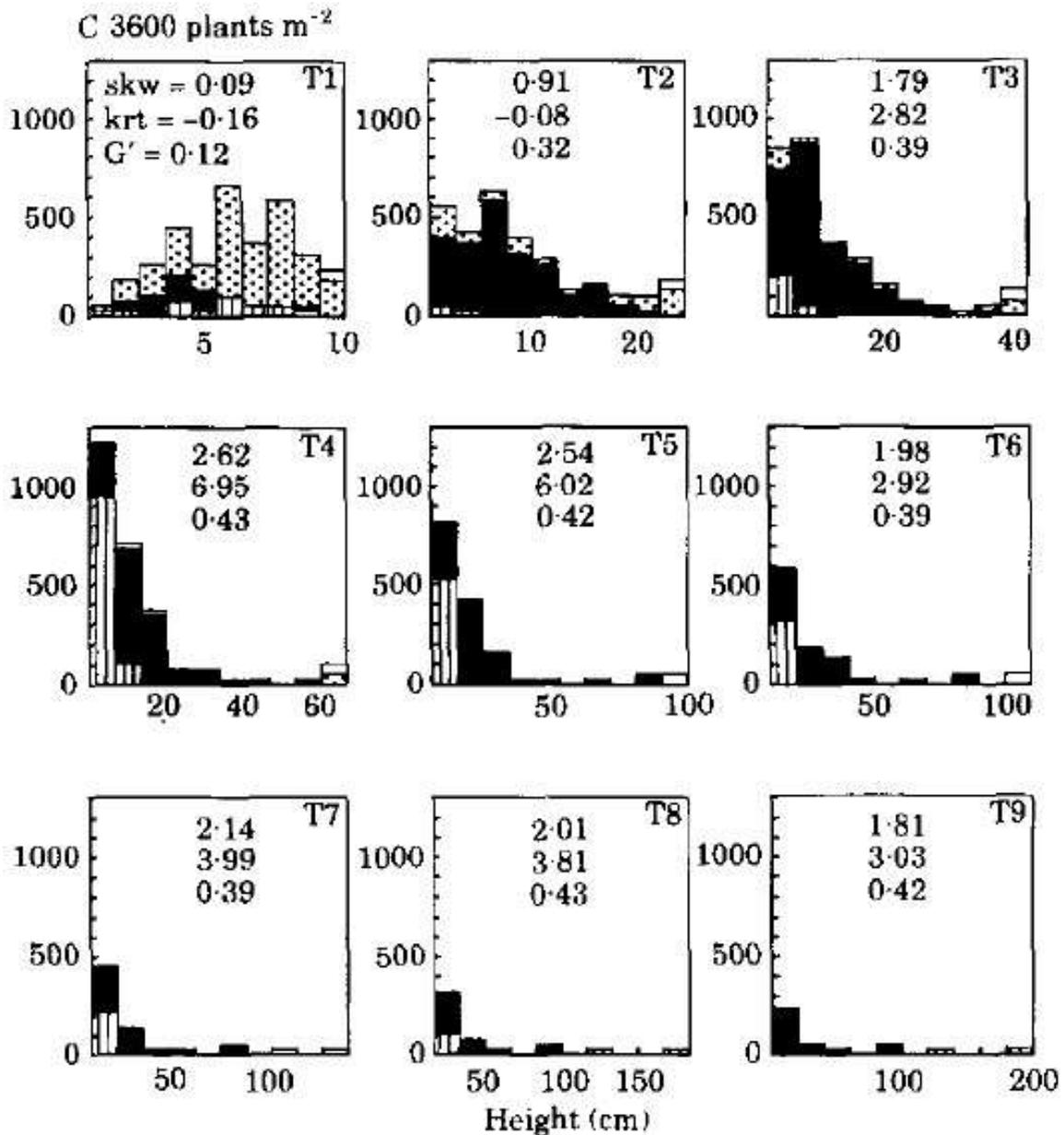


Figura 1.7: Distribuição do tamanho (altura) de uma população de *Chenopodium album* L. (ansarina branca). A densidade é de 3600 plantas por metro quadrado. A distribuição é mostrada para diferentes estágios de desenvolvimento, representados por  $T_1$  (estágio inicial, próximo à germinação) até o estágio final  $T_9$ . Valores da assimetria, da curtose e do índice de Gini são mostrados em cada painel. Fonte: Referência [20] (H. Nagashima, I. Terashima and S. Katoh, 1995).

populações de soja, girassol e milho. Em todos os casos (incluindo para diferentes densidades), as distribuições de tamanho são assimétricas com assimetria negativa.

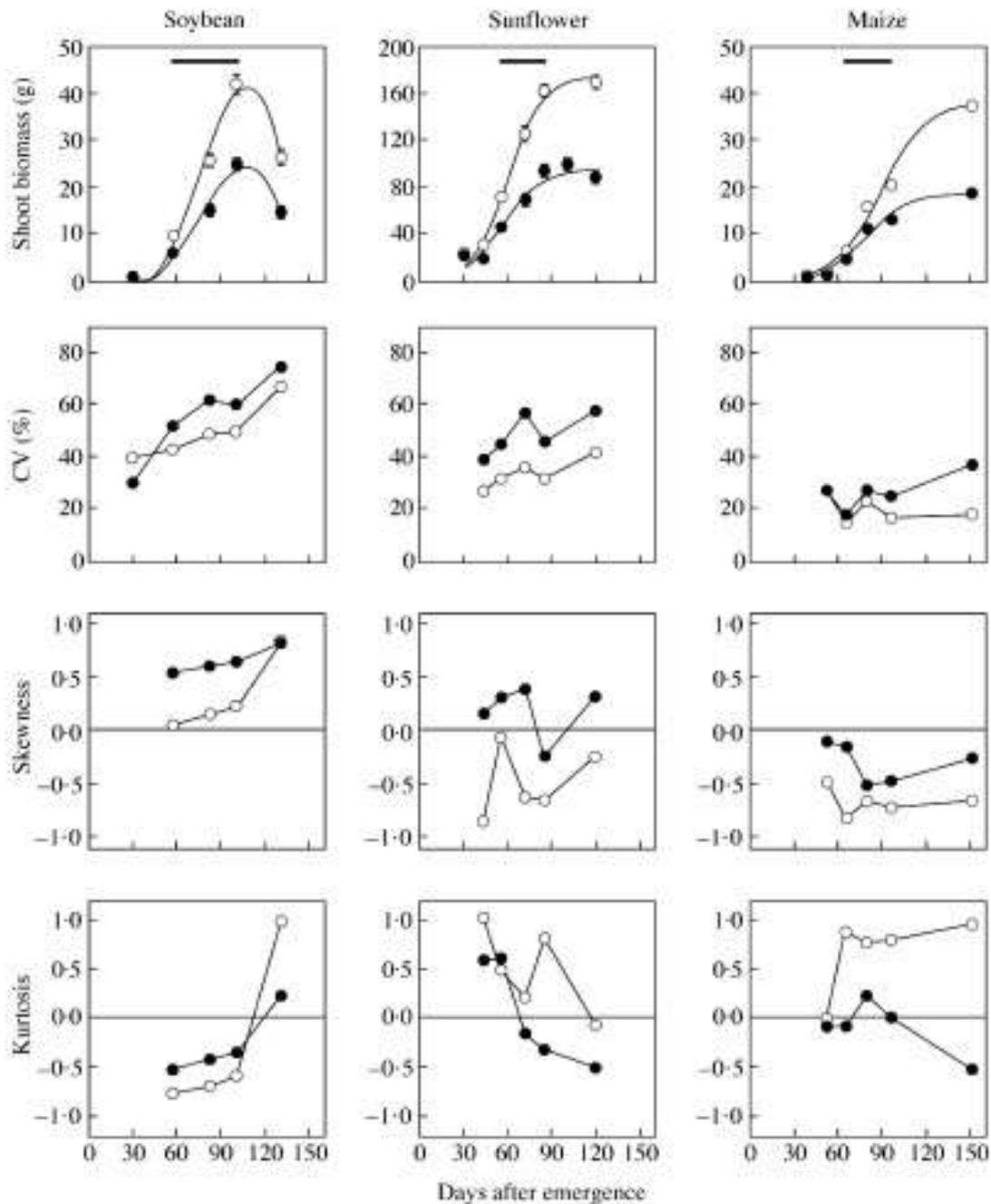


Figura 1.8: Evolução temporal do valor médio, do coeficiente de variação, da assimetria e da curtose relativos à biomassa (peso seco) de populações de soja, girassol e milho. Para cada espécie de planta, os dados se referem a duas populações com diferentes densidades: símbolos abertos (densidade padrão,  $\rho$ ) e símbolos fechados (alta densidade,  $2\rho$ ). Valores de  $\rho$  (em plantas por metro quadrado): 30 para a soja, 6 para o girassol e 8,5 para o milho. Fonte: Referência [23] (C. R. C. Vega and V. O. Sadras, 2003).

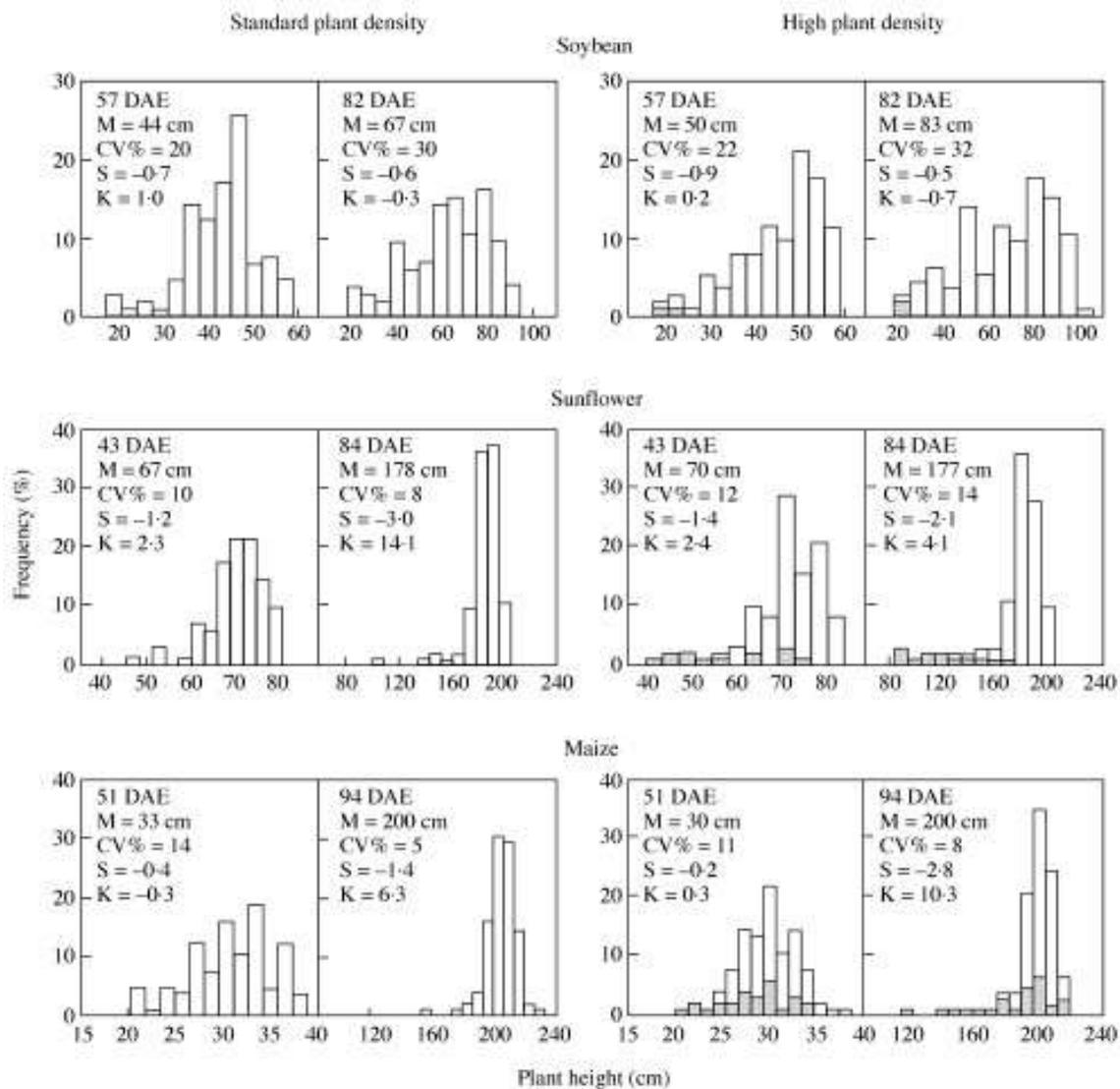


Figura 1.9: Distribuições de tamanho (altura, em cm) de populações de soja, girassol e milho para populações com diferentes densidades: densidade padrão ( $\rho$ ) e alta densidade ( $2\rho$ ). Valores de  $\rho$  (em plantas por metro quadrado): 30 para a soja, 6 para o girassol e 8,5 para o milho. Cada painel contém as seguintes informações: dias após a germinação (DAE), moda (M), coeficiente de variação (CV), assimetria (S) e curtose (K). Fonte: Referência [23] (C. R. C. Vega and V. O. Sadras, 2003).

# Capítulo 2

## Métodos

Neste capítulo, descrevemos a base de dados que foi utilizada para o estudo da variabilidade do tamanho de plantas. Além disso, definimos e descrevemos os índices que foram utilizados na análise desses dados. Esses índices incluem medidas padrão (que têm sido utilizadas há décadas na análise da variabilidade de tamanho em monoculturas) calculados diretamente a partir dos tamanhos das plantas. Incluem também índices calculados a partir i) dos incrementos de tamanho entre plantas sucessivas; ii) dos valores absolutos dos incrementos de tamanho (ou diferenças de tamanho); e iii) dos intervalos de retorno entre diferenças de tamanho que excedem um certo limite.

### 2.1 Base de dados

A base de dados utilizada neste trabalho foi obtida experimentalmente, no ano de 2007, a partir de medições realizadas em plantações de milho (*Zea mays* L.) na cidade de Maringá-PR, localizada na região sul do Brasil (23°20S, 51°57W) e com altitude de aproximadamente 450 metros. Os dados referem-se ao tamanho (altura) de plantas sucessivas ao longo de fileiras aproximadamente lineares localizadas no interior (não nas bordas) de plantações comerciais de milho. A medição foi feita de modo direto, com o auxílio de uma trena. Cada planta foi medida da superfície do solo até a parte mais alta do pendão. Os dados foram obtidos em quatro plantações distintas, cada uma correspondendo a um híbrido de milho (tabela 2.1): P 30f87 (amostra 1), P 30k75 (amostra 2), AG 9010 (amostra 3) e AS 1570 (amostra 4). O espaçamento entre fileiras adjacentes nas lavouras de milho em que as medições foram feitas é de aproximadamente 90cm (espaçamento convencional utilizado na época). Essa base de dados é a mesma utilizada na ref. [36]. Na figura 2.1, vemos uma sequência de imagens de uma das lavouras de milho e do processo de medição da altura das plantas.

O resultado da medição da altura (em m) da  $n$ ésima planta em uma amostra é representado por  $s_n$ ,  $n = 1, \dots, N_s$ , em que  $n$  representa a posição sequencial da planta



Figura 2.1: Em cima, à esquerda: uma das lavouras de milho em que as medições foram feitas, com a cidade de Maringá ao fundo. Em cima, à direita: Vista interior da lavoura, em que o espaçamento entre fileiras adjacentes é de aproximadamente 90cm. Em baixo, à esquerda: Pendões de plantas individuais (as medições foram feitas do solo até a parte mais alta do pendão). Em baixo, à direita: Processo de medição da altura das plantas, em que foi utilizada uma trena fixada em uma régua de madeira.

na amostra e  $N_s$  é o número total de plantas na amostra. De acordo com essa definição,  $s_1$  representa o tamanho da primeira planta da amostra,  $s_2$  representa o tamanho da segunda e assim por diante. Incrementos de tamanho entre plantas sucessivas são definidos como [36]

$$\Delta s_n = s_{n+1} - s_n, \quad (2.1)$$

$n = 1, \dots, N_s - 1$ , em que  $s_n$  é o tamanho da  $n$ -ésima planta e  $s_{n+1}$  o tamanho da planta seguinte. Nessa definição, o elemento seguinte depende do sentido em que a fileira de plantas é percorrida. Aqui utilizamos o sentido original percorrido no processo de medição. De qualquer modo, os módulos dos incrementos  $\Delta s$  não dependem do sentido escolhido. Definimos também os incrementos normalizados,

$$\Delta s'_n = \frac{\Delta s_n - \overline{\Delta s}}{\sigma_{\Delta s}}, \quad (2.2)$$

Tabela 2.1: Descrição das amostras.  $N_s$  é o número total de plantas na amostra.  $\rho$  é a densidade linear de plantas (em plantas por metro) na parte da fileira em que as medições foram feitas.  $\bar{s}$  é o valor médio do tamanho (altura) das plantas em metros.  $\sigma_s$  é o desvio padrão dos tamanhos de plantas.  $\sigma_{\Delta s}$  é o desvio padrão dos incrementos de tamanho  $\Delta s$ .  $\bar{d}$  é o valor médio das diferenças de tamanho  $d$ .  $\sigma_d$  é o desvio padrão das diferenças de tamanho.

Amostra	Híbrido	Fabricante	$N_s$	$\rho$	$\bar{s}$	$\sigma_s$	$\sigma_{\Delta s}$	$\bar{d}$	$\sigma_d$
1	P 30f87	Pioneer	1803	4.21	2.06	0.23	0.31	0.19	0.25
2	P 30k75	Pioneer	2109	4.78	1.69	0.23	0.27	0.20	0.18
3	AG 9010	Agrocere	2111	4.22	2.05	0.31	0.40	0.25	0.31
4	AS 1570	Agroeste	2203	5.64	2.32	0.14	0.18	0.12	0.14

em que  $\bar{\Delta s}$  é o valor médio e  $\sigma_{\Delta s}$  é o desvio padrão calculados a partir dos incrementos  $\Delta s$ .

O valor absoluto de  $\Delta s_n$ ,

$$d_n = |\Delta s_n|, \quad (2.3)$$

refere-se à diferença de tamanho (em módulo) entre plantas sucessivas. A seguir, definimos as diferenças de tamanho normalizadas,

$$d'_n = \frac{d_n}{\sigma_d}, \quad (2.4)$$

em que  $\sigma_d$  é o desvio padrão das diferenças de tamanho  $d$ . O número total de termos na sequência das diferenças de tamanho  $d$  é igual a  $N_s - 1$ , assim como no caso dos incrementos de tamanho  $\Delta s$ . Note que a equação 2.3, que define  $d_n$  como o valor absoluto de uma diferença entre valores consecutivos de uma série, apresenta uma certa similaridade com a definição de volatilidade de preços do mercado financeiro [52–57]. Devido à essa similaridade, as diferenças de tamanho entre plantas vizinhas poderiam ser vistas como um tipo de ‘volatilidade’ de tamanhos (de natureza espacial) em analogia com a volatilidade de preços (de natureza temporal).

Na tabela 2.1 descrevemos algumas propriedades estatísticas dos tamanhos das plantas, dos incrementos de tamanho e do valor absoluto desses incrementos. As figuras 2.2, 2.3 e 2.4 mostram as sequências de valores de  $s$ ,  $\Delta s$  e  $d'$ , respectivamente, para cada amostra.

A seguir, vamos definir os intervalos de retorno entre diferenças de tamanho que excedem um certo limite. Para isso, vamos construir uma variável binária cujos valores

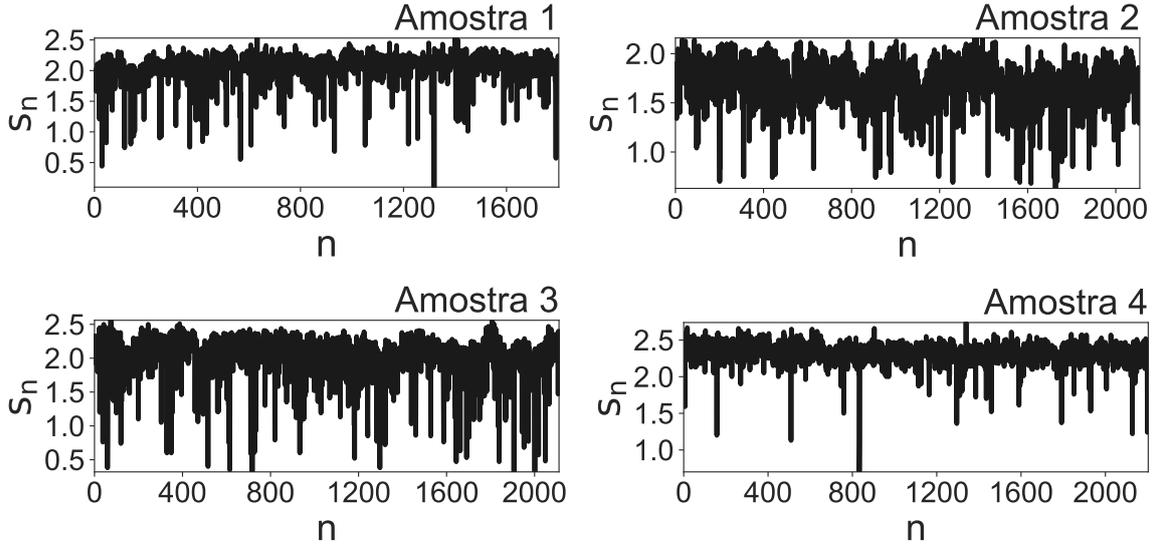


Figura 2.2: Sequência dos tamanhos de plantas sucessivas em cada uma das amostras,  $s_n$  versus  $n$ . O tamanho (altura)  $s$  é dado em metros.

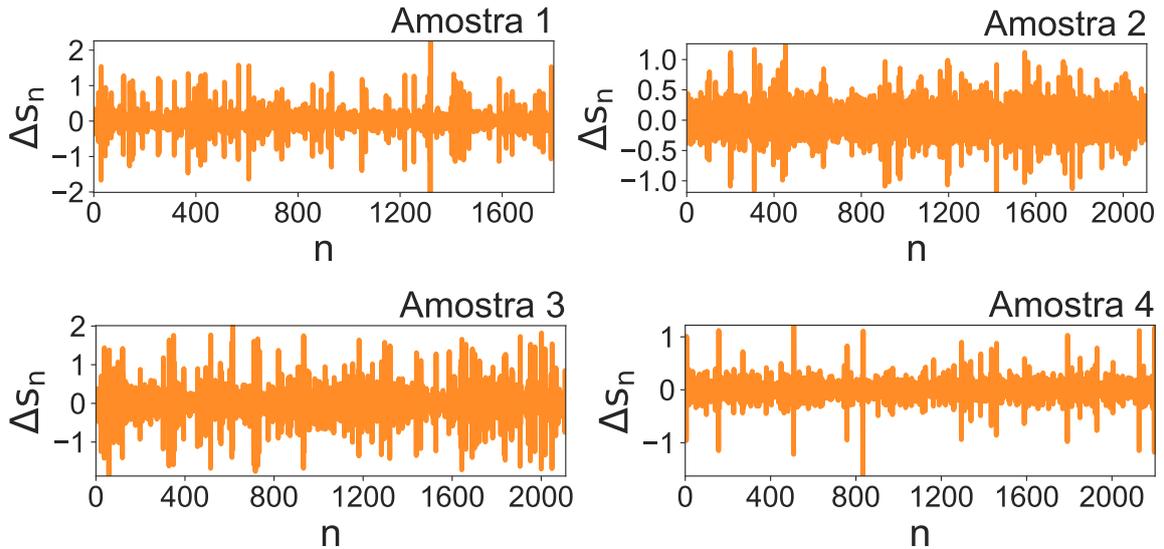


Figura 2.3: Sequência dos incrementos de tamanho entre plantas sucessivas,  $\Delta s_n$  versus  $n$ .

numéricos  $z_n$ ,  $n = 1, \dots, N_s - 1$ , são dados por

$$z_n = \begin{cases} 0, & \text{se } d'_n < h \\ 1, & \text{se } d'_n \geq h. \end{cases} \quad (2.5)$$

Aqui  $h$  é um valor limite para a diferenças de tamanho normalizada  $d'$ . O resultado  $z_n = 1$  indica que  $d'$  excedeu o valor limite  $h$  na posição  $n$ . Em outras palavras, dizemos que ocorreu um cruzamento de nível das diferenças de tamanho normalizadas (um evento). Por outro lado, o resultado  $z_n = 0$  indica que  $d'$  não excedeu  $h$  na posição

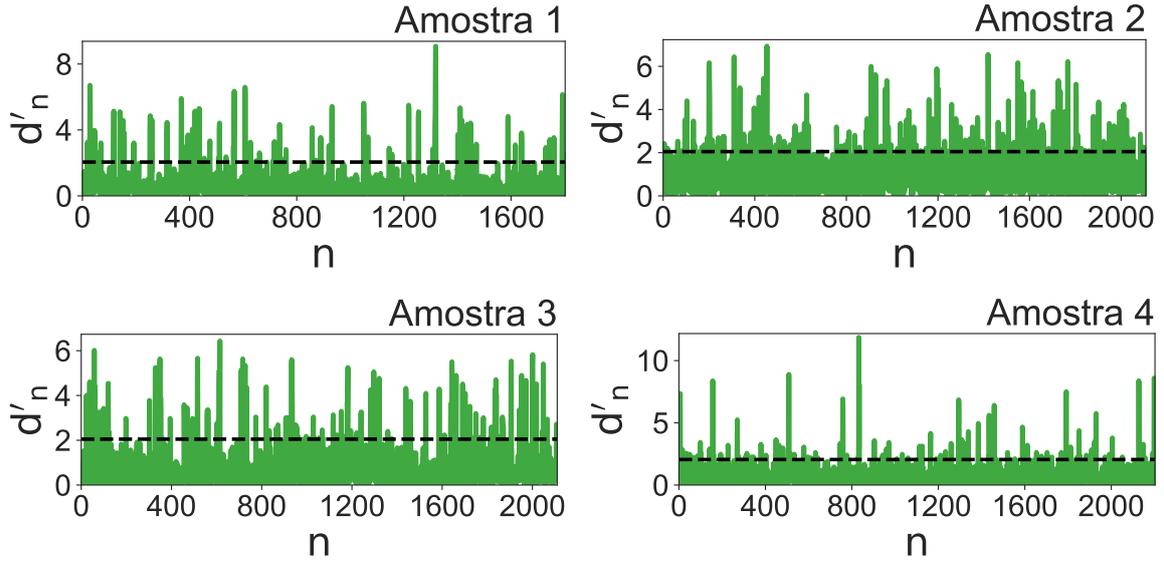


Figura 2.4: Sequência das diferenças de tamanho normalizadas,  $d'_n$  versus  $n$ . A linha tracejada representa o valor limite  $h = 2,05$ .



Figura 2.5: Ocorrência de eventos ao longo de cada amostra. As barras verticais marcam as posições em que  $d'_n > h$ , com  $h = 2,05$ . A distância horizontal entre barras verticais sucessivas representa o intervalo de retorno  $\lambda$ .

$n$  e, portanto, não houve cruzamento de nível. O conjunto de posições  $n$  em que  $d'$  excede o valor limite  $h$  é representado por  $n_l$ ,  $l = 0, \dots, L$ . Nessa notação,  $n_0$  marca a posição do primeiro evento da amostra e  $n_L$  marca a posição do último evento. Os intervalos de retorno das diferenças de tamanho normalizadas, para um dado valor de  $h$ , são definidos como

$$\lambda_l = n_l - n_{l-1}, \quad (2.6)$$

$l = 1, \dots, L$ . Note que essa definição de  $\lambda$  exclui os efeitos de borda. Mais especificamente, o intervalo entre a posição da primeira planta da amostra e a posição do

primeiro cruzamento de nível não entra na análise. Da mesma forma, o intervalo entre a posição do último cruzamento de nível e a posição da última planta da amostra também é desconsiderado. Essa definição de intervalo de retorno, excluindo os efeitos de borda, é comumente usada na literatura sobre sistemas complexos. A figura 2.5 mostra, para cada amostra, a sequência das posições dos eventos ao longo da fileira calculados para  $h = 2,05$  (em desvios-padrões). Definimos também os intervalos de retorno normalizados,

$$\lambda'_l = \frac{\lambda_l}{\sigma_\lambda}, \quad (2.7)$$

em que  $\sigma_\lambda$  é o desvio padrão dos  $\lambda_l$ 's.

## 2.2 Caracterização da variabilidade do tamanho de plantas em uma fileira

Nessa seção, descrevemos diversos índices para caracterizar a variabilidade do tamanho de plantas ao longo de fileiras ou linhas de plantio. Esses índices incluem medidas tradicionais comumente utilizadas na literatura para descrever a variabilidade do tamanho de plantas em monoculturas. Além disso, introduzimos índices baseados nos incrementos de tamanho entre plantas vizinhas, nos valores absolutos dos incrementos de tamanho (ou diferenças de tamanho) e nos intervalos de retorno entre diferenças de tamanho que excedem um certo valor limite.

### 2.2.1 Assimetria, coeficiente de variação, índice de Gini e coeficiente de autocorrelação

Dentre os índices mais utilizados para caracterizar a variabilidade de tamanho de plantas em monoculturas estão aqueles relacionados à distribuição de tamanho das plantas,  $P(s)$ . Esses índices incluem o coeficiente de variação e a assimetria calculados a partir dos tamanhos das plantas,  $s$ . Mais recentemente, o índice de Gini também tem sido utilizado para esse objetivo. No presente trabalho, vamos definir índices que se referem não apenas à distribuição dos tamanhos,  $P(s)$ , mas também à distribuição dos incrementos,  $P(\Delta s)$ , à distribuição das diferenças de tamanho,  $P(d)$ , e à distribuição dos intervalos de retorno,  $P(\lambda)$ . Inicialmente, para evitar definições repetitivas, vamos utilizar a seguinte notação:  $x_n$ ,  $n = 1, \dots, N$ , pode representar  $s_n$  (com  $N = N_s$ ),  $\Delta s_n$  (com  $N = N_s - 1$ ),  $d_n$  (com  $N = N_s - 1$ ) bem como valores normalizados dessas variáveis. Nessa notação,  $P(x)$  pode representar  $P(s)$ ,  $P(\Delta s)$ ,  $P(d)$  ou a distribuição dos valores normalizados correspondentes. Além disso, vamos utilizar coeficientes descritos pela  $k$ -estatística de Fisher para o cálculo dos momentos estatísticos das

distribuições [71]. A definição dos momentos  $k_2$ ,  $k_3$  e  $k_4$ , relativos à variável  $x$ , está descrita no apêndice A.

O coeficiente de variação, relativo à variável  $x$ , é definido como

$$CV_x = \frac{\sigma_x}{\bar{x}} \quad (2.8)$$

em que  $\bar{x}$  é o valor médio e  $\sigma_x = k_2^{1/2}$  é o desvio padrão dos  $x_n$ 's.  $CV_x = 0$  indica total igualdade — todos os valores  $x_n$  são iguais ( $\sigma_x = 0$ ). Se  $P(x)$  é uma distribuição exponencial ( $\sigma_x = \bar{x}$ ), então  $CV_x = 1$ . Se  $CV_x < 1$ , diz-se que  $P(x)$  é uma distribuição de baixa variância. Se  $CV_x > 1$ , diz-se que  $P(x)$  é uma distribuição de alta variância. A assimetria da amostra (também conhecida como  $g_1$ ) é definida como

$$Sk_x = \frac{k_3}{k_2^{3/2}}. \quad (2.9)$$

A assimetria  $Sk_x$  pode ser negativa, positiva ou nula. Distribuições simétricas apresentam  $Sk_x = 0$ , como ocorre, por exemplo, na distribuição normal (gaussiana).  $Sk_x > 0$  indica que a cauda direita da distribuição  $P(x)$  é mais longa que a esquerda.  $Sk_x < 0$  indica que a cauda esquerda da distribuição  $P(x)$  é mais longa que a direita. O excesso de curtose (ou simplesmente curtose) da amostra é definida como

$$K_x = \frac{k_4}{k_2^2} - 3. \quad (2.10)$$

Uma distribuição  $P(x)$  normal (gaussiana) é caracterizada por  $K_x = 0$  (distribuição mesocúrtica). Valores positivos da curtose,  $K_x > 0$ , podem indicar que  $P(x)$  é mais afunilada que a distribuição normal e com caudas mais longas (distribuição leptocúrtica). Valores negativos de curtose,  $K_x < 0$ , podem indicar que  $P(x)$  é mais achatada que uma distribuição normal e com caudas mais curtas (distribuição platicúrtica). O índice de Gini pode ser definido como

$$G_x = \frac{\sum_{i=1}^N \sum_{j=1}^N |x_i - x_j|}{2N^2\bar{x}}, \quad (2.11)$$

e assume valores no intervalo  $[0, 1]$ .  $G_x = 0$  indica igualdade perfeita ( $x_i = x_j$  para todo  $i$  e  $j$ ).  $G_x = 1$  indica máxima desigualdade entre os valores de  $x_n$ . Por exemplo, se  $x_1 = 1$  e  $x_n = 0$  para  $n = 2, \dots, N$ , temos  $G_x = 1$ . O índice de Gini também pode ser definido como uma razão de áreas associadas à curva de Lorenz. No apêndice B, apresentamos mais detalhes sobre a relação entre o índice de Gini e a curva de Lorenz. Exemplos gráficos do índice de Gini e outros índices comuns na literatura

de monoculturas foram apresentados na revisão e nos resultados. Visto que os índices definidos a seguir ainda não tinham sido empregados no contexto de plantas, eles serão ilustrados somente nos resultados.

Todos os índices definidos até agora nesse capítulo estão relacionados com a distribuição  $P(x)$  e, portanto, são insensíveis ao ordenamento sequencial dos valores  $x_n$ . Mudar aleatoriamente ou embaralhar a posição dos elementos  $x_n$  não muda a distribuição  $P(x)$ . O próximo índice está relacionado ao ordenamento sequencial dos valores  $x_n$  e é dado pelo coeficiente de autocorrelação de ordem 1 da série. Um estimador do coeficiente de autocorrelação de ordem 1 da variável  $x_n$  é definido como

$$r_x = \frac{1}{N-1} \sum_{n=1}^{N-1} \frac{(x_n - \mu_{x1})(x_{n+1} - \mu_{x2})}{\sigma_{x1}\sigma_{x2}}, \quad (2.12)$$

em que  $\mu_{x1}$  e  $\sigma_{x1}$  são o valor médio e o desvio padrão da sequência  $\{x_1, x_2, \dots, x_{N-1}\}$ ;  $\mu_{x2}$  e  $\sigma_{x2}$  são o valor médio e o desvio padrão da sequência  $\{x_2, x_3, \dots, x_N\}$ . Os valores de  $r_x$  se encontram no intervalo  $[-1, 1]$ .  $r_x$  próximo de zero indica que valores sucessivos de  $x_n$  não são correlacionados.  $r_x < 0$  indica que valores sucessivos de  $x_n$  são anti-correlacionados (correlações negativas) e  $r_x > 0$  indica que valores sucessivos de  $x_n$  são correlacionados (correlações positivas).

Os índices nesta seção estão definidos em termos dos valores genéricos  $x_n$ . Agora, vamos especificar alguns dos índices que utilizamos nesse trabalho. Os índices calculados a partir dos tamanhos  $s$  são os seguintes:  $CV_s$  (coeficiente de variação),  $G_s$  (coeficiente de Gini),  $Sk_s$  (assimetria) e  $r_s$  (coeficiente de autocorrelação). Os índices baseados nos incrementos de tamanho  $\Delta s$  são:  $Sk_{\Delta s}$  (assimetria) e  $K_{\Delta s}$  (excesso de curtose). Por fim, os índices baseados nas diferenças de tamanho  $d$  são:  $CV_d$  (coeficiente de variação),  $G_d$  (coeficiente Gini),  $Sk_d$  (assimetria) e  $r_d$  (coeficiente de autocorrelação).

## 2.2.2 Variação local, *burstiness* e memória

Nesta seção, vamos definir índices baseados nos intervalos de retorno das diferenças de tamanho,  $\lambda$ . A variação local é definida como [58, 59]

$$LV_\lambda = \frac{3}{(L-1)} \sum_{i=1}^{L-1} \frac{(\lambda_i - \lambda_{i+1})^2}{(\lambda_i + \lambda_{i+1})^2}. \quad (2.13)$$

$LV_\lambda$  pode assumir valores no intervalo  $[0, 3]$ .  $LV_\lambda = 0$  indica que os eventos ocorrem em perfeita regularidade ou periodicidade ( $\lambda_l = \lambda_{l+1}$ , para todo  $l$ ).  $LV_\lambda = 1$  indica que os eventos ocorrem aleatoriamente, como em processos de Poisson em que a distribuição  $P(\lambda)$  é exponencial e  $\lambda$  é uma variável não-correlacionada (veja o apêndice D).  $LV_\lambda$  está

relacionada a dois fatores: a distribuição  $P(\lambda)$  e às correlações entre valores sucessivos de  $\lambda$ . O coeficiente *burstiness* é definido como [60, 72]

$$B_\lambda = \frac{\sigma_\lambda - \bar{\lambda}}{\sigma_\lambda + \bar{\lambda}} = \frac{CV_\lambda - 1}{CV_\lambda + 1}, \quad (2.14)$$

em que  $\sigma_\lambda$  é o desvio padrão e  $\bar{\lambda}$  é a média dos  $\lambda_i$ 's. Note que o coeficiente *burstiness* pode ser escrito em termos do coeficiente de variação da variável  $\lambda_i$ ,  $CV_\lambda$ .  $B_\lambda$  pode assumir valores no intervalo  $[-1, 1]$ .  $B_\lambda = -1$  ( $\sigma_\lambda = 0$ ) indica que os eventos são totalmente regulares ou periódicos.  $B_\lambda = 0$  ( $\sigma_\lambda = \bar{\lambda}$ ) indica que os eventos ocorrem aleatoriamente, como em um processo de Poisson.  $B_\lambda > 0$  ( $\sigma_\lambda > \bar{\lambda}$ ) indica que  $P(\lambda)$  decai mais lentamente que uma exponencial, indicando uma probabilidade maior de grandes valores de  $\lambda$  (esse comportamento tem sido chamado de *bursty behavior*). Uma definição alternativa para o coeficiente *burstiness*, que leva em conta efeitos de tamanho finito, é dada por [72]

$$BC_\lambda = \frac{\sqrt{L+1}CV_\lambda - \sqrt{L-1}}{(\sqrt{L+1}-2)CV_\lambda + \sqrt{L-1}}, \quad (2.15)$$

em que  $L$  é o número total de eventos. Aqui, vamos chamar  $BC_\lambda$  de *burstiness* corrigido. O coeficiente memória é definido como [60]

$$M_\lambda = \frac{1}{L-1} \sum_{i=1}^{L-1} \frac{(\lambda_i - \mu_{\lambda_1})(\lambda_{i+1} - \mu_{\lambda_2})}{\sigma_{\lambda_1}\sigma_{\lambda_2}} \quad (2.16)$$

em que  $\mu_{\lambda_1}$  e  $\sigma_{\lambda_1}$  são o valor médio e o desvio padrão da sequência  $\{\lambda_1, \lambda_2, \dots, \lambda_{L-1}\}$ ;  $\mu_{\lambda_2}$  e  $\sigma_{\lambda_2}$  são o valor médio e o desvio padrão da sequência  $\{\lambda_2, \lambda_3, \dots, \lambda_L\}$ . Note que  $M_\lambda$  corresponde a um estimador do coeficiente de auto-correlação de ordem 1 para a variável  $\lambda_i$ .  $M_\lambda$  pode assumir valores no intervalo  $[-1, 1]$ .  $M_\lambda = 0$  indica que a variável  $\lambda_i$  não é correlacionada, como em um processo de Poisson.  $M_\lambda < 0$  indica a presença de anti-correlações em  $\lambda_i$ . Isso significa que intervalos de retorno longos tendem a ser seguidos por intervalos de retorno curtos e intervalos de retorno curtos tendem a ser seguidos por intervalos de retorno longos.  $M_\lambda > 0$  indica a presença de correlações positivas em  $\lambda_i$ . Intervalos de retorno longos tendem a ser seguidos por intervalos de retorno longos e intervalos de retorno curtos tendem a ser seguidos por intervalos de retorno curtos.

## 2.3 Correlações entre pares de índices

Para investigar a presença de possíveis correlações entre pares de índices, utilizamos o coeficiente de correlação de Pearson,  $r$ . No apêndice [C](#) definimos o coeficiente  $r$  e discutimos sua interpretação. Essa análise pode ajudar na seleção de índices não redundantes para a caracterização de diferentes aspectos do sistema em estudo.

# Capítulo 3

## Análise dos dados e resultados

Neste capítulo, descrevemos os principais resultados obtidos a partir da análise dos dados descritos no capítulo anterior, para cada uma das quatro amostras independentes. Iniciamos calculando as distribuições empíricas dos tamanhos, dos incrementos de tamanho, das diferenças de tamanho e dos intervalos de retorno das diferenças de tamanho. Posteriormente, destacamos aspectos qualitativos dessas distribuições. A seguir, calculamos sistematicamente os valores empíricos dos índices definidos no capítulo anterior. Por fim, investigamos possíveis correlações entre pares de índices.

### 3.1 Distribuições, correlações e outras propriedades estatísticas

Na figura 3.1, vemos a distribuição dos tamanhos das plantas,  $P(s)$ , calculada para cada uma das amostras. Visualmente, a distribuição  $P(s)$  parece ser assimétrica com uma cauda mais longa voltada para a esquerda. Isso indicaria a presença de plantas com tamanho bem abaixo da média e a ausência de plantas com tamanho muito acima da média. Na figura 3.2, vemos a distribuição dos incrementos normalizados,  $P(\Delta s')$ . Nessa figura,  $P(\Delta s')$  é comparada com uma distribuição normal com média zero e variância unitária. Visualmente,  $P(\Delta s')$  parece apresentar comportamento não-normal ou não-gaussiano para as quatro amostras. Entretanto, o comportamento não gaussiano parece ser mais perceptível nas amostras 1 e 3. Nas amostras 2 e 4, o comportamento não-gaussiano pode ser percebido principalmente nas caudas da distribuição. Na figura 3.3, vemos a distribuição das diferenças de tamanho normalizadas,  $P(d')$ , calculada para cada uma das amostras. A distribuição  $P(d')$  parece ser assimétrica com uma cauda mais longa à direita. Para comparação, vemos também uma

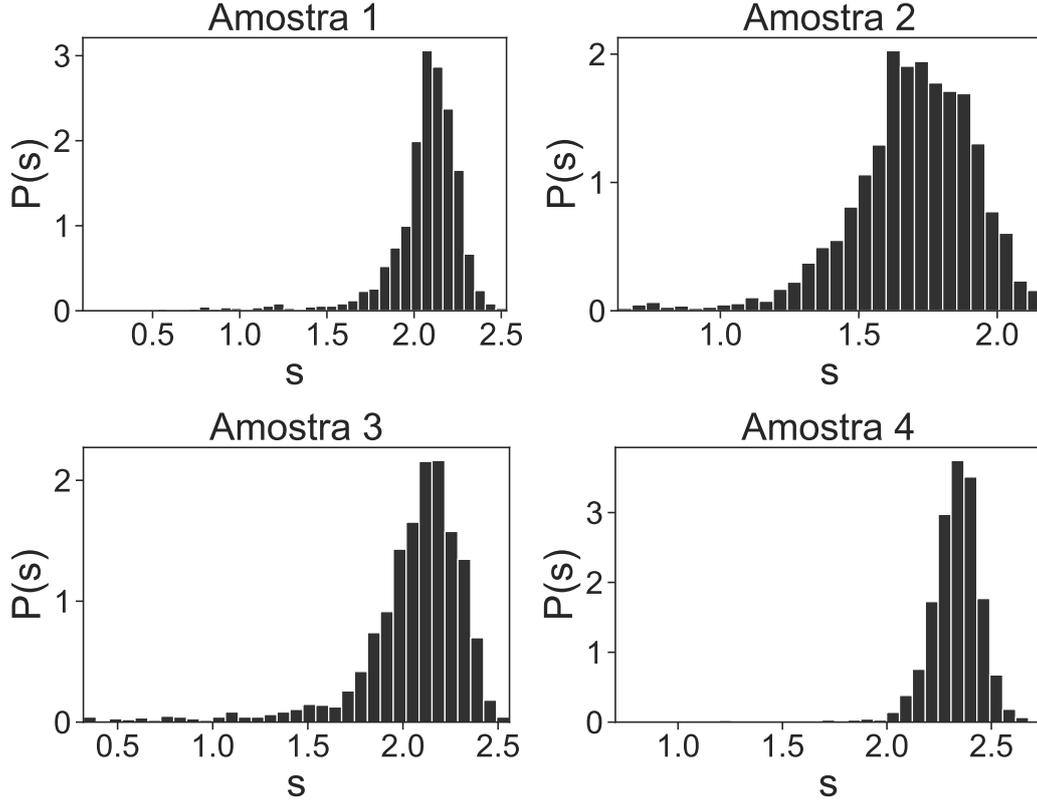


Figura 3.1: Distribuição do tamanho de plantas,  $P(s)$ , para cada uma das amostras. O tamanho  $s$  é dado em metros.

distribuição log-normal,

$$P(d') = \frac{1}{d' \sigma \sqrt{2\pi}} \exp\left(-\frac{(\ln d' - \mu)^2}{2\sigma^2}\right). \quad (3.1)$$

Os parâmetros  $\mu$  e  $\sigma$  da distribuição log-normal que aparece na figura 3.3, obtidos pelo método da máxima verossimilhança, são os seguintes:  $\mu = -0.82$ ,  $\sigma = 1.09$  (amostra 1);  $\mu = -0.27$ ,  $\sigma = 1.00$  (amostra 2);  $\mu = -0.80$ ,  $\sigma = 1.16$  (amostra 3);  $\mu = -0.51$ ,  $\sigma = 0.98$  (amostra 4). Na figura 3.4, vemos a distribuição dos logaritmos de  $d'$ , normalizados para que tenham média zero e variância unitária. Note que as distribuições empíricas das 4 amostras aproximadamente colapsam em uma curva — uma distribuição normal com média zero e variância unitária. Esses resultados sugerem que  $P(d')$  exibe um comportamento que pode ser aproximadamente descrito por uma distribuição log-normal. Por fim, na figura 3.5, vemos a distribuição acumulada dos intervalos de retorno normalizados,  $P_c(\lambda')$ , calculada para cada uma das amostras. Comparamos as distribuições empíricas com uma distribuição exponencial acumulada, com média e variância unitárias. Na maioria dos casos, é possível ver que  $P_c(\lambda')$  desvia do comportamento exponencial.

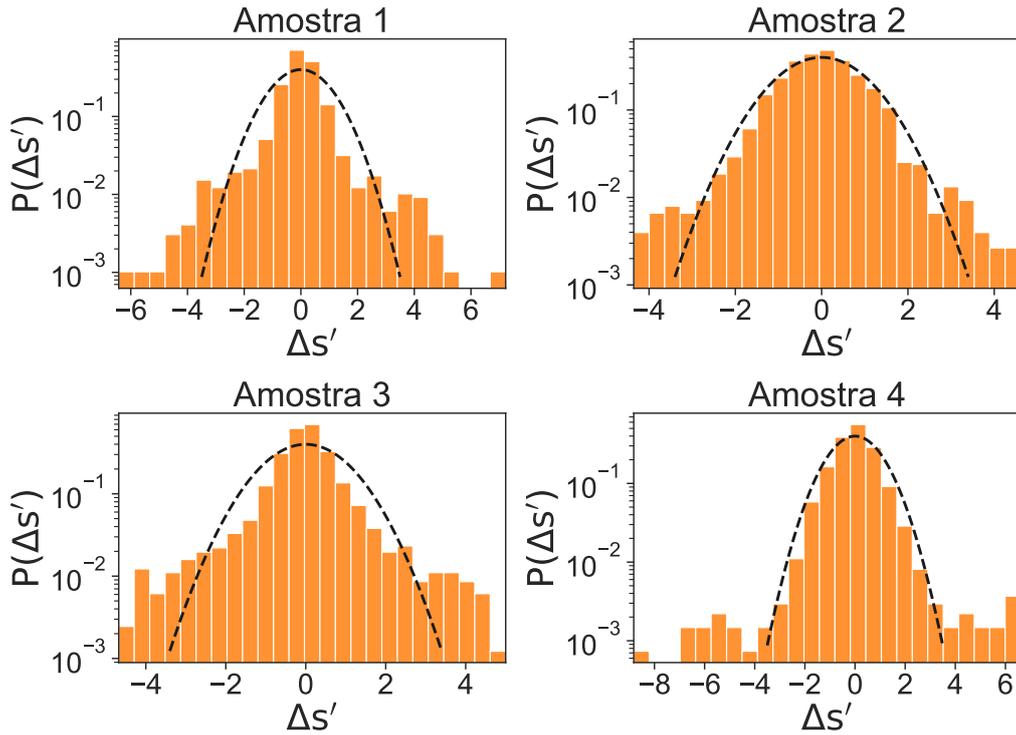


Figura 3.2: Distribuição dos incrementos normalizados,  $P(\Delta s')$ . As linhas tracejadas representam distribuições normais com média zero e variância unitária.

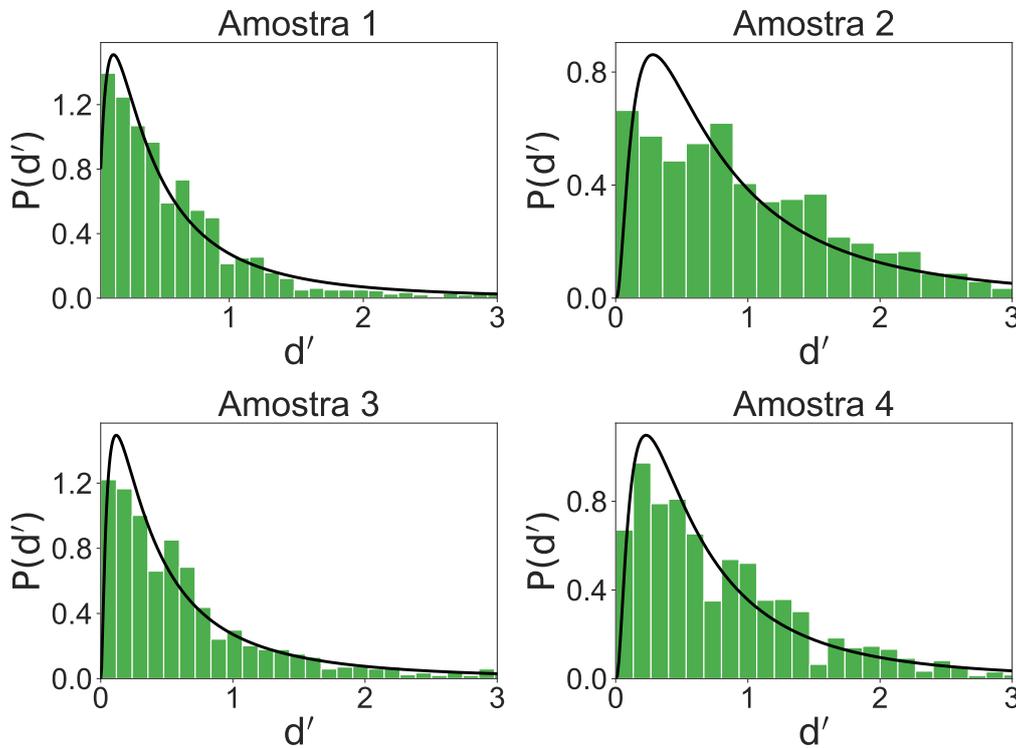


Figura 3.3: Distribuição das diferenças de tamanho normalizadas,  $P(d')$ . As linhas representam log-normais ajustadas aos dados pelo método da máxima verossimilhança.

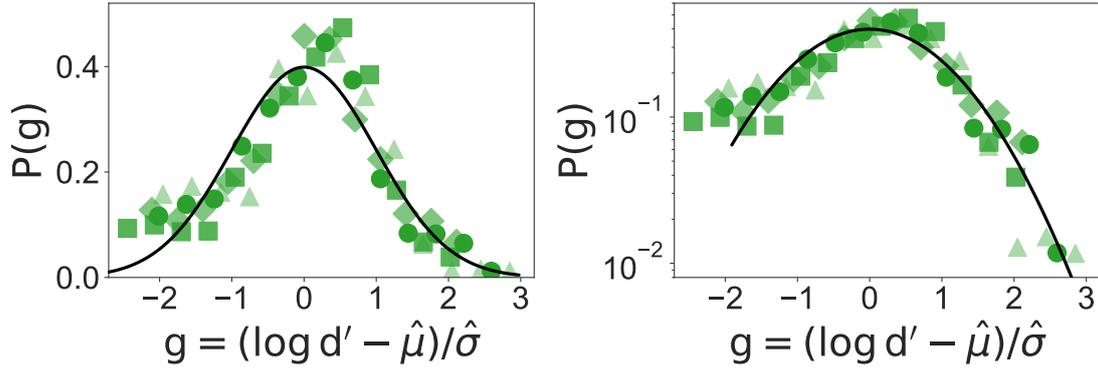


Figura 3.4: Distribuição  $P(g)$ , com  $g = (\ln d' - \hat{\mu})/\hat{\sigma}$ . Aqui,  $\hat{\mu}$  é o valor médio de  $\ln d'$  e  $\hat{\sigma}$  é o desvio padrão de  $\ln d'$ . Os dados referem-se às amostras: 1 (círculos); 2 (quadrados); 3 (diamantes); e 4 (triângulos). As linhas sólidas representam uma distribuição normal com média zero e variância unitária.

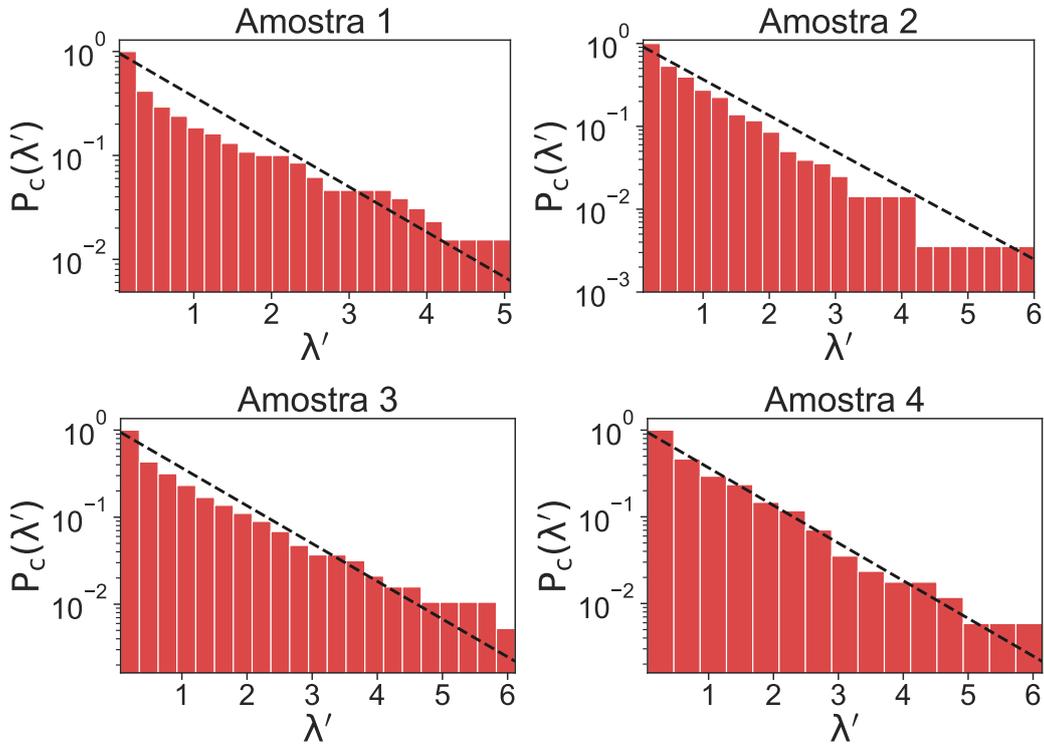


Figura 3.5: Distribuição acumulada dos intervalos de retorno normalizados,  $P_c(\lambda')$ . As linhas tracejadas representam uma distribuição acumulada exponencial com variância unitária,  $\exp(-x)$ .

A seguir, passamos para a análise dos índices, alguns deles relacionados às distribuições de probabilidade que acabamos de mencionar. As figuras 3.6 - 3.9 apresentam os valores obtidos de todos os índices definidos no capítulo anterior, calculados para cada uma das amostras. A partir de agora, vamos discutir esses resultados por grupo de índices, começando com os índices tradicionais, calculados a partir dos tamanhos

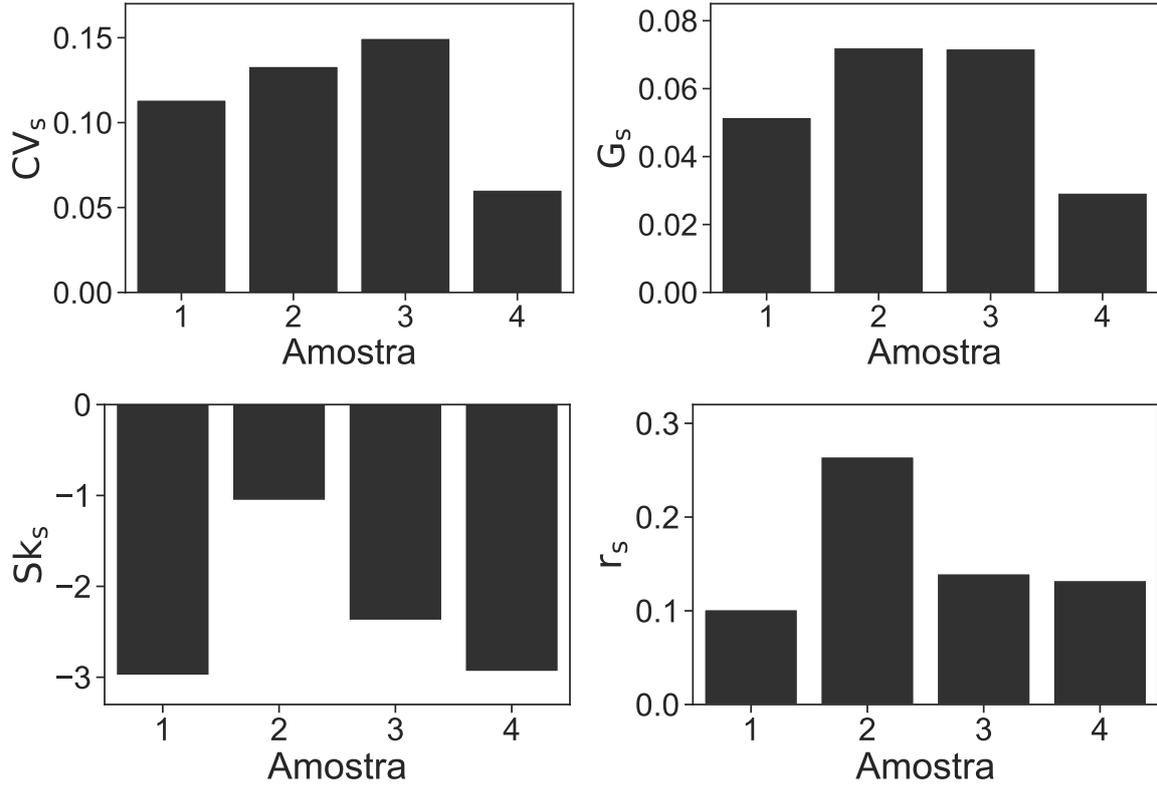


Figura 3.6: Coeficiente de variação  $CV_s$ , índice de Gini  $G_s$ , assimetria  $Sk_s$  e coeficiente de autocorrelação  $r_s$  calculados para a sequência dos tamanhos das plantas. Os dados empíricos referem-se às amostras 1, 2, 3 e 4.

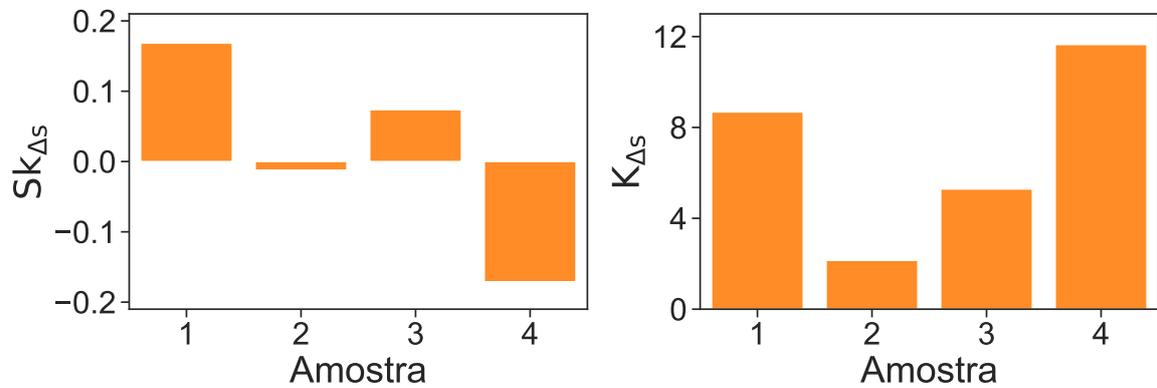


Figura 3.7: Assimetria  $Sk_{\Delta s}$  e curtose  $K_{\Delta s}$  calculadas para as diferenças de tamanho entre plantas vizinhas. Os dados empíricos referem-se às amostras 1, 2, 3 e 4.

das plantas. A figura 3.6 apresenta os valores empíricos obtidos para os índices  $CV_s$ ,  $G_s$ ,  $Sk_s$  e  $r_s$  para cada uma das amostras. Os valores obtidos para o coeficiente de variação,  $CV_s$ , estão no intervalo entre 0.06 e 0.15 (com valor médio de 0.11). Esses valores são menores do que 1, indicando que o desvio padrão dos tamanhos é menor que o valor médio dos tamanhos. Podemos dizer então que  $P(s)$  é uma distribuição de baixa variância. Com respeito ao índice de Gini,  $G_s$ , os valores obtidos estão no inter-

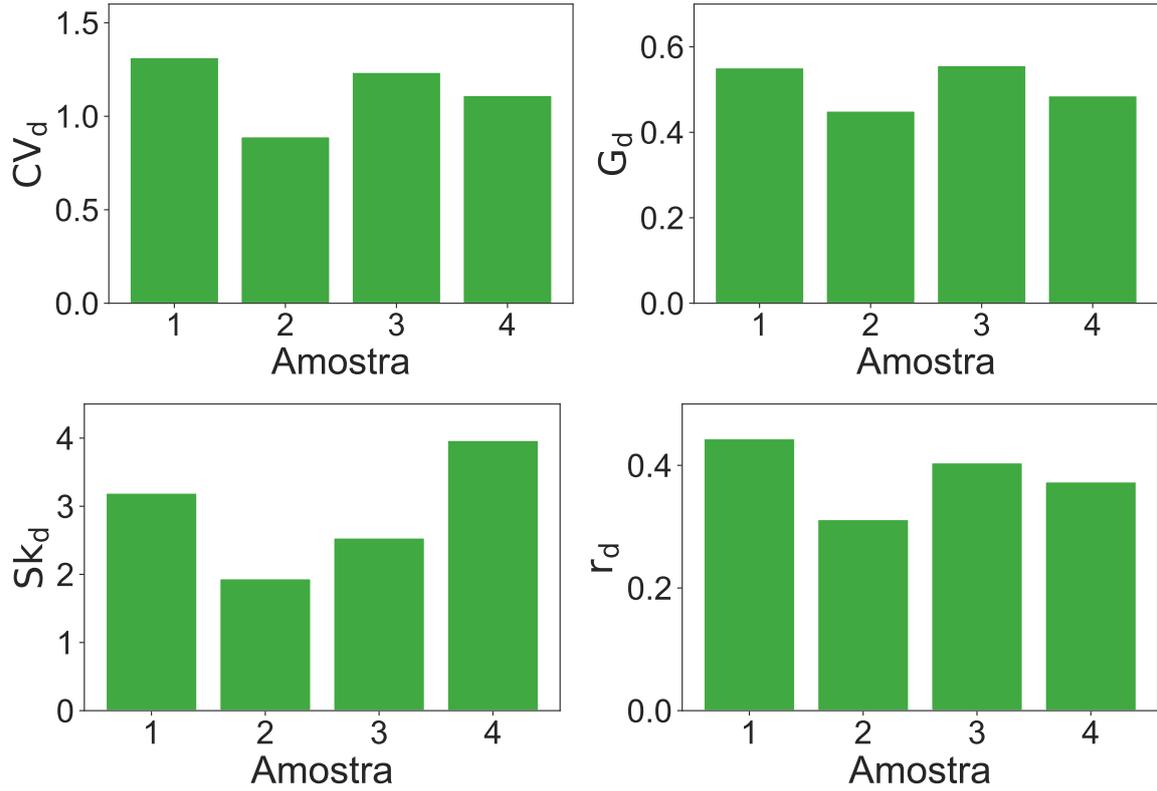


Figura 3.8: Coeficiente de variação  $CV_v$ , índice de Gini  $G_v$ , assimetria  $Sk_v$  e autocorrelação  $r_v$  para as "volatilidades" normalizadas  $v_n$ . Os dados empíricos referem-se às amostras 1, 2, 3 e 4.

valo de 0.03 a 0.07 (com valor médio de 0.06). Esses valores são relativamente baixos, indicando uma baixa desigualdade de tamanho entre as plantas. As curvas de Lorentz associadas a  $G_s$  são apresentadas no apêndice B. No caso da assimetria,  $Sk_s$ , os valores obtidos estão no intervalo de  $-2.98$  a  $-1.06$  (com valor médio de  $-2.34$ ). Esses valores negativos indicam que  $P(s)$  é uma distribuição assimétrica, com a cauda esquerda mais alongada que a cauda direita. Isso reforça que as amostras contém plantas bem abaixo do tamanho médio mas não contém plantas muito maiores do que o tamanho médio (veja também a figura 3.1). Os valores obtidos para o coeficiente de autocorrelação,  $r_s$ , estão no intervalo de 0.10 e 0.26 (com valor médio de 0.16). Esse resultado ( $r_s > 0$ ) indica uma autocorrelação positiva: plantas altas (baixas) tendem a ser seguidas por plantas altas (baixas). Para verificar esse último resultado, repetimos a análise para várias sequências embaralhadas dos dados originais. O embaralhamento da série dos tamanhos consiste em colocar os valores de  $s$  em ordem aleatória. Calculamos  $r_s$  para  $10^3$  versões aleatórias dos dados originais e obtivemos valores no intervalo de  $-0.02$  a 0.02 (com valor médio de 0.00). Essa análise reforça o resultado obtido anteriormente: os valores de  $r_s$  obtidos indicam uma correlação positiva nos valores de  $s$ .

A figura 3.7 mostra os valores obtidos de  $Sk_{\Delta_s}$  e  $K_{\Delta_s}$  para cada uma das amostras.

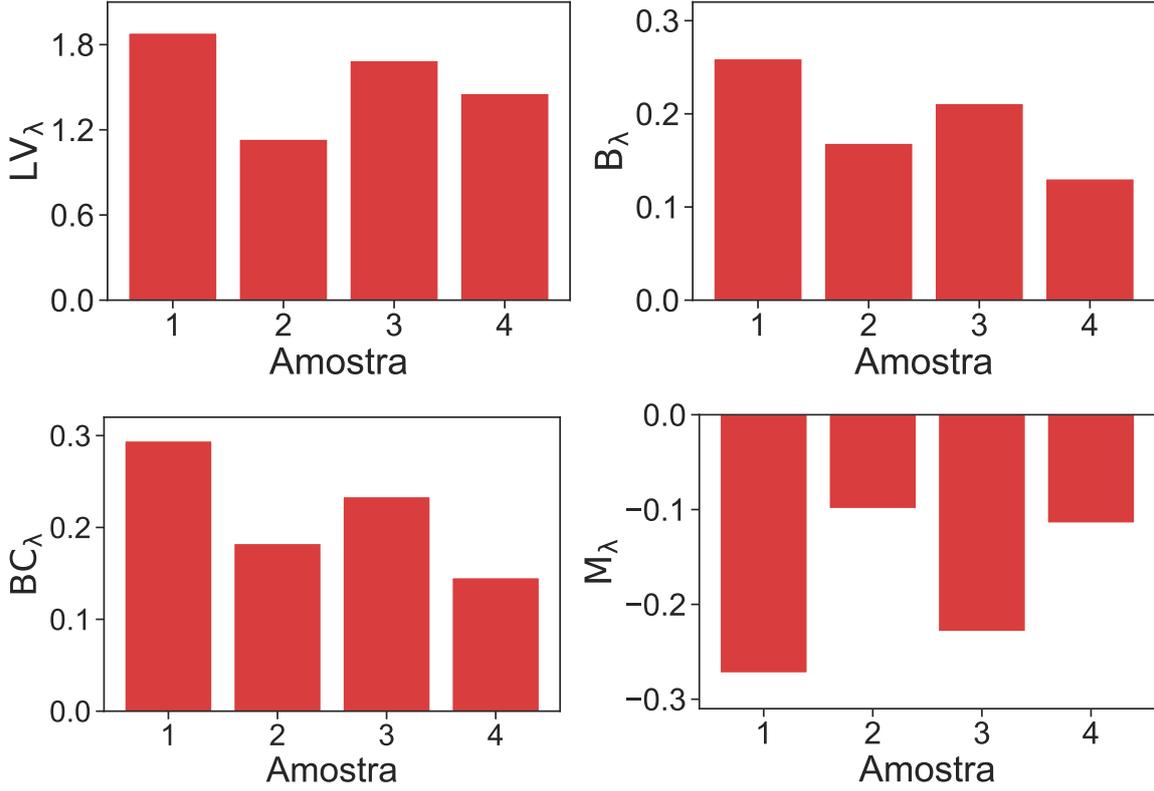


Figura 3.9: Variação local  $LV_\lambda$ , *burstiness*  $B_\lambda$ , *burstiness* corrigido  $BC'_\lambda$  e memória  $M_\lambda$ , calculados a partir dos intervalos de retorno  $\lambda$  (valor limite  $h = 2.05$ ), para cada uma das amostras.

Para a assimetria,  $Sk_{\Delta s}$ , encontramos valores no intervalo entre  $-0.17$  e  $0.17$  (com valor médio de  $0.01$ ). Os valores de  $Sk_{\Delta s}$  flutuam em torno do zero, sugerindo que a distribuição dos incrementos tende a ser aproximadamente simétrica. No caso do excesso de curtose,  $K_{\Delta s}$ , os valores obtidos estão no intervalo entre  $2.15$  e  $11.65$  (valor médio  $6.94$ ). Esses valores de  $K_{\Delta s}$  maiores que zero indicam que a  $P(\Delta s)$  é mais afunilada e/ou possui caudas mais longas que a distribuição normal. Dizemos que  $P(\Delta s)$  é leptocúrtica (veja também a figura 3.2).

A figura 3.8 apresenta os resultados obtidos para os índices  $CV_d$ ,  $Sk_d$ ,  $G_d$  e  $r_d$ , calculados a partir das diferenças de tamanho entre plantas vizinhas,  $d$ , para cada uma das amostras. Para o coeficiente de variação  $CV_d$ , os valores obtidos estão no intervalo entre  $0.89$  e  $1.31$  (com valor médio de  $1.14$ ). Note que os valores de  $CV_d$  flutuam em torno da unidade, indicando que o valor médio e o desvio padrão das diferenças de tamanho são aproximadamente iguais. Para o índice de Gini,  $G_d$ , os valores obtidos estão entre  $0.45$  e  $0.56$  (valor médio de  $0.51$ ). Esse resultado indica uma desigualdade considerável nas diferenças de tamanho entre plantas vizinhas. As curvas de Lorentz associadas a  $G_d$  são apresentadas no apêndice B. Para a assimetria,  $Sk_d$ , os valores obtidos estão entre  $1.94$  e  $3.97$  com valor médio de  $2.91$ . O resultado

$Sk_d > 0$  indica que  $P(d)$  é uma distribuição assimétrica, com a cauda direita mais longa que a esquerda (veja também a figura 3.3). Para o coeficiente de autocorrelação,  $r_d$ , os valores obtidos estão entre 0.31 e 0.44, sendo a média igual a 0.38. Esses valores são positivos, indicando uma autocorreção positiva na série das diferenças. Para verificar esse resultado, repetimos o cálculo para  $10^3$  sequências embaralhadas dos valores de  $d$  e obtivemos valores no intervalo de  $-0.02$  a  $0.02$  (com valor médio de 0.00). Essa análise reforça que há uma correlação positiva nos valores das diferenças de tamanho. Isso indica que diferenças grandes (pequenas) tendem a ser seguidas por diferenças grandes (pequenas).

A figura 3.9 mostra os valores dos índices  $LV_\lambda$ ,  $B_\lambda$ ,  $BC_\lambda$  e  $M_\lambda$ , referentes aos intervalos de retorno  $\lambda$ , com valor limite  $h = 2.05$ . O papel de  $h$  nos resultados vai ser considerado na próxima seção. Os valores obtidos para a variação local,  $LV_\lambda$ , estão no intervalo de 1.13 a 1.88, com valor médio de 1.54. Esse resultado ( $LV_\lambda > 1$ ) indica que os cruzamentos de nível da volatilidade não ocorrem aleatoriamente ao longo da fileira. Para o coeficiente *burstiness*,  $B_\lambda$ , os valores obtidos estão no intervalo entre 0.13 e 0.26, com valor médio de 0.19. Para o *burstiness* corrigido,  $BC_\lambda$ , os valores empíricos estão no intervalo de 0.15 a 0.29, com valor médio de 0.21. Valores de  $B_\lambda > 0$  (ou  $BC_\lambda > 0$ ) indicam que  $P(\lambda)$  decai mais lentamente que uma distribuição exponencial (veja também a figura 3.5). Os valores obtidos para a memória,  $M_\lambda$ , estão no intervalo de  $-0.10$  a  $-0.27$ , com valor médio de  $-0.18$ . Valores negativos do coeficiente memória ( $M_\lambda < 0$ ) indicam a presença de anti-correlações. Esse resultado sugere que intervalos de retorno longos tendem a ser seguidos por intervalos de retorno curtos e intervalos curtos tendem a ser seguidos por intervalos longos.

Para verificar comparativamente os resultados apresentados na figura 3.9, calculamos os índices  $LV_\lambda$ ,  $B_\lambda$ ,  $BC_\lambda$  e  $M_\lambda$  para versões embaralhadas do sinal binário dado por  $z_n$ . Como vimos,  $z_n = 1$  localiza os cruzamentos de nível das diferenças de tamanho que excedem o valor limite  $h$ . O embaralhamento dos valores  $z_n$  corresponde a colocar os 0's e 1's em posições aleatórias ao longo da sequência. A partir do sinal binário embaralhado, calculamos novamente os intervalos de retorno usando a eq. (2.6). Vamos chamar esses intervalos de retorno de  $\lambda^{emb}$ . A seguir, calculamos os 4 índices definidos nas equações 2.13 a 2.16 usando  $\lambda^{emb}$  no lugar de  $\lambda$ . Para cada índice e cada amostra, repetimos esse procedimento  $10^3$  vezes e tomamos o valor médio do valor do índice. Valores médios típicos obtidos para  $LV_\lambda^{emb}$  estão no intervalo de 0.79 a 0.87, com valor médio calculado sobre as amostras de 0.84. Esse valor é menor que o valor médio de 1.54 obtido para as sequências originais. Valores típicos obtidos para  $B_\lambda^{emb}$  e  $BC_\lambda^{emb}$  estão no intervalo de  $-0.04$  a  $-0.02$ , com valor médio de  $-0.03$ . Esses valores são próximos de zero e menores que os valores médios de 0.19 e 0.22, respectivamente, obtidos para as sequências originais. Por sua vez, valores típicos obtidos para  $M_\lambda^{emb}$

estão no intervalo de  $-0.007$  a  $-0.003$ , com valor médio de  $-0.006$ . Em módulo, esse valor é menor que o valor médio de  $-0.18$  obtido para as sequências originais. Essa análise reforça que as sequências originais de eventos exibem padrões não triviais, diferentes do que seria esperado se os eventos ocorressem aleatoriamente.

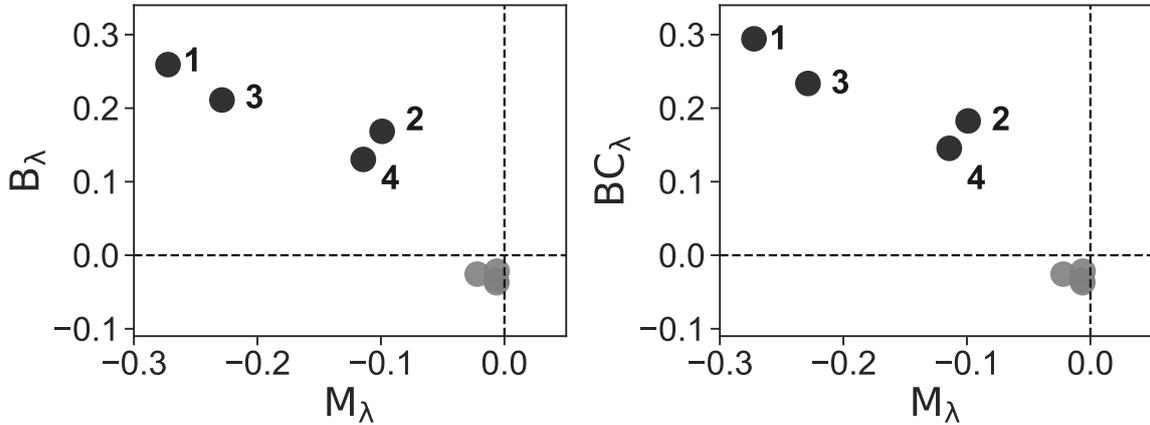


Figura 3.10: Diagrama  $MB$ . À esquerda: *Burstiness*  $B_\lambda$  versus memória  $M_\lambda$  para as amostras 1, 2, 3 e 4. À direita: *Burstiness* corrigido  $BC_\lambda$  versus  $M_\lambda$ . Próximo à origem (em cinza), mostramos os mesmos pontos calculados a partir das sequências de eventos embaralhadas.

Na análise das séries embaralhadas descrita no parágrafo anterior, os valores médios da variação local, do coeficiente *burstiness* e do coeficiente memória foram  $0.84$ ,  $-0.03$  e  $-0.006$ , respectivamente. Para processos de Poisson homogêneos, os valores esperados para esses índices são  $1$ ,  $0$  e  $0$ , respectivamente. Essas diferenças entre os resultados obtidos com as séries embaralhadas e os resultados esperados para um processo de Poisson ocorrem devido à natureza discreta e finita dos dados (efeitos de tamanho finito). Note também que a memória das séries embaralhadas fica bem próxima de zero. Isso sugere que esse índice é mais robusto em relação a efeitos de tamanho finito. Por outro lado, a variação local das séries embaralhadas fica relativamente distante de  $1$ , sugerindo que esse índice é mais sensível a efeitos de tamanho finito.

Na figura 3.10, apresentamos os valores de  $M_\lambda$  em função de  $B_\lambda$  para cada uma das amostras. Mostramos também os valores de  $M_\lambda$  em função de  $BC_\lambda$ . Além disso, os mesmos coeficientes, calculados a partir do sinal binário embaralhado, também são mostrados na figura. Fica nítida a diferença entre o padrão empírico que ocorre nas fileiras de milho e o comportamento esperado se os eventos fossem aleatórios. Note também que os pontos empíricos na figura estão no segundo quadrante, em que o coeficiente *burstiness* é positivo e o coeficiente memória é negativo. Esse tipo de gráfico, conhecido como diagrama  $MB$ , permite visualizar os resultados de forma concisa e facilitam a comparação com outros sistemas complexos.

Em resumo, a tabela 3.1 apresenta os valores obtidos de todos os índices conside-

Tabela 3.1: Breve descrição do conjunto de índices utilizados no presente trabalho. A tabela inclui os valores obtidos para cada amostra (1, 2, 3 e 4) e o valor médio calculado sobre as amostras.

Índice	Descrição	Variável	Valores (1, 2, 3 e 4)				Média
$CV_s$	coeficiente de variação	$s$	0.11	0.13	0.15	0.06	0.11
$G_s$	Gini	$s$	0.05	0.07	0.07	0.03	0.06
$Sk_s$	assimetria	$s$	-2.98	-1.06	-2.37	-2.94	-2.34
$r_s$	coeficiente de correlação	$s$	0.10	0.26	0.14	0.13	0.16
$Sk_{\Delta s}$	assimetria	$\Delta s$	0.17	-0.01	0.07	-0.17	0.01
$K_{\Delta s}$	curtose	$\Delta s$	8.68	2.15	5.29	11.65	6.94
$CV_d$	coeficiente de variação	$d$	1.31	0.89	1.24	1.11	1.14
$G_d$	Gini	$d$	0.55	0.45	0.56	0.49	0.51
$Sk_d$	assimetria	$d$	3.19	1.94	2.54	3.97	2.91
$r_d$	coeficiente de correlação	$d$	0.44	0.31	0.40	0.37	0.38
$LV_\lambda$	variação local	$\lambda$	1.88	1.13	1.69	1.46	1.54
$B_\lambda$	<i>burstiness</i>	$\lambda$	0.26	0.17	0.21	0.13	0.19
$BC_\lambda$	<i>burstiness</i> corrigido	$\lambda$	0.29	0.18	0.23	0.15	0.21
$M_\lambda$	memória	$\lambda$	-0.27	-0.10	-0.23	-0.11	-0.18

rados nesse capítulo, para cada uma das amostras, incluindo o valor médio de cada índice calculado sobre as amostras.

### 3.2 Análise das correlações entre pares de índices

A seguir, fazemos uma análise das correlações entre pares de índices. Para investigar a intensidade e a direção da associação entre pares de índices, utilizamos o coeficiente de correlação de Pearson,  $r$  (veja o apêndice C). Na figura 3.11, apresentamos os valores obtidos de  $r$  para todos os pares de índices. Além disso, vemos na figura os  $p$ -valores associados a cada par de índices, calculados a partir de um teste estatístico paramétrico, o teste  $t$  de *student*. Quando o  $p$ -valor é menor ou igual a 0,05, dizemos que a associação linear entre o par de índices é estatisticamente significativa. Temos ao todo 91 pares de índices representados na figura 3.11 (excluindo-se as combinações de cada índice com ele mesmo). Desses 91 pares, apenas 13 apresentam associações estatisticamente significativas (esse resultado pode estar relacionado ao pequeno número de amostras independentes). De modo geral, independentemente do

coeficiente  $r$  ser ou não estatisticamente significativo, 32 pares de índices apresentam  $|r| \geq 0.8$  (cerca de 35% dos pares), sugerindo associações fortes. Por outro lado, 34 pares de índices apresentam  $|r| \leq 0.5$  (cerca de 37% dos pares), sugerindo associações moderadas a fracas.

Na figura 3.12, apresentamos os gráficos de dispersão dos pares de índices que apresentaram associações estatisticamente significativas ( $p \leq 0.05$ ). Esses pares são os seguintes:  $Sk_s$  e  $r_s$ , com  $r = 0.97$  e  $p = 0.03$ ;  $CV_s$  e  $G_s$ , com  $r = 0.97$  e  $p = 0.03$ ;  $G_s$  e  $Sk_d$ , com  $r = -0.96$  e  $p = 0.04$ ;  $Sk_{\Delta_s}$  e  $B_\lambda$ , com  $r = 0.98$  e  $p = 0.02$ ;  $Sk_{\Delta_s}$  e  $BC_\lambda$ , com  $r = 0.97$  e  $p = 0.03$ ;  $K_{\Delta_s}$  e  $Sk_d$ , com  $r = 1.00$  e  $p = 0.003$ ;  $CV_d$  e  $G_d$ , com  $r = 0.95$  e  $p = 0.05$ ;  $CV_d$  e  $r_d$ , com  $r = 0.99$  e  $p = 0.01$ ;  $CV_d$  e  $LV_\lambda$ , com  $r = 0.99$  e  $p = 0.01$ ;  $G_d$  e  $LV_\lambda$ , com  $r = 0.95$  e  $p = 0.05$ ;  $G_d$  e  $M_\lambda$ , com  $r = -0.95$  e  $p = 0.05$ ;  $r_d$  e  $LV_\lambda$ , com  $r = 1.00$  e  $p = 0.002$ ;  $B_\lambda$  e  $BC_\lambda$ , com  $r = 1.00$  e  $p = 0.001$ .

Ao longo da diagonal principal da figura 3.11 é fácil identificar 5 grupos de índices fortemente correlacionados: i)  $CV_s$  e  $G_s$ ; ii)  $Sk_s$  e  $r_s$ ; iii)  $Sk_d$  e  $K_{\Delta_s}$ ; iv)  $CV_d$ ,  $G_d$ ,  $r_d$  e  $LV_\lambda$ ; e v)  $B_\lambda$  e  $BC_\lambda$ . Esse resultado sugere que utilizar mais de um índice do mesmo grupo na análise da presente base de dados poderia ser redundante. Nessa direção, um grupo de índices com potencial para acessar diferentes aspectos do sistema em estudo poderia incluir, por exemplo, um índice de cada um desses grupos e  $M_\lambda$ . Nesse sentido, na figura 3.13, apresentamos os coeficientes de correlação  $r$  para todas as combinações de pares dentro de um grupo de seis índices selecionados:  $CV_s$ ,  $Sk_s$ ,  $CV_d$ ,  $Sk_d$ ,  $B_\lambda$  e  $M_\lambda$ .

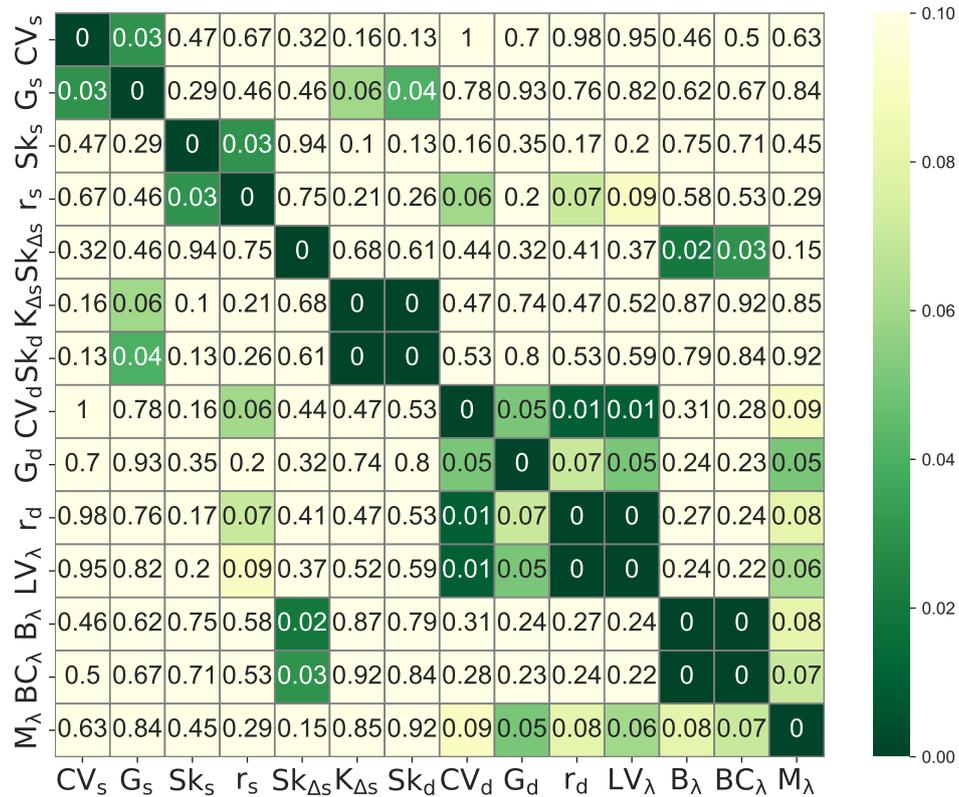
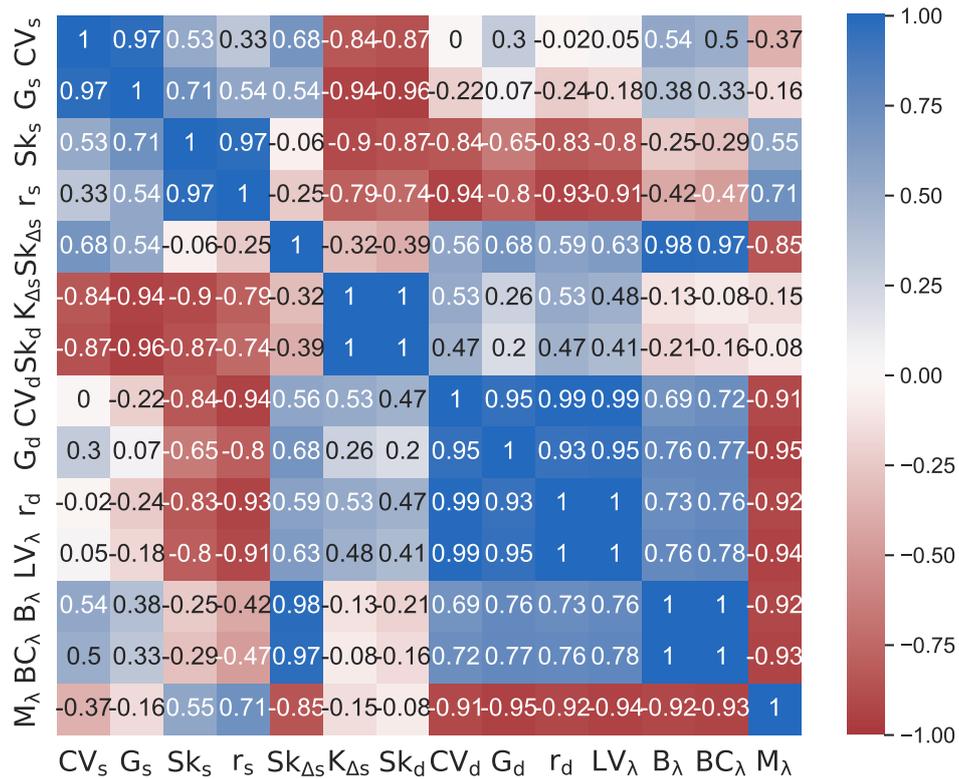


Figura 3.11: Em cima: Coeficiente de correlação de Pearson,  $r$ , para todos os pares de índices. Os índices relacionados aos intervalos de retorno  $\lambda$  são calculados com valor limite  $h = 2,05$ . Em baixo:  $p$ -valores correspondentes obtidos por meio do teste  $t$  de *student*.

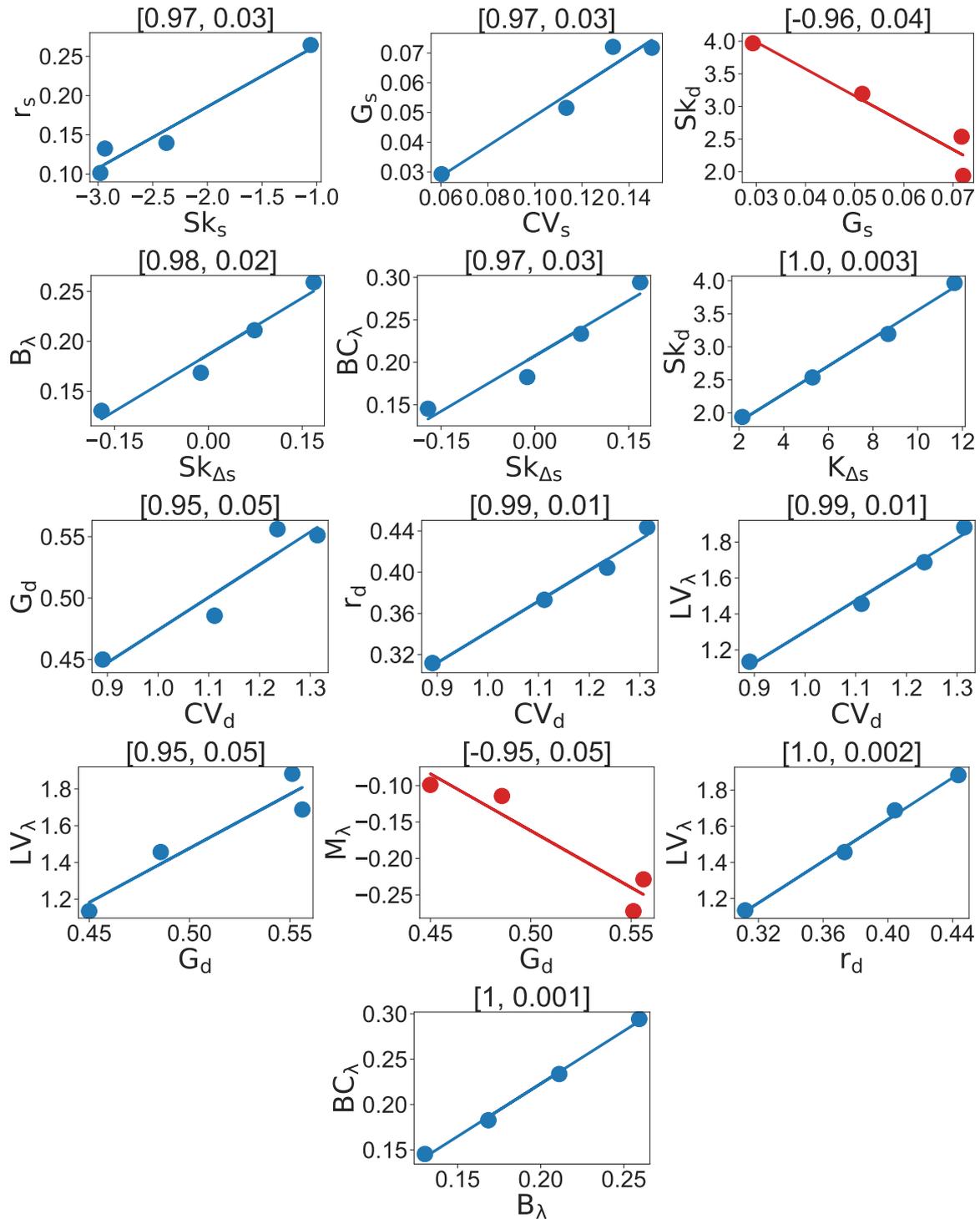


Figura 3.12: Gráficos de dispersão dos 13 pares de índices cujos valores do coeficiente de correlação de Pearson,  $r$ , são estatisticamente significativos. O coeficiente  $r$  e o  $p$ -valor são mostrados na parte superior dos gráficos.

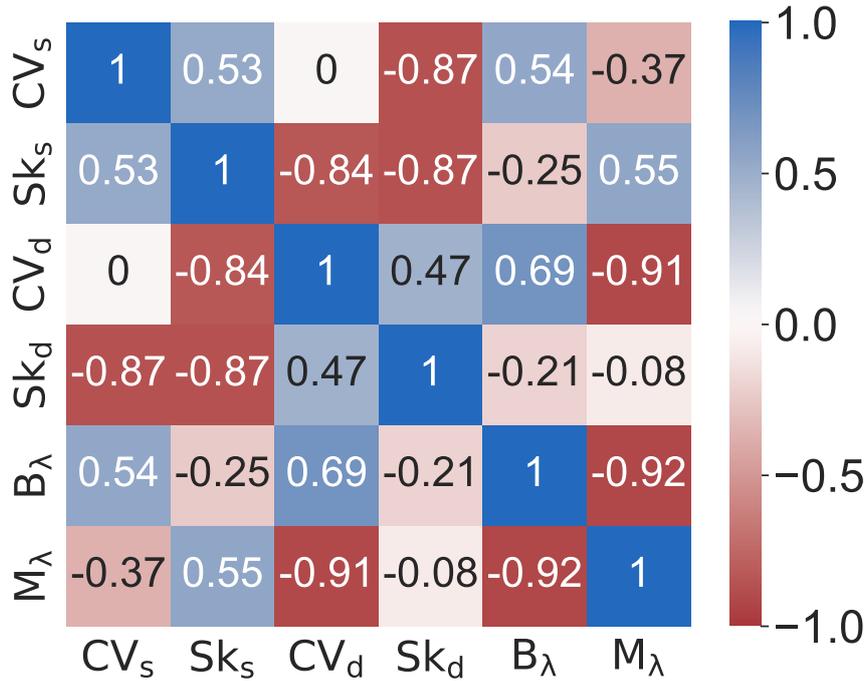


Figura 3.13: Coeficiente de correlação de Pearson,  $r$ , para todas as combinações de pares dentro de um grupo de seis índices selecionados.

### 3.3 Variações do valor limite $h$ das diferenças de tamanho

Até agora, os resultados relacionados aos índices  $LV_\lambda$ ,  $B_\lambda$ ,  $BC_\lambda$  e  $M_\lambda$  foram obtidos para um dado valor limite  $h$  ( $h = 2,05$ ). Nessa seção, vamos apresentar resultados para diferentes valores de  $h$ . Nas figuras 3.14-3.17, vemos  $LV_\lambda$ ,  $B_\lambda$ ,  $BC_\lambda$  e  $M_\lambda$ , para cada uma das amostras, com  $h$  variando no intervalo de 1.50 e 2.50 e com incrementos de 0.05. Essas figuras também mostram os respectivos valores médios dos índices calculados sobre  $10^2$  versões embaralhadas das sequências originais. De modo geral, os índices são relativamente robustos com a variação de  $h$ , sugerindo que nossos principais resultados não dependem da escolha de um valor particular de  $h$ . Nas figuras 3.14-3.17, vemos também que os efeitos de tamanho finito ficam mais evidentes para pequenos valores de  $h$ . Valores menores de  $h$  correspondem a mais eventos (cruzamentos de nível) e, como o tamanho da sequência é o mesmo, um número maior de eventos aumenta o efeito da discretização dos intervalos de retorno. Por outro lado, à medida que  $h$  aumenta o número de eventos diminui e isso pode afetar o cálculo dos momentos estatísticos da distribuição. Por esse motivo, o valor de  $h$  deve ser escolhido de modo a minimizar os efeitos de tamanho finito e ao mesmo tempo fornecer um número de eventos suficiente para o cálculo dos índices. No presente caso, verificamos que  $h \simeq 2$  atende a esses requisitos.

A seguir, investigamos o efeito da variação de  $h$  sobre a análise das correlações entre pares de índices. Na figura 3.18, mostramos os coeficientes de correlação de Pearson para um grupo selecionado de pares de índices, calculados para diferentes valores de  $h$ . Especificamente, os resultados referem-se a valores de  $h$  variando no intervalo de 1.50 e 2.50, com incrementos de 0.20. De modo geral, os resultados são robustos com a variação de  $h$  no intervalo escolhido. Em particular, o coeficiente  $r$  tende a diminuir com o aumento de  $h$  para alguns pares de índices. Esse é o caso, por exemplo, dos pares  $(Sk_s, B_\lambda)$ ,  $(CV_d, B_\lambda)$ ,  $(CV_d, M_\lambda)$  e  $(B_\lambda, M_\lambda)$ .

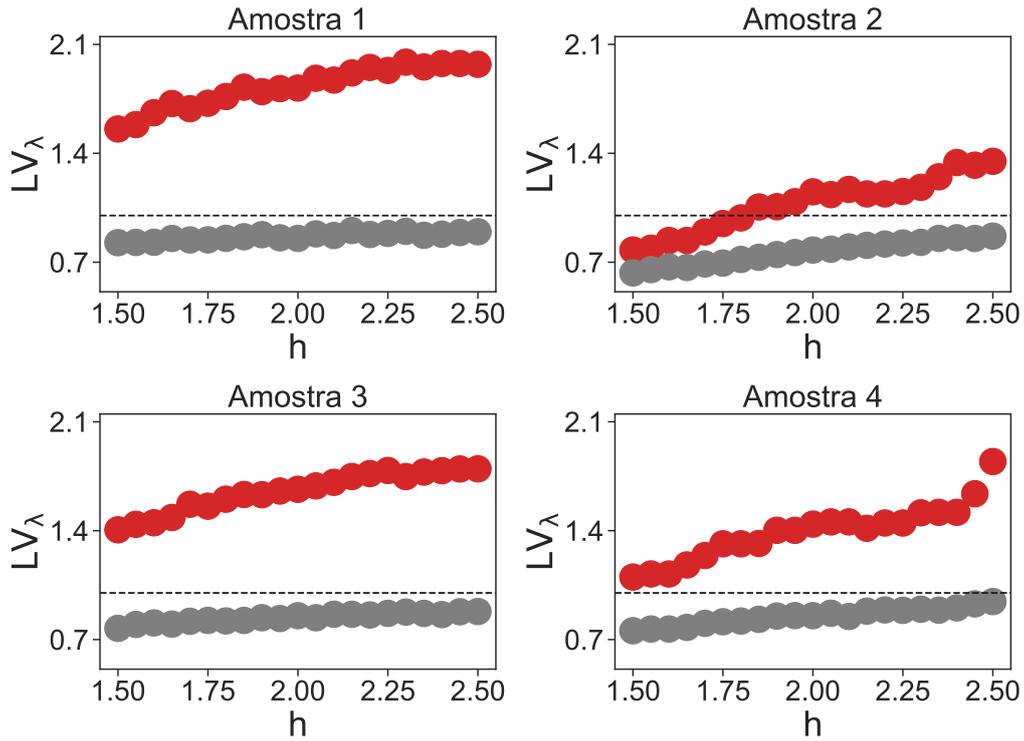


Figura 3.14:  $LV_\lambda$  em função de  $h$ , para  $h$  variando de 1.50 a 2.50, com incrementos 0.05. Valores correspondentes calculados a partir de seqüências embaralhadas dos eventos são mostrados em cinza. A linha tracejada corresponde a  $LV_\lambda = 1$ .

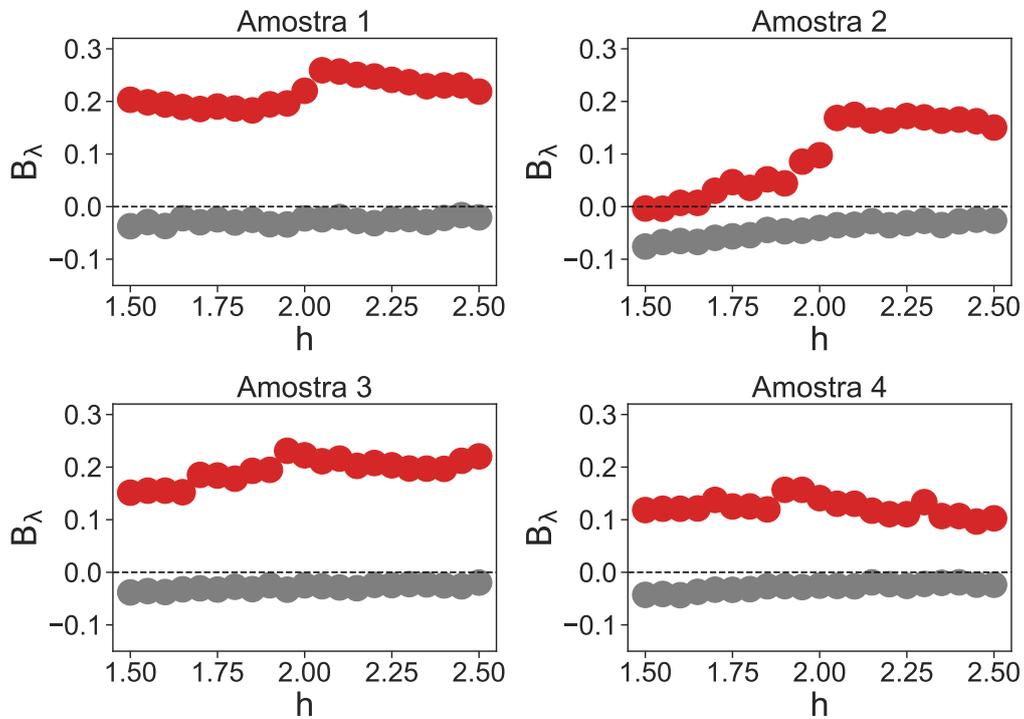


Figura 3.15:  $B_\lambda$  em função de  $h$ , para  $h$  variando de 1.50 a 2.50, com incrementos 0.05. Valores correspondentes calculados a partir de seqüências embaralhadas dos eventos são mostrados em cinza.

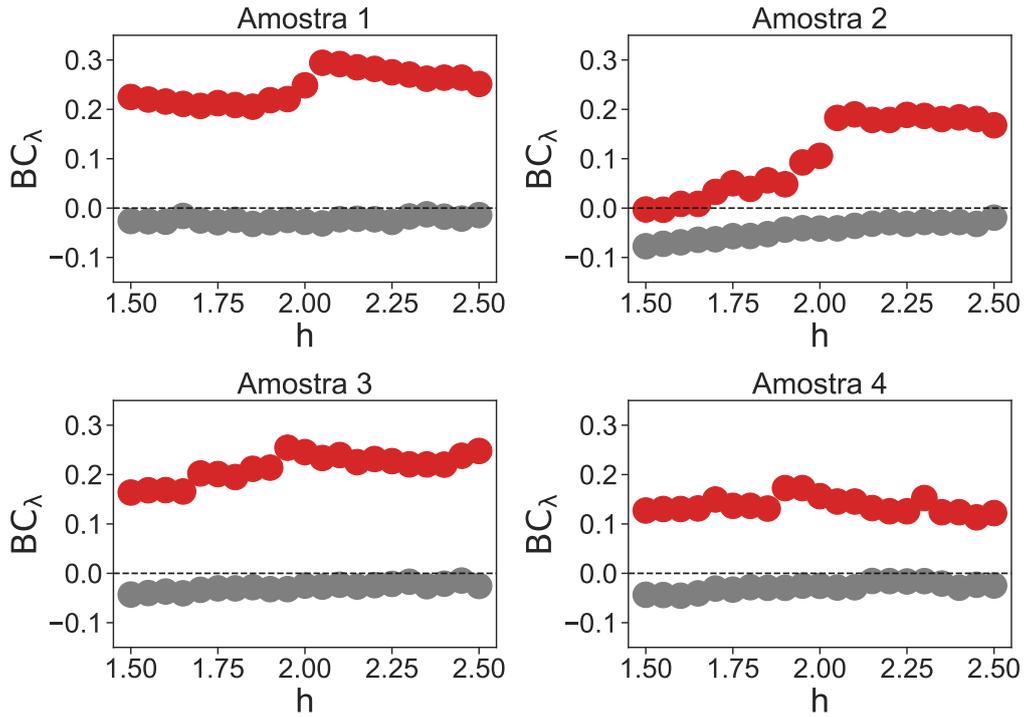


Figura 3.16:  $BC_\lambda$  em função de  $h$ , para  $h$  variando de 1.50 a 2.50, com incrementos 0.05. Valores correspondentes calculados a partir de seqüências embaralhadas dos eventos são mostrados em cinza.

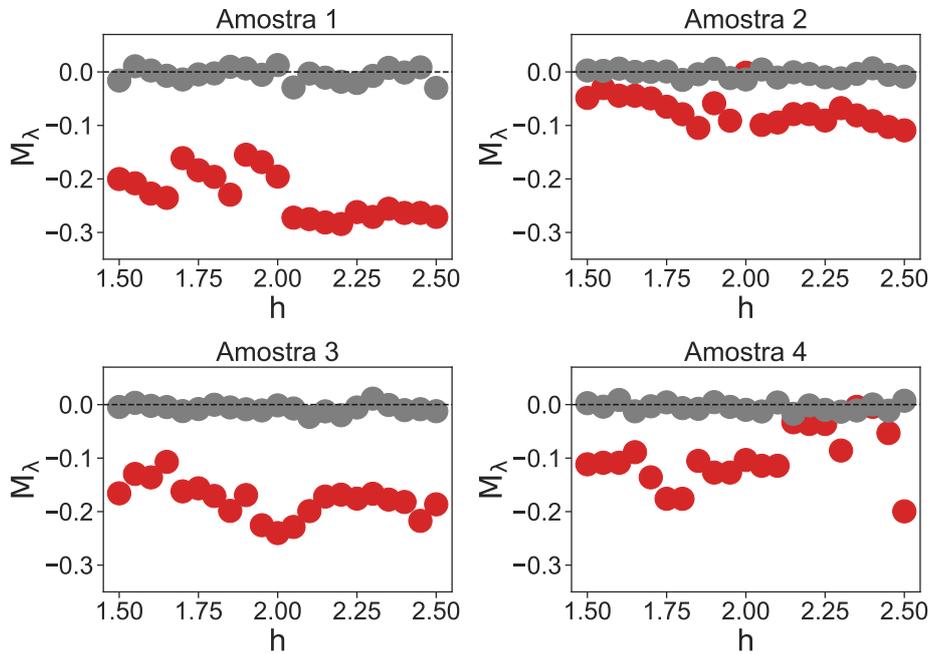


Figura 3.17:  $M_\lambda$  em função de  $h$ , para  $h$  variando de 1.50 a 2.50, com incrementos 0.05. Valores correspondentes calculados a partir de seqüências embaralhadas dos eventos são mostrados em cinza.

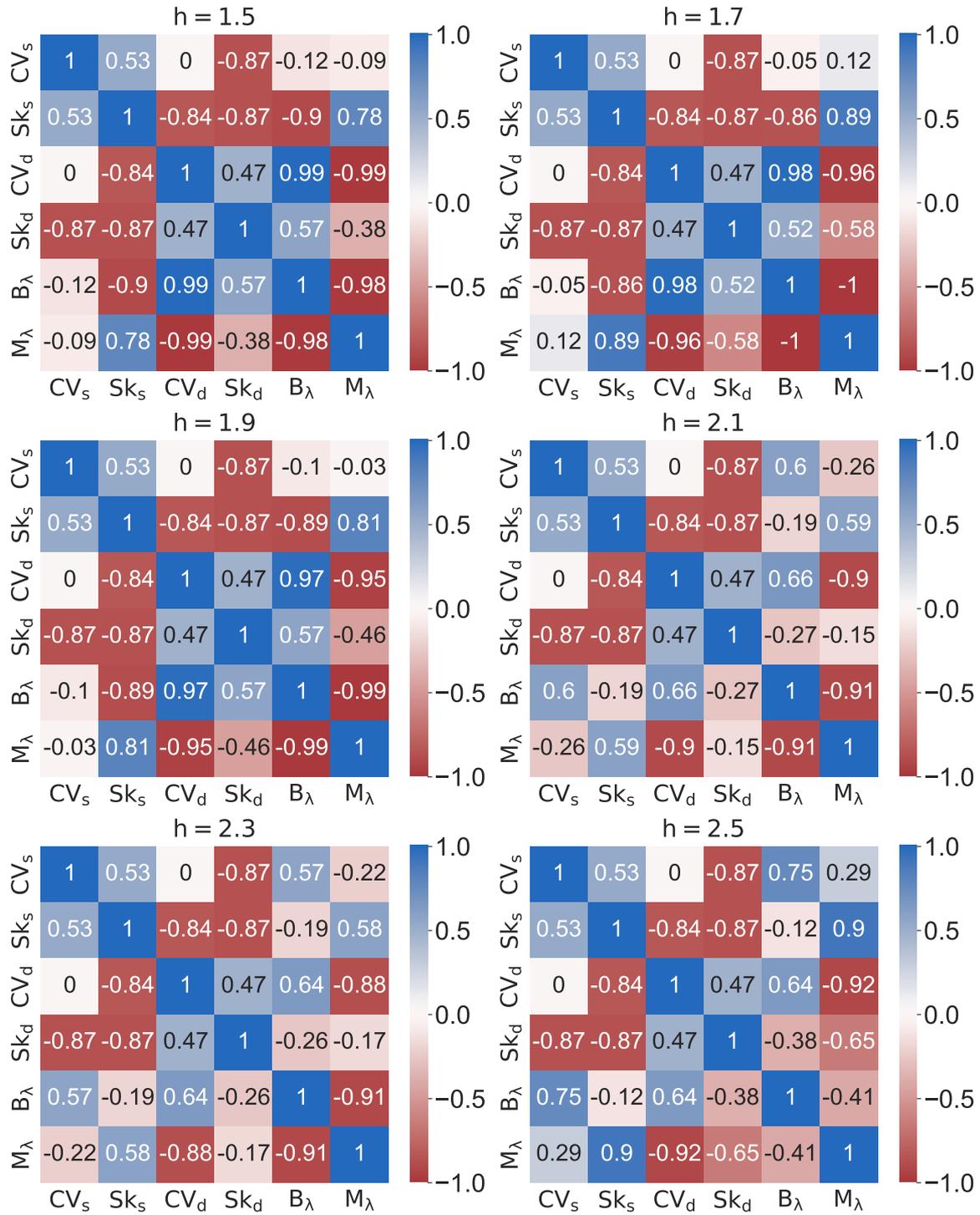


Figura 3.18: Coeficiente de correlação de Pearson,  $r$ , para um grupo selecionado de pares de índices, calculados para diferentes valores do valor limite  $h$ . Aqui, mostramos  $h$  variando de 1.5 a 2.5 com incrementos de 0.2.

# Capítulo 4

## Discussão e conclusões

Neste capítulo vamos resumir e discutir alguns dos principais resultados obtidos neste trabalho a partir da análise da variabilidade dos tamanhos de plantas sucessivas em fileiras de milho.

A distribuição dos tamanhos das plantas,  $P(s)$ , e os índices  $CV_s$ ,  $G_s$ ,  $Sk_s$  e  $r_s$  calculados a partir dos tamanhos  $s$  são apresentados nas figuras 3.1 e 3.6. Verificamos que  $P(s)$  é assimétrica, com  $Sk_s \simeq -2,3$  (assimetria negativa), e apresenta baixa variância, com  $CV_s \simeq 0,1$  ( $\simeq 10\%$ ). Além disso, as fileiras são caracterizadas por uma baixa desigualdade de tamanho, com índice de Gini  $G_s \simeq 0,06$ . Esses resultados são consistentes com trabalhos anteriores sobre variabilidade de tamanho de plantas em culturas de milho, que também apontam para uma assimetria negativa, baixa variância e índice de Gini relativamente baixo [9, 12, 23]. Sobre as autocorrelações na série dos tamanhos, obtivemos  $r_s \simeq 0,16$ , indicando a presença de correlações positivas nos tamanhos das plantas. Esse resultado sugere que, ao se percorrer a fileira, plantas altas tendem a ser seguidas por plantas altas e plantas baixas tendem a ser seguidas por plantas baixas. De modo geral, esse mesmo padrão tem sido observado em vários tipos de populações de plantas [24, 73, 74]. Entretanto, como geralmente as amostras analisadas têm caráter bidimensional, o índice que costuma ser utilizado para quantificar correlações espaciais em populações de plantas é a estatística  $I$  de Moran [24, 73, 74].

A distribuição dos incrementos de tamanho,  $P(\Delta s)$ , e os índices  $Sk_{\Delta s}$  e  $K_{\Delta s}$  calculados a partir dos incrementos  $\Delta s$  são apresentados nas figuras 3.2 e 3.7. Os resultados indicam que  $P(\Delta s)$  é aproximadamente simétrica, com  $Sk_{\Delta s} \simeq 0$ . Além disso, os valores positivos para o excesso de curtose,  $K_{\Delta s} \simeq 6,9$ , indicam que  $P(\Delta s)$  é mais afunilada e/ou possui caudas mais longas que a distribuição normal. Esses resultados são consistentes com a ref. [36] que também aponta para um comportamento não gaussiano (com caudas mais longas que a distribuição gaussiana) para a distribuição dos incrementos de tamanho em fileiras de milho. Em particular, o uso do excesso

de curtose  $K_{\Delta_s}$  permite quantificar esse aspecto da distribuição dos incrementos de tamanho sem a necessidade de escolher uma distribuição específica como no caso da ref. [36].

A distribuição das diferenças (normalizadas) de tamanho entre plantas vizinhas,  $P(d')$ , e os índices  $CV_d$ ,  $G_d$ ,  $Sk_d$  e  $r_d$  calculados a partir das diferenças de tamanho  $d$  são apresentados nas figuras 3.3 e 3.8. A distribuição  $P(d')$  é assimétrica, com  $Sk_d \simeq 2,9$  (assimetria positiva). Além disso,  $P(d)$  apresenta coeficiente de variação próximo da unidade,  $CV_d \simeq 1,1$ , e índice de Gini  $G_d \simeq 0,5$ . Esse último resultado indica uma desigualdade relativamente grande nas diferenças de tamanho entre plantas vizinhas. Para fins de comparação, relativo à renda da população, o Brasil apresenta um coeficiente de Gini de  $\simeq 0,5$  [75]. Além disso, verificamos que a distribuição empírica  $P(d')$  se assemelha a uma distribuição log-normal (figuras 3.3 e 3.4). Como vimos,  $d'$  pode ser visto como um tipo de 'volatilidade' de tamanhos (de natureza espacial). É interessante notar que a distribuição da volatilidade de preços no mercado financeiro, para valores moderados da volatilidade, também tem sido descrita como uma log-normal [52, 54]. De modo geral, distribuições log-normais têm sido identificadas em uma grande variedade de sistemas naturais e costumam ser associadas a processos multiplicativos [76].

A distribuição acumulada dos intervalos de retorno normalizados,  $P_c(\lambda')$ , e os índices  $LV_\lambda$ ,  $B_\lambda$ ,  $BC_\lambda$  e  $M_\lambda$  calculados a partir dos intervalos de retorno  $\lambda$  são apresentados nas figuras 3.5 e 3.9. Os resultados obtidos,  $LV_\lambda \simeq 1,5$ ,  $B_\lambda \simeq 0,2$ ,  $BC_\lambda \simeq 0,2$  e  $M_\lambda \simeq -0,2$  indicam que as grandes diferenças de tamanho ( $d' > 2.05$ ) não ocorrem aleatoriamente ao longo das fileiras de milho. Para comparação, em um processo de Poisson, teríamos  $LV_\lambda = 1$ ,  $B_\lambda = 0$ ,  $BC_\lambda = 0$  e  $M_\lambda = 0$ . O resultado  $B_\lambda > 0$  indica que a distribuição dos intervalos de retorno decai mais lentamente que uma exponencial. O resultado  $M_\lambda < 0$  indica que os intervalos de retorno apresentam anti-correlações: intervalos longos tendem a ser seguidos por intervalos curtos e intervalos curtos tendem a ser seguidos por intervalos longos.

Podemos comparar nossos resultados envolvendo  $B_\lambda$  e  $M_\lambda$  com resultados análogos obtidos no estudo da volatilidade de preços no mercado financeiro. Alguns resultados indicam que a distribuição dos intervalos de retorno da volatilidade de preços decai mais lentamente que uma exponencial [55]. Esse resultado é consistente com nossos resultados sobre tamanho de plantas ( $B_\lambda > 0$ ). Por outro lado, em contraste com nossos resultados sobre tamanho de plantas ( $M_\lambda < 0$ ), os intervalos de retorno da volatilidade de preços exibem um comportamento consistente com memória positiva [55]. Essas diferenças na memória dos intervalos de retorno talvez possam refletir diferenças estruturais entre a volatilidade de preços no mercado financeiro e as diferenças de tamanho de plantas ao longo de fileiras de milho. No mercado financeiro é usual que

uma grande variação de preço (alta volatilidade) seja seguida por várias outras grandes variações, gerando vários cruzamentos de nível próximos um do outro e, portanto, vários intervalos de retorno sucessivos relativamente pequenos. Isso pode contribuir para uma memória positiva nos intervalos de retorno. Em contraste, plantas com tamanho muito abaixo da média em uma fileira de milho costumam ocorrer sozinhas, gerando apenas dois cruzamentos de nível das diferenças de tamanho, e portanto, apenas um intervalo de retorno relativamente pequeno. Esse comportamento pode contribuir para uma memória negativa nos intervalos de retorno das diferenças de tamanho.

Como vimos, os resultados que obtivemos envolvendo  $B_\lambda$  e  $M_\lambda$  seguem um padrão: *burstiness* positivo e memória negativa. Isso corresponde ao segundo quadrante em um diagrama  $MB$  (figura 3.10). É interessante também comparar esses resultados com padrões de ocorrência de eventos observados em outros sistemas complexos. Por exemplo, diversas atividades humanas, como enviar e-mails ou iniciar chamadas telefônicas, seguem padrões descritos por um coeficiente *burstiness* positivo e memória positiva mas próxima de zero. Isso localiza essas atividades principalmente no primeiro quadrante do diagrama  $MB$  [60]. Outros exemplos incluem fenômenos naturais como terremotos e precipitação que também exibem coeficiente *burstiness* positivo e memória positiva, sendo localizados próximo à diagonal do primeiro quadrante em um diagrama  $MB$  [60]. Por outro lado, no caso dos intervalos de tempo entre batidas do coração a localização ocorre no quarto quadrante, com *burstiness* negativo e memória positiva [60]. De acordo com nossos resultados, os cruzamentos de nível das diferenças de tamanho entre plantas vizinhas em culturas de milho (localizados no segundo quadrante) exibem um padrão diferente do observado nesses outros sistemas citados.

Sobre a análise da correlação entre pares de índices, verificamos que alguns dos 14 índices utilizados neste trabalho são fortemente correlacionados (figura 3.11). Sabemos que índices fortemente correlacionados ou fortemente anti-correlacionados podem conter informações redundantes. Por esse motivo, um conjunto de índices fracamente correlacionado seria mais indicado para descrever diferentes aspectos do sistema. Com base nesse tipo de análise, sugerimos um conjunto de 6 índices que poderia ser usado para capturar diferentes aspectos do sistema em estudo. Esse grupo inclui os seguintes índices:  $CV_s$ ,  $Sk_s$ ,  $CV_d$ ,  $Sk_d$ ,  $B_\lambda$  e  $M_\lambda$  (figura 3.13). Uma limitação da presente análise das correlações entre índices é o pequeno número de amostras independentes que utilizamos (4, ao todo). Seria desejável repetir essa análise para um número maior de amostras. Além disso, é bom lembrar que as correlações entre índices podem diferir dependendo do tipo de planta que está sendo analisada e da medida de tamanho que está sendo utilizada [30].

Neste trabalho adotamos uma abordagem unidimensional para a análise da variabilidade do tamanho de plantas em culturas de milho. Desse modo, foi possível

utilizar índices que usualmente são aplicados na análise de séries temporais como  $B_\lambda$  e  $M_\lambda$ . Por outro lado, esses índices relacionados aos intervalos de retorno  $\lambda$  têm uma desvantagem prática visto que necessitam de amostras de tamanho considerável. Por exemplo, o número total de plantas em cada amostra utilizada neste trabalho é  $\simeq 2056$  (valor médio calculado sobre as 4 amostras). Entretanto, o número total de intervalos de retorno associado ao valor limite  $h = 2,05$  é  $\simeq 193$ . Isso ilustra que as amostras devem ser suficientemente grandes para que a análise dos intervalos de retorno seja viável. É importante ressaltar aqui a dificuldade de se obter dados sobre tamanho de plantas em monoculturas. Geralmente, medidas obtidas de modo direto, como as utilizadas neste trabalho, envolvem um considerável custo em termos de tempo e mão de obra. Por outro lado, recentes avanços tecnológicos têm o potencial de facilitar a obtenção de dados relacionados a monoculturas agrícolas. Os equipamentos utilizados incluem diferentes tipos de sensores, como sensores ultrassônicos e câmeras de alta resolução, combinados com veículos terrestres e aéreos (drones). Em particular, esse tipo de técnica tem sido utilizada para a obtenção de dados sobre tamanho de plantas em monoculturas de milho [77]. O acesso a dados obtidos por meios automatizados poderia contribuir para o aprimoramento do estudo da variabilidade do tamanho de plantas em fileiras de milho.

Em monoculturas agrícolas, a competição entre plantas é um fator chave que tem sido relacionado à variabilidade de tamanho dos indivíduos da população. Entretanto, outros trabalhos são necessários para quantificar até que ponto nossos resultados refletem em algum grau efeitos de competição. No presente trabalho, os dados analisados foram obtidos em plantações de milho com densidade de plantas praticamente constante. Como vimos no capítulo de revisão, vários trabalhos no contexto de monoculturas exploram o efeito da densidade de plantas (relacionada à intensidade da competição) sobre a estrutura de tamanho dos indivíduos da população. Nessa direção, trabalhos futuros poderiam explorar a dependência dos índices utilizados aqui com a densidade de plantas. Além disso, os dados analisados neste trabalho se referem a um determinado estágio do desenvolvimento das plantas. Um aspecto que tem sido abordado no contexto de monoculturas é o acompanhamento temporal das populações (como destacado no capítulo de revisão). Desse modo, outra possível extensão desse trabalho poderia incluir uma análise sobre o comportamento dos índices propostos calculados em diferentes estágios de desenvolvimento da planta, desde a germinação até a fase pré-colheita. Além disso, na literatura sobre variabilidade de tamanho de plantas em monoculturas, outros tipos de plantas têm sido analisadas incluindo plantas que são utilizadas comercialmente, como a soja (*Glycine max* L.), o girassol (*Helianthus annuus* L.), a linhaça (*Linum usitatissimum* L.) e a ervilha (*Pisum sativum* L.). Trabalhos futuros poderiam também explorar o comportamento dos índices utilizados aqui

para estudar diferentes tipos de monocultura.

Como um comentário final, vários modelos têm sido propostos para descrever o processo de crescimento dos indivíduos de uma população sob o efeito da competição com plantas vizinhas [18, 78, 79]. O tipo de análise empregada no presente trabalho poderia ser utilizada, por exemplo, para comparar os resultados de diferentes modelos de crescimento de plantas com os padrões empíricos observados em monoculturas.

# Apêndice A

## Momentos das distribuições e $k$ -estatística de Fisher

Neste apêndice, definimos os coeficientes descritos pela  $k$ -estatística de Fisher para o cálculo dos momentos estatísticos das distribuições [71]. Correções dos momentos estatísticos determinados pela  $k$ -estatística de Fisher costumam ser utilizadas para levar em conta amostras com tamanho finito. Os momentos das distribuições são utilizados nesse trabalho para quantificar a variabilidade do tamanho de plantas em fileiras de milho.

Aqui estamos interessados nos momentos da distribuição  $P(x)$ , associados aos valores observados  $x_n$ ,  $n = 1, \dots, N$ . Os momentos  $k_2$ ,  $k_3$  e  $k_4$ , relativos à variável  $x$ , são definidos como

$$k_2 = \frac{N}{N-1} m_2, \quad (\text{A.1})$$

$$k_3 = \frac{N^2}{(N-1)(N-2)} m_3, \quad (\text{A.2})$$

$$k_4 = \frac{N^2[(N+1)m_4 - 3(N-1)m_2^2]}{(N-1)(N-2)(N-3)}, \quad (\text{A.3})$$

em que  $m_2$ ,  $m_3$  e  $m_4$  são os momentos centrais dados por

$$m_k = \frac{1}{N} \sum_{n=1}^N (x_n - \bar{x})^k, \quad (\text{A.4})$$

em que  $k = 2, 3$  e  $4$ , respectivamente.

## Apêndice B

# Índice de Gini e curva de Lorenz

Neste apêndice, apresentamos uma breve introdução sobre a relação entre a curva de Lorenz e o índice de Gini, suas representações gráficas e interpretação.

Considere uma variável  $x_n$ ,  $n = 1, 2, \dots, N$ , organizada de tal forma que  $x_1 \leq x_2 \leq x_3 \leq \dots \leq x_N$ . A proporção acumulada da população até o  $n$ -ésimo elemento é dada por

$$p_n = \frac{n}{N}. \quad (\text{B.1})$$

Por sua vez, a proporção acumulada da variável  $x_n$  é definida como

$$\Phi_n = \frac{1}{N\bar{x}} \sum_{j=1}^n x_j, \quad (\text{B.2})$$

em que  $\bar{x}$  é o valor médio de  $x$ . Os pares de valores  $(p_n, \Phi_n)$  correspondem a pontos que, quando unidos, formam a chamada curva de Lorenz.

Na figura B.1, apresentamos uma curva de Lorenz típica. Para ajudar na interpretação, vamos considerar alguns casos particulares. Note que os valores máximos de  $p_n$  e  $\Phi_n$  são iguais a 1. Suponha que todos os  $x_n$  sejam iguais ( $x_n = \bar{x}$ ), caracterizando uma igualdade perfeita. Nesse caso, a curva de Lorenz é representada por uma linha reta diagonal que une os pontos (0,0) e (1,1) (aqui estamos considerando que o número de termos é suficientemente grande). Outro caso particular é o de perfeita desigualdade, em que um elemento possui tudo e todos os demais elementos são nulos. Nesse caso extremo, a curva de Lorenz coincide com o eixo horizontal do gráfico em quase toda sua extensão. No último ponto,  $\Phi_n$  dá um salto e atinge a unidade. Nos casos intermediários a curva de Lorenz estará abaixo da diagonal, iniciando próximo de (0,0) e terminando próximo do ponto (1,1).

A área de desigualdade,  $\alpha$ , é definida como a área da região entre a linha de perfeita igualdade e a curva de Lorenz (veja a figura B.1). Em um caso extremo de desigualdade, em que  $N - 1$  elementos são nulos e apenas um é não nulo, a região de

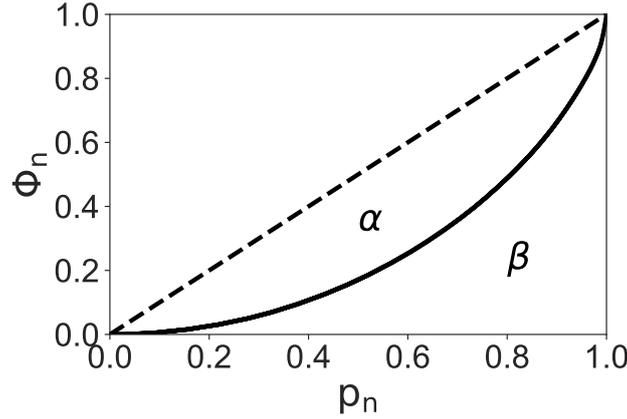


Figura B.1: Proporção acumulada da população,  $p_n$  versus proporção acumulada da variável  $x$ ,  $\Phi_n$ . A linha tracejada representa uma curva de Lorenz para o caso de perfeita igualdade. A linha sólida representa uma curva de Lorenz típica, um caso intermediário entre a perfeita igualdade e a total desigualdade. As áreas  $\alpha$  e  $\beta$  estão identificadas no gráfico.

desigualdade é descrita por um triângulo de altura igual a 1 e base  $1 - 1/N$ . Dessa forma, o valor máximo de  $\alpha$  para uma distribuição discreta é dada pela relação

$$\alpha_{max} = \frac{1}{2}\left(1 - \frac{1}{N}\right), \quad (\text{B.3})$$

que corresponde à área do triângulo mencionado. Note que  $\alpha_{max}$  tende para  $1/2$  quando o número de termos  $N$  tende para infinito. O índice de Gini é definido como o quociente entre  $\alpha$  e esse limite. Assim,

$$G = \frac{\alpha}{1/2} = 2\alpha. \quad (\text{B.4})$$

Como  $0 \leq \alpha \leq \alpha_{max}$ , temos que  $0 \leq G \leq (1 - \frac{1}{N})$ . No limite de  $N$  grande, o valor máximo de  $G$  tende a 1.

O cálculo do índice de Gini pode ser feito sem precisar calcular explicitamente a área  $\alpha$  em um gráfico da curva de Lorenz. Para verificar isso, vamos definir a região entre o eixo das abcissas e a curva de Lorenz como uma área  $\beta$  (veja a figura B.1). Por construção temos que  $\alpha + \beta = \frac{1}{2}$ . Podemos decompor  $\beta$  em  $N - 1$  trapézios. Para cada trapézio, a base maior e a base menor são, respectivamente,  $\Phi_n$  e  $\Phi_{n-1}$  e a altura é dada pela relação  $p_n - p_{n-1} = 1/N$ . Assumindo que  $\Phi_0 = 0$ , temos que a área de cada trapézio é dada por

$$S_n = \frac{1}{2N}(\Phi_n + \Phi_{n-1}). \quad (\text{B.5})$$

A soma das áreas de todos os trapézios representa a área  $\beta$ . Assim,

$$\beta = \sum_{n=1}^N S_n = \frac{1}{2N} \sum_{n=1}^N (\Phi_n + \Phi_{n-1}). \quad (\text{B.6})$$

Como  $G = 2\alpha$  e  $\alpha + \beta = \frac{1}{2}$ , temos que

$$G = 1 - 2\beta = 1 - \frac{1}{N} \sum_{n=1}^N (\Phi_n + \Phi_{n-1}). \quad (\text{B.7})$$

O somatório que aparece nessa equação, em termos de  $\Phi_n$ , pode ser reescrito em termos de  $x_n$ , de tal forma que

$$\sum_{n=1}^N (\Phi_n + \Phi_{n-1}) = \frac{1}{N\bar{x}} \sum_{n=1}^N (2N - 2n + 1)x_n. \quad (\text{B.8})$$

Substituindo a relação [B.8](#) na equação [B.7](#), temos que

$$G = \frac{2}{N^2\bar{x}} \sum_{n=1}^N nx_n - \left(1 - \frac{1}{N}\right). \quad (\text{B.9})$$

Por fim, essa expressão pode ser reescrita como

$$G = \frac{\sum_n \sum_j |x_n - x_j|}{2\bar{x}N^2}, \quad (\text{B.10})$$

que corresponde à média dos valores absolutos das diferenças entre dois valores quaisquer da variável  $x_n$ . Essa última expressão é a que utilizamos nesse trabalho para o cálculo do índice de Gini (veja a equação [2.11](#)).

Nas figuras [B.2](#) e [B.3](#), apresentamos respectivamente as curvas de Lorenz para os tamanhos e os incrementos de tamanho das plantas investigadas no trabalho

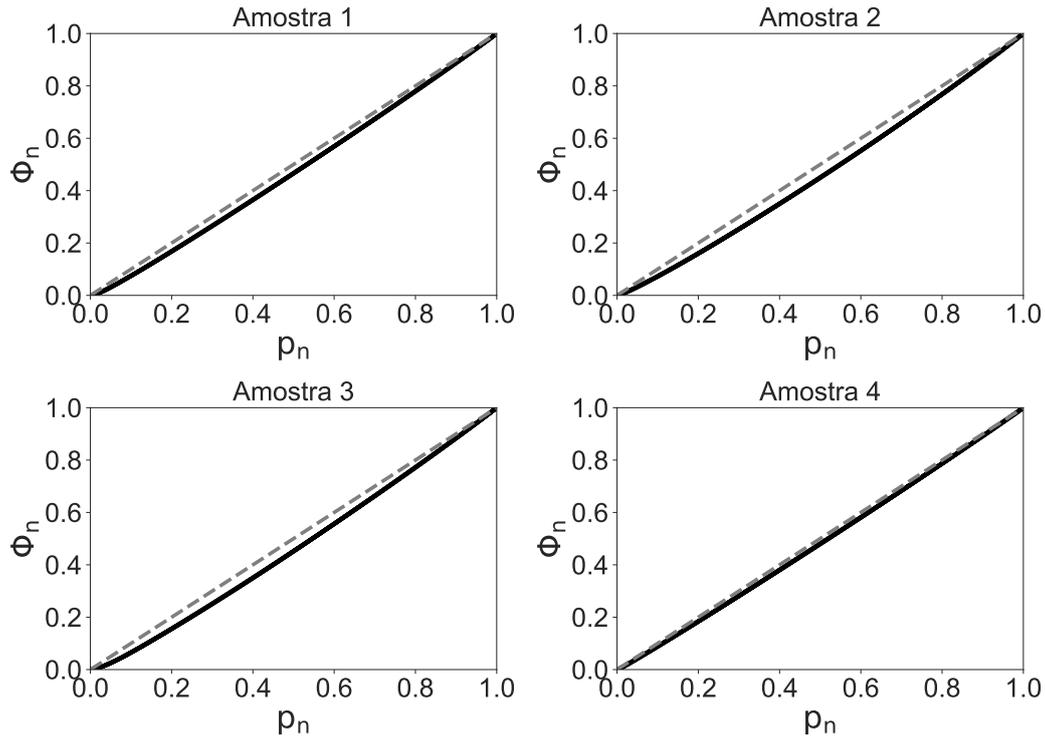


Figura B.2: Curvas de Lorenz referentes ao tamanho das plantas  $s$ , calculadas para cada uma das amostras. O valor médio dos índices de Gini correspondentes é igual a 0.06.

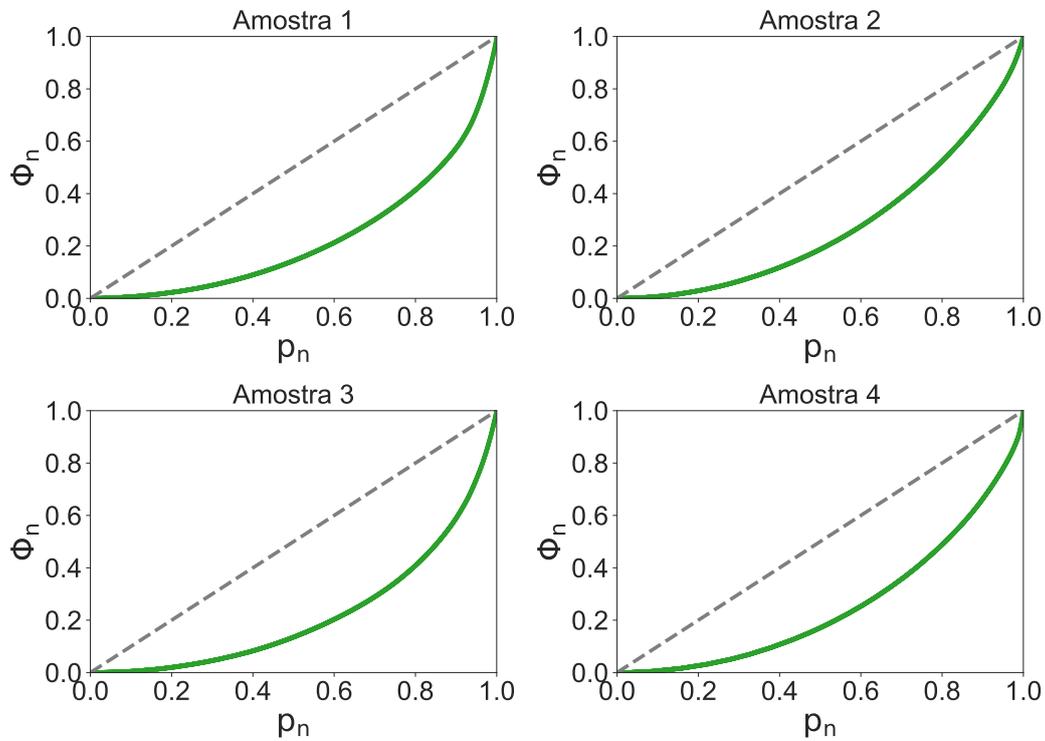


Figura B.3: Curvas de Lorenz referentes às diferenças de tamanho entre plantas vizinhas  $d$ , calculadas para cada uma das amostras. O valor médio dos índices de Gini correspondentes é igual a 0.51.

# Apêndice C

## Coeficiente de correlação de Pearson

Neste apêndice, descrevemos o procedimento para o cálculo do coeficiente de correlação de Pearson,  $r$ , entre cada par de índices. Também descrevemos o teste  $t$  de *student* utilizado para verificar se  $r$  é estatisticamente significativo.

Considere um par genérico de índices,  $(W_i, Y_i)$ , calculado a partir da amostra  $i$ , com  $i = 1, \dots, m$ . No presente caso,  $m = 4$  visto que temos 4 amostras independentes. O coeficiente de correlação de Pearson entre  $W_i$  e  $Y_i$  é definido como

$$r = \frac{\sum_{i=1}^m (W_i - \bar{W})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^m (W_i - \bar{W})^2} \sqrt{\sum_{i=1}^m (Y_i - \bar{Y})^2}}, \quad (\text{C.1})$$

em que  $\bar{W}$  e  $\bar{Y}$  são os valores médios de  $W_i$  e  $Y_i$ , respectivamente. O coeficiente  $r$  é uma medida da associação ou correlação linear entre  $W$  e  $Y$ , fornecendo informação sobre a magnitude e a direção da associação ou correlação linear. O coeficiente  $r$  pode assumir valores no intervalo  $[-1, 1]$ . Valores de  $r$  iguais a  $-1$  ou  $+1$  implicam uma relação linear perfeita. Valores positivos de  $r$  implicam uma associação positiva (se uma variável aumenta, então a outra também aumenta). Valores negativos de  $r$  implicam uma associação negativa (se uma variável aumenta, a outra diminui).

O coeficiente de correlação de Pearson pode ser usado para testar se a associação linear entre as duas variáveis é estatisticamente significativa. O coeficiente  $r$  (o coeficiente de Pearson da amostra) é uma estimativa de  $\rho$ , o coeficiente de Pearson da população. Conhecendo-se  $r$  e o número de pontos  $m$ , podemos inferir se  $\rho$  é significativamente diferente de 0. Assim, temos:

$$\text{Hipótese nula } H_0: \rho = 0; \quad (\text{C.2})$$

$$\text{Hipótese alternativa } H_a: \rho \neq 0. \quad (\text{C.3})$$

Para testar a hipótese nula, calculamos o valor da estatística  $t$ ,

$$t = \frac{r}{\sigma_r}, \quad (\text{C.4})$$

em que  $\sigma_r$  é um tipo de erro padrão dado por

$$\sigma_r = \left( \frac{m-2}{1-r^2} \right). \quad (\text{C.5})$$

O próximo passo é obter o valor crítico de  $t$ ,  $t^*$ . O valor crítico  $t^*$  pode ser obtido a partir de uma tabela  $t$ , usando o número de graus de liberdade dado por  $df = m - 2$ , o nível de significância  $\alpha$  (usualmente,  $\alpha = 0.05$ ) e escolhendo o teste bilateral. Se  $|t| > t^*$ , podemos rejeitar a hipótese nula: a associação quantificada por  $r$  é estatisticamente significativa ( $p < \alpha$ ) em que  $p$  representa o  $p$ -valor.

# Apêndice D

## Processos de Bernoulli e Poisson

Um ensaio de Bernoulli produz um resultado para a variável aleatória discreta  $z$ , descrita pela distribuição de probabilidades

$$P(z) = \begin{cases} p, & z = 1 \\ 1 - p, & z = 0 \\ 0, & \text{outros casos.} \end{cases} \quad (\text{D.1})$$

Tipicamente, o resultado de um ensaio de Bernoulli é chamado de ‘sucesso’ se  $z = 1$  e de ‘falha’ se  $z = 0$ . Um processo de Bernoulli é uma série de ensaios de Bernoulli independentes, cada um com a mesma probabilidade de sucesso  $p$ . Vamos representar um processo de Bernoulli, composto de  $N$  ensaios independentes, como  $z_n$ ,  $n = 1, \dots, N$ . Usualmente, é conveniente chamar os sucessos em um processo de Bernoulli como ‘chegadas’. Considere uma variável aleatória discreta  $\lambda$  dada pelo número de ensaios de Bernoulli até (e incluindo) o primeiro sucesso. A variável aleatória  $\lambda$  é conhecida como tempo entre chegadas de primeira ordem. Essa variável pode assumir valores  $\lambda_l$ ,  $l = 1, 2, \dots, L$ , em que  $\lambda_l$  é inteiro e maior ou igual a 1.

Não é difícil ver que a distribuição de probabilidades dos tempos entre chegadas,  $P(\lambda)$ , é dada por [80]

$$P(\lambda) = p(1 - p)^{\lambda-1}, \quad (\text{D.2})$$

com  $\lambda = 1, 2, \dots, L$ . Visto que seus termos sucessivos decrescem em progressão geométrica, a distribuição (D.2) é conhecida como distribuição geométrica. Note que  $P(\lambda)$  foi obtida como a distribuição do número de ensaios até (e incluindo) o primeiro sucesso. Entretanto, em vista dos ensaios de Bernoulli serem independentes (sem memória), a variável aleatória  $\lambda$  também representa o número de ensaios entre um sucesso e o próximo sucesso.

Um processo de Poisson pode ser definido a partir de uma descrição probabilística de chegadas em processos de Bernoulli. Em um processo de Poisson, as chegadas ocor-

rem de modo contínuo (como em pontos sobre uma linha). Por conveniência, vamos nos referir a essa linha como o eixo do tempo,  $t$ . Utilizando essa interpretação, um processo de Poisson pode ser definido como o limite, em que  $\Delta t \rightarrow 0$ , de uma série de ensaios de Bernoulli idênticos e independentes que ocorrem a cada intervalo de tempo  $\Delta t$ . Cada um desses ensaios tem uma probabilidade de sucesso dada por  $p = \alpha\Delta t$ , em que  $\alpha$  é uma constante. Nesse limite, o tempo entre chegadas  $\lambda$  será uma variável aleatória contínua. É possível mostrar que, nesse limite, a distribuição  $P(\lambda)$  será exponencial, com valor médio igual a  $1/\alpha$  [80,81]. Podemos dizer então que um processo de Bernoulli, de tempo discreto, converge para um processo de Poisson homogêneo, de tempo contínuo. Processos de Bernoulli e processos de Poisson homogêneos descrevem eventos completamente aleatórios, independentes entre si. Por esse motivo, eles são muito utilizados como modelos de referência no estudo de sistemas temporalmente heterogêneos.

# Referências Bibliográficas

- [1] D. Sornette, *Critical phenomena in natural sciences — Chaos, fractals, self-organization and disorder: Concepts and tools*. Springer International Publishing, 2006.
- [2] M. Mitchell, *Complexity: A guided tour*. Oxford University Press, 2009.
- [3] N. Boccarda, *Modeling complex systems*. Springer Science & Business Media, 2010.
- [4] J. L. Harper, *Population biology of plants*. Academic Press., 1977.
- [5] C. G. N. Turnbull, *Plant architecture and its manipulation*. Wye Campus, Kent UK, 2005.
- [6] M. Anand, A. Gonzalez, F. Guichard, J. Kolasa, and L. Parrott, “Ecological systems as complex systems: challenges for an emerging science,” *Diversity*, vol. 2, no. 3, pp. 395–410, 2010.
- [7] E. Ford, “Competition and stand structure in some even-aged plant monocultures,” *The Journal of Ecology*, pp. 311–333, 1975.
- [8] R. E. Naylor, “Changes in the structure of plant populations,” *Journal of Applied Ecology*, pp. 513–521, 1976.
- [9] G. O. Edmeades and T. B. Daynard, “The development of plant-to-plant variability in maize at different planting densities,” *Canadian Journal of Plant Science*, 2011.
- [10] E. Gorham, “Shoot height, weight and standing crop in relation to density of monospecific plant stands,” *Nature*, vol. 279, no. 5709, pp. 148–150, 1979.
- [11] D. Rabinowitz, “Bimodal distributions of seedling weight in relation to density of *Festuca paradoxa* desv.,” *Nature*, vol. 277, no. 5694, 1979.
- [12] T. Daynard and J. Muldoon, “Plant-to-plant variability of maize plants grown at different densities,” *Canadian Journal of Plant Science*, vol. 63, no. 1, pp. 45–59, 1983.

- [13] W. R. Hawthorn and P. B. Cavers, “Dry weight and resource allocation patterns among individuals in populations of *Plantago major* and *P. rugelii*,” *Canadian Journal of Botany*, vol. 60, p. 2424, 1982.
- [14] C. L. Hedley, M. J. Ambrose, and K. A. Pyke, *The physiology, genetics and nodulation of temperate legumes*, ch. Developing an improved model for the pea crop. London: Pitman, 1983.
- [15] J. Weiner and O. T. Solbrig, “The meaning and measurement of size hierarchies in plant populations,” *Oecologia*, vol. 61, no. 3, pp. 334–336, 1984.
- [16] J. Weiner, “Size hierarchies in experimental populations of annual plants,” *Ecology*, vol. 66, no. 3, pp. 743–752, 1985.
- [17] L. R. Benjamin and R. C. Hardwick, “Sources of variation and measures of variability in even-aged stands of plants,” *Annals of Botany*, pp. 757–778, 1986.
- [18] J. Weiner and S. C. Thomas, “Size variability and competition in plant monocultures,” *Oikos*, pp. 211–222, 1986.
- [19] R. Santos, “Size structure and inequality in a commercial stand of the seaweed *gelidium-sesquipedale*,” *Marine Ecology Progress Series*, vol. 119, no. 01-mar, pp. 253–263, 1995.
- [20] H. Nagashima, I. Terashima, and S. Katoh, “Effects of plant density on frequency distributions of plant height in *Chenopodium album* stands: analysis based on continuous monitoring of height-growth of individual plants,” *Annals of Botany*, vol. 75, no. 2, pp. 173–180, 1995.
- [21] A. B. Yastrebov, “Different types of heterogeneity and plant competition in monospecific stands,” *Oikos*, pp. 89–97, 1996.
- [22] C. Damgaard and J. Weiner, “Describing inequality in plant size or fecundity,” *Ecology*, pp. 1139–1142, 2000.
- [23] C. R. C. Vega and V. O. Sadras, “Size-dependent growth and the development of inequality in maize, sunflower and soybean,” *Annals of Botany*, vol. 91, no. 7, pp. 795–805, 2003.
- [24] T. Wyszomirski and J. Weiner, “Variation in local density results in a positive correlation between plant neighbor sizes,” *Am. Nat.*, 2009.

- [25] H. Nagashima and K. Hikosaka, “Plants in a crowded stand regulate their height growth so as to maintain similar heights to neighbours even when they have potential advantages in height growth,” *Annals of botany*, vol. 108, no. 1, pp. 207–214, 2011.
- [26] O. F. Falade and S. B. Oguntona, “Tree size structure of *Tectona grandis* (linn f.) stand in hilltop and valley-bottom of omo forest reserve,” *Environmental Science Proceedings*, vol. 13, p. 21, 2022.
- [27] T. Hiroi, “Dry matter economy of *Helianthus annuus* communities grown at varying densities on light intensities,” *Journal of the Faculty of Science, Tokyo University, Series*, no. 9, pp. 241–285, 1966.
- [28] W. J. Leverich and D. A. Levin, “Age-specific survivorship and reproduction in *Phlox drummondii*,” *The American Naturalist*, vol. 113, no. 6, pp. 881–903, 1979.
- [29] O. T. Solbrig, “Studies on the population biology of the *genus Viola*. ii. the effect of plant size on fitness in *Viola sororia*,” *Evolution*, pp. 1080–1093, 1981.
- [30] R. B. Bendel, S. S. Higgins, J. E. Teberg, and D. A. Pyke, “Comparison of skewness coefficient, coefficient of variation, and Gini coefficient as inequality measures within populations,” *Oecologia*, vol. 78, no. 3, pp. 394–400, 1989.
- [31] C.-K. Peng, S. V. Buldyrev, A. L. Goldberger, S. Havlin, F. Sciortino, M. Simons, and H. E. Stanley, “Long-range correlations in nucleotide sequences,” *Nature*, vol. 356, no. 6365, pp. 168–170, 1992.
- [32] C.-K. Peng, S. V. Buldyrev, S. Havlin, M. Simons, H. E. Stanley, and A. L. Goldberger, “Mosaic organization of dna nucleotides,” *Physical Review E*, vol. 49, no. 2, p. 1685, 1994.
- [33] S. Buldyrev, A. Goldberger, S. Havlin, R. Mantegna, M. Matsa, C.-K. Peng, M. Simons, and H. Stanley, “Long-range correlation properties of coding and noncoding dna sequences: Genbank analysis,” *Physical Review E*, vol. 51, no. 5, p. 5084, 1995.
- [34] E. G. Altmann, G. Cristadoro, and M. D. Esposti, “On the origin of long-range correlations in texts,” *Proceedings of the National Academy of Sciences*, vol. 109, no. 29, pp. 11582–11587, 2012.
- [35] D. S. Vieira, S. Picoli, and R. S. Mendes, “Robustness of sentence length measures in written texts,” *Physica A: Statistical mechanics and its applications*, vol. 506, pp. 749–754, 2018.

- [36] S. Picoli Jr, R. Mendes, E. Lenzi, and L. Malacarne, “Scale-invariant structure of size fluctuations in plants,” *Scientific Reports*, vol. 2, no. 1, p. 328, 2012.
- [37] A.-L. Barabási, *Bursts: The hidden patterns behind everything we do, from your e-mail to bloody crusades*. Plume; Reprint edition, 2011.
- [38] M. Karsai, H.-H. Jo, and K. Kaski, *Bursty human dynamics*. Springer International Publishing, 2018.
- [39] A. Corral, “Long-term clustering, scaling, and universality in the temporal occurrence of earthquakes,” *Physical Review Letters*, vol. 92, p. 108501, 2004.
- [40] A.-L. Barabási, “The origins of bursts and heavy tails in human dynamics,” *Nature*, vol. 435, pp. 207–11, 2005.
- [41] L. De Arcangelis, E. Godano, C. and Lippiello, and M. Nicodemi, “Universality in solar flare and earthquake occurrence,” *Physical Review Letters*, vol. 96, no. 5, p. 051102, 2006.
- [42] A. Paraschiv-Ionescu, E. Buchser, B. Rutschmann, , and K. Aminian, “Nonlinear analysis of human physical activity patterns in health and disease,” *Physical Review E*, vol. 77, p. 021913, 2008.
- [43] M. I. Bogachev, I. S. Kireenkov, , E. M. Nifontov, and A. Bunde, “Statistics of return intervals between long heartbeat intervals and their usability for online prediction of disorders,” *New Journal of Physics*, vol. 11, no. 6, p. 063036, 2009.
- [44] S. Picoli, M. d. Castillo-Mussot, H. V. Ribeiro, E. K. Lenzi, and R. S. Mendes, “Universal bursty behaviour in human violent conflicts,” *Scientific Reports*, vol. 4, no. 1, p. 4773, 2014.
- [45] W. Wang, N. Yuan, L. Pan, P. Jiao, W. Dai, G. Xue, and D. Liu, “Temporal patterns of emergency calls of a metropolitan city in China,” *Physica A: Statistical Mechanics and its Applications*, vol. 436, pp. 846–55, 2015.
- [46] P. Manshour, M. Anvari, N. Reinke, M. Sahimi, and M. R. R. Tabar, “Interoccurrence time statistics in fully-developed turbulence,” *Scientific Reports*, vol. 6, p. 27452, 2016.
- [47] L. Liu and F. HU, “Finescale clusterization intermittency of turbulence in the atmospheric boundary layer,” *Journal of the Atmospheric Sciences*, vol. 77, p. 2375, 2020.

- [48] J. F. Eichner, E. Koscielny-Bunde, A. Bunde, S. Havlin, and H.-J. Schellnhuber, “Power-law persistence and trends in the atmosphere: A detailed study of long temperature records,” *Physical Review E*, vol. 68, no. 4, p. 046133, 2003.
- [49] A. Bunde, J. F. Eichner, S. Havlin, and J. W. Kantelhardt, “Return intervals of rare events in records with long-term persistence,” *Physica A: Statistical Mechanics and its Applications*, vol. 342, no. 1-2, pp. 308–314, 2004.
- [50] J. F. Eichner, J. W. Kantelhardt, A. Bunde, and S. Havlin, “Statistics of return intervals in long-term correlated records,” *Physical Review E*, vol. 75, no. 1, p. 011128, 2007.
- [51] S. Hallerberg and H. Kantz, “Influence of the event magnitude on the predictability of an extreme event,” *Physical Review E*, vol. 77, no. 1, p. 011108, 2008.
- [52] P. Cizeau, Y. Liu, M. Meyer, C.-K. Peng, and H. E. Stanley, “Volatility distribution in the S&P500 stock index,” *Physica A: Statistical Mechanics and its Applications*, vol. 245, no. 3-4, pp. 441–445, 1997.
- [53] Y. Liu, P. Cizeau, M. Meyer, C.-K. Peng, and H. E. Stanley, “Correlations in economic time series,” *Physica A: Statistical Mechanics and its Applications*, vol. 245, no. 3-4, pp. 437–440, 1997.
- [54] Y. Liu, P. Gopikrishnan, H. E. Stanley, *et al.*, “Statistical properties of the volatility of price fluctuations,” *Physical Review E*, vol. 60, no. 2, p. 1390, 1999.
- [55] K. Yamasaki, L. Muchnik, S. Havlin, A. Bunde, and H. E. Stanley, “Scaling and memory in volatility return intervals in financial markets,” *Proceedings of the National Academy of Sciences*, vol. 102, no. 26, pp. 9424–9428, 2005.
- [56] F. Wang, K. Yamasaki, S. Havlin, and H. E. Stanley, “Scaling and memory of intraday volatility return intervals in stock markets,” *Physical Review E*, vol. 73, no. 2, p. 026117, 2006.
- [57] F. Wang, P. Weber, K. Yamasaki, S. Havlin, and H. Stanley, “Statistical regularities in the return intervals of volatility,” *The European Physical Journal B*, vol. 55, pp. 123–133, 2007.
- [58] S. Shinomoto, K. Shima, and J. Tanji, “Differences in spiking patterns among cortical neurons,” *Neural Computation*, vol. 15, no. 12, pp. 2823–42, 2003.
- [59] S. Shinomoto, K. Miura, and S. Koyama, “A measure of local variation of interspike intervals,” *BioSystems*, vol. 79, p. 67, 2005.

- [60] K.-I. Goh and A.-L. Barabási, “Burstiness and memory in complex systems,” *Europhysics Letters*, vol. 81, no. 4, p. 48002, 2008.
- [61] J. Yu, Y. Hu, M. Yu, and Z. Di, “Analyzing netizens’ view and reply behaviors on the forum,” *Physica A: Statistical Mechanics and its Applications*, vol. 389, pp. 3267–3273, 2010.
- [62] H.-H. Jo, M. Karsai, J. Kertész, and K. Kaski, “Circadian pattern and burstiness in mobile phone communication,” *New Journal of Physics*, vol. 14, p. 013055, 2012.
- [63] T. Yasseri, R. Sumi, A. Rung, A. Kornai, and J. Kertész, “Dynamics of conflicts in wikipedia,” *PloS One*, vol. 7, no. 6, p. e38869, 2012.
- [64] A. Paraschiv-Ionescu, E. Buchser, and K. Aminian, “Unraveling dynamics of human physical activity patterns in chronic pain conditions,” *Scientific Reports*, vol. 3, p. 2019, 2013.
- [65] C. Sanli and R. Lambiotte, “Local variation of hashtag spike trains and popularity in twitter,” *PLoS One*, vol. 10, no. 7, p. e0131704, 2015.
- [66] T. Aoki, T. Takaguchi, R. Kobayashi, and R. Lambiotte, “Input-output relationship in social communications characterized by spike train analysis,” *Physical Review E*, vol. 94, p. 042313, 2016.
- [67] M. Schleiss and J. Smith, “Two simple metrics for quantifying rainfall intermittency: The burstiness and memory of interamount times,” *Journal of Hydrometeorology*, vol. 17, pp. 421–36, 2016.
- [68] D. Yan, Z.-W. Wei, X.-P. Han, and B. Wang, “Empirical analysis on the human dynamics of blogging behavior on github,” *Physica A: Statistical Mechanics and its Applications*, vol. 465, pp. 775 – 81, 2017.
- [69] S. Picoli, E. S. D. Santos, P. P. Deprá, and R. S. Mendes, “Quantifying postural sway dynamics using burstiness and interevent time distributions,” *The European Physical Journal B*, vol. 92, no. 7, p. 156, 2019.
- [70] S. Picoli, G. Bombo, E. S. Santos, P. P. Deprá, and R. S. Mendes, “Characterizing postural sway signals by the analysis of zero-crossing patterns,” *Physica A: Statistical Mechanics and its Applications*, vol. 596, p. 127160, 2022.
- [71] H. Cramér, *Mathematical methods of statistics*, vol. 43. Princeton university press, 1999.

- [72] E.-K. Kim and H.-H. Jo, "Measuring burstiness for finite event sequences," *Physical Review E*, vol. 94, p. 032311, 2016.
- [73] D. D. Reed and H. E. Burkhart, "Spatial autocorrelation of individual tree characteristics in loblolly pine stands," *Forest Science*, vol. 31, p. 575, 1985.
- [74] C. Wilson and J. Gurevitch, "Plant size and spatial pattern in a natural population of *Myosotis micrantha*," *Journal of Vegetation Science*, vol. 6, p. 847, 1995.
- [75] Ipea - Instituto de Pesquisa Econômica Aplicada, 2023. <http://www.ipea.gov.br>.
- [76] E. Limpert, W. A. Stahel, and M. Abbt, "Log-normal distributions across the sciences: Keys and clues," *BioScience*, vol. 51, p. 341, 2001.
- [77] X. W. et al., "Dynamic plant height qtl revealed in maize through remote sensing phenotyping using a highthroughput unmanned aerial vehicle (uav)," *Scientific Reports*, vol. 9, p. 3458, 2019.
- [78] D. P. Aikman, *Modeling of growth and competition in plant monocultures in Individual-based models and approaches in ecology*. Chapman and Hall/CRC, 1992.
- [79] S. E. Park, L. R. Benjamin, and A. R. Watkinson, "The theory and application of plant competition models: An agronomic perspective," *Annals of Botany*, vol. 92, p. 741, 2003.
- [80] A. W. Drake, *Fundamentals of applied probability theory*. Mcgraw-Hill College, 1967.
- [81] E. G. Altmann, E. C. da Silva, and I. L. Caldas, "Recurrence time statistics for finite size intervals," *Chaos*, vol. 14, no. 4, pp. 975–81, 2004".