

---

Processos de ramificação com aplicações em  
biologia

*Cristel Ecaterin Vera Tapia*

---

SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito:

Assinatura: \_\_\_\_\_

## Processos de ramificação com aplicações em biologia

**Cristel Ecaterin Vera Tapia**

*Orientador: Prof. Dr. Pablo Martín Rodríguez*

Dissertação apresentada ao Instituto de Ciências Matemáticas e de Computação - ICMC-USP e ao Departamento de Estatística – DEs-UFSCar, como parte dos requisitos para obtenção do título de Mestre em Estatística – Interinstitucional de Pós-Graduação em Estatística. *EXEMPLAR DE DEFESA*

**USP/UFSCar – São Carlos**

**Fevereiro de 2015**

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi  
e Seção Técnica de Informática, ICMC/USP,  
com os dados fornecidos pelo(a) autor(a)

T172p

Tapia, Cristel Ecaterin Vera

Processos de ramificação com aplicações em  
biologia / Cristel Ecaterin Vera Tapia; orientador  
Pablo Martín Rodríguez. -- São Carlos, '2015.  
63 p.

Dissertação (Mestrado - Programa  
Interinstitucional de Pós-graduação em Estatística) --  
Instituto de Ciências Matemáticas e de Computação,  
Universidade de São Paulo, '2015.

1. Processos de ramificação. I. Rodríguez, Pablo  
Martín, orient. II. Título.

Com todo meu amor, aos meus pais, que fizeram tudo na vida para que eu pudesse alcançar meus sonhos, por me dar força e coragem nos momentos difíceis, para vocês sempre meu coração e minha gratidão.



# Agradecimentos

---

---

Quero agradecer, em primeiro lugar, a Deus, por estar comigo em todos os momentos, sendo meu refúgio e fortaleza nos momentos mais difíceis durante esta caminhada.

Agradeço especialmente, aos meus pais: Hipolito e Ruth, por sua capacidade de acreditar em mim. Mãe seu carinho e apoio incondicional, foi que me deram, força e coragem para seguir. Pai sua presença sempre foi importante na minha vida. Aos meus irmãos: Maureem e Peter, pelo carinho e confiança. Ao meu tio Godo pela preocupação e incentivo. E a toda minha família.

Ao professor doutor Pablo Martín Rodríguez, meu orientador, por todo o conhecimento transmitido, dedicação, paciência e principalmente pelo apoio e preocupação durante esta jornada.

Aos colegas e professores do mestrado, por tudo o que com eles aprendi e por partilharem a construção do meu estudo.

A todos os meus amigos, pela amizade, por todos os momentos de divertimento que passamos juntos, pela ajuda e também pelas horas de estudo compartilhadas com alguns deles.

A CAPES, pelo apoio financeiro.

A todos muito obrigada.



Estudamos a teoria de processos de ramificação de Galton-Watson a tempo discreto e as ferramentas probabilísticas necessárias para analisá-los. Na primeira etapa, demos um tratamento básico de processos de ramificação, isto é, assumimos que as partículas são iguais e que a distribuição do número de descendentes diretos de cada partícula é sempre a mesma. Também incluímos resultados sobre o comportamento limite para os casos subcrítico, crítico e supercrítico. Posteriormente, consideramos uma generalização das características assumidas na etapa anterior, baseada em processos de Galton-Watson em meios variáveis, onde a distribuição do número de descendentes diretos de uma partícula varia de geração em geração. Estudamos e provamos teoremas limite. Finalmente, discutimos dois modelos de processos de ramificação binária com aplicações em biologia.





# Abstract

---

---

We study the theory of Galton-Watson branching processes at discrete time and the necessary probabilistic tools to analyze them. In the first stage, was given a basic treatment of the branching processes, that is, it was assumed that all the particles are equal and that the distribution of the number of offspring produced by a particle is always the same. Also were included some results about the asymptotic behavior for the subcritical, critical and supercritical cases. Afterwards, was considered a generalization of the characteristics assumed in the previous stage, based on Galton-Watson processes in varying environments, where the distribution of offspring produced by a particle varies from generation to generation. Were studied and proved limit theorems. Finally, were discussed two models of binary branching processes with applications in biology.



# Sumário

---

---

<b>Dedicatória</b>	<b>i</b>
<b>1 Introdução</b>	<b>1</b>
<b>2 Processos ramificação de Galton-Watson</b>	<b>5</b>
2.1 Definições e resultados preliminares . . . . .	5
2.1.1 O espaço de probabilidade . . . . .	7
2.1.2 Momentos . . . . .	9
2.1.3 Extinção do processo . . . . .	12
2.2 Resultados Assintóticos . . . . .	16
2.2.1 Caso Supercrítico . . . . .	16
2.2.2 Caso Subcrítico . . . . .	21
2.2.3 Caso Crítico . . . . .	24
<b>3 Processos de Galton-Watson em meios variáveis</b>	<b>29</b>
3.1 Noções preliminares . . . . .	29
3.2 Resultados . . . . .	31
<b>4 Aplicações</b>	<b>43</b>
4.1 Probabilidade de resistência aos medicamentos . . . . .	43
4.1.1 Descrição do modelo . . . . .	44
4.2 Probabilidade de resistência com duas drogas . . . . .	50
4.2.1 Descrição do modelo . . . . .	50
<b>A Martingale</b>	<b>57</b>



---

## Introdução

---

Nesse capítulo mencionamos alguns fatos acerca do surgimento dos processos de ramificação. A seguinte descrição está baseada nos livros de Allen (2003) e Gut (2009). Pode-se dizer que o estudo de processos de ramificação, um tipo especial de processo estocástico, foi iniciado na década de 1840, por Irenée-Jules Bienaymé, um probabilista e estatístico e foi avançado na década de 1870 com o trabalho do reverendo Henry William Watson, um clérigo e matemático, e Francis Galton, um biometricista, primo de Charles Darwin, que estudou o decaimento da nobreza Inglesa e outros nomes de famílias de distinção. Em 1873, Galton propôs o seguinte problema na “Educational Times”, respeito à sobrevivência do sobrenome de uma família.

**Problema 4001:** Considera-se uma população de tamanho  $N$ , da qual estamos apenas interessados nos adultos do sexo masculino, cada um tendo sobrenome diferente. A regra da população é tal que, em cada geração,  $p_0$  por cento dos homens adultos não tem nenhum filho do sexo maculino que atinge a idade adulta,  $p_1$  por cento dos homens adultos tem um filho que alcança a idade adulta,  $p_2$  tem dois filhos que alcançam a idade adulta e assim por diante até  $p_5$  por cento dos homens adultos que tem cinco filhos do sexo masculino que atingem a idade adulta. Queremos encontrar: (1) Que proporção dos sobrenomes chegará a ser extinta após  $r$  gerações; (2) Quantos casos haverá, de  $m$  pessoas tendo o mesmo sobrenome?

Em outras palavras, Galton propôs estudar a sobrevivência do sobrenome de uma família quando este é passado do pai para o filho, assumindo que cada homem tem  $j$  filhos com probabilidade  $p_j$ ,  $j = 0, 1, \dots, 5$ .

A primeira tentativa de uma resposta foi dada por Watson, que reformulou o problema em termos de funções geradoras, mas ele concluiu incorretamente que a probabilidade de extinção do sobrenome é sempre 1. Por sorte, o erro que ele cometeu foi algébrico e seus argumentos de funções geradoras são usados até hoje para encontrar a solução correta do problema proposto por Galton e muitas das suas variantes. Mais tarde esses processos foram chamados de processos de ramificação de Galton-Watson. Para mais detalhes sobre o surgimento e a evolução histórica desta teoria pode-se ver o livro de Harris (2002). Durante os últimos anos muita atenção tem-sido dedicada a esses processos, devido ao interesse crescente nas aplicações destes em Biologia, Física, etc. Por exemplo, no estudo do crescimento de populações, a propagação de epidemias ou rumores e a cadeia de reações nucleares, entre outros.

Para ilustrar este tipo de processos, imaginemos partículas que podem gerar partículas adicionais do mesmo tipo de acordo com cópias independentes de uma variável aleatória discreta, com valores no conjunto dos inteiros não negativos. Assim, um conjunto inicial de partículas, o qual chamamos de geração 0, gera certo número de partículas que são chamadas de primeira geração, seus filhos são chamados de segunda geração e assim por diante. Alguns exemplos conhecidos na literatura (ver Gut (2009)) e que ilustram aplicações dos processos de ramificação são:

**Sobrenome de uma família:** Assuma que uma mulher e um homem são casados e que a mulher muda o sobrenome dela para o de seu marido (como nos tempos antigos), de modo que o sobrenome da família sobreviverá só através dos filhos homens. Se os filhos nascem de acordo ao processo descrito anteriormente, este representa um processo de ramificação.

**Reações nucleares:** A fissão provocada pela colisão de neutróns resulta em um número de novos neutróns, os mesmos que quando colidem produzem outros novos neutróns e assim por diante. Desta forma, temos que o número de neutróns cresce como um processo de ramificação.

**Filas de clientes em espera:** Um determinado cliente que chega em um servidor desocupado pode ser visto como uma partícula inicial, os clientes chegando quando ele está sendo atendido representam os filhos deste (descendentes da partícula inicial) e assim por diante. O processo continuará enquanto houver pessoas na fila, esperando para serem atendidas.

No capítulo 2 estudamos noções e definições básicas necessárias para o melhor desenvolvimento do modelo original dos processos de Galton-Watson. Também provamos teoremas limite para os casos subcrítico, crítico e supercrítico. No capítulo 3 apresentamos uma generalização das características de processos de Galton-Watson que foram assumidas no capítulo 2, baseada em processos de ramificação em meios variáveis. Finalmente, no capítulo 4 estudamos dois modelos de processos de ramificação binária com aplicações em biologia, revisamos a formulação matemática e os principais resultados destes.





---

# Processos ramificação de Galton-Watson

---

Iniciamos este capítulo com as noções básicas de processo estocástico e cadeia de Markov, que serão necessárias para um melhor desenvolvimento do modelo de Galton-Watson. Os livros de Karlin e Taylor (1998) e Schinazi (1999) são ótimas referências.

## 2.1 Definições e resultados preliminares

**Definição 2.1.1.** *Um processo estocástico a tempo discreto é uma sequência de variáveis aleatórias  $(X_n)_{n \geq 0}$ ,  $n \in \mathbb{N}$ , definidas no mesmo espaço de probabilidade e com valores em algum conjunto enumerável  $S$ , chamado espaço de estados. Neste trabalho assumiremos  $S = \mathbb{Z}_+ := \{0, 1, 2, \dots\}$ .*

O nome “tempo discreto” é consequência de que com frequência o subíndice  $n$  representa unidades de tempo e o processo  $(X_n)_{n \geq 0}$  representa a evolução de um determinado fenômeno aleatório ao longo desse tempo. Assim, suponha que observamos a evolução de uma nova doença em uma população. Se a variável aleatória  $Y_n$  representa o número de infectados pela doença no  $n$ -ésimo dia após de o primeiro registro da infecção, a sequência  $(Y_n)_{n \geq 1}$  é um processo estocástico a tempo discreto.

**Definição 2.1.2.** *Uma cadeia de Markov a tempo discreto é um processo estocástico a tempo discreto  $(X_n)_{n \geq 0}$ , tal que,*

$$\mathbb{P}[X_{n+1} = j \mid X_0 = i_0, X_1 = i_1, X_2 = i_2, \dots, X_{n-1} = i_{n-1}, X_n = i] = \mathbb{P}[X_{n+1} = j \mid X_n = i], \quad (2.1)$$

para todo  $n \geq 1$  e para todo subconjunto de estados  $\{i_0, i_1, \dots, i_{n-1}, i, j\}$ .

Em palavras, uma cadeia de Markov é um processo estocástico para o qual afirmações probabilísticas sobre o futuro dependem somente do estado presente do processo.

**Observação 2.1.1.** *Da definição anterior podemos ver que, para todo  $i, j \in S$  e para todo  $n \geq 0$ , as probabilidades*

$$p(i, j) := \mathbb{P}[X_{n+1} = j \mid X_n = i],$$

chamadas probabilidades de transição, são fundamentais para definir a cadeia de Markov.

De modo geral podemos pensar num processo de Galton-Watson como segue. Suponha que temos partículas que geram novas partículas de acordo a uma variável aleatória discreta  $X$  com valores no conjunto  $\mathbb{Z}_+$  e função de distribuição de probabilidades dada por

$$\mathbb{P}[X = k] = p_k, \quad k = 0, 1, 2, \dots \quad (2.2)$$

Dizemos que uma partícula no tempo  $n = 0$  constitui a geração 0 e as partículas que nascem no tempo  $n$  constituem a  $n$ -ésima geração do processo,  $n \geq 1$ . Seja  $Z_n$  a variável aleatória que conta o número de partículas na  $n$ -ésima geração. Agora podemos formalizar a definição de processo de ramificação de Galton-Watson.

**Definição 2.1.3.** *Seja  $X$  uma variável aleatória discreta com distribuição de probabilidades dada por (2.2). Chamamos processo de ramificação de Galton-Watson a cadeia de Markov  $(Z_n)_{n \geq 0}$  com valores no conjunto  $\{0, 1, 2, \dots\}$  e probabilidades de transição dadas por*

$$p(i, j) = \begin{cases} \mathbb{P}[Z_{n+1} = j \mid Z_n = i] = \mathbb{P}\left[\sum_{r=1}^i X_r = j\right] & \text{para } i \geq 1 \text{ e } j \geq 0, \\ 0, & \text{para } i = 0 \text{ e } j > 0, \\ 1, & \text{para } i = 0 \text{ e } j = 0, \end{cases} \quad (2.3)$$

onde  $X_1, X_2, \dots, X_i$  são variáveis aleatórias *i.i.d.* à variável aleatória  $X$ .

A definição de processo de ramificação de Galton-Watson foi dada em função das suas probabilidades de transição. Pela Observação 2.1.1 sabemos que ele está completamente definido a partir destas probabilidades e da distribuição do seu estado inicial  $Z_0$ . Para simplificar a exposição assumamos que  $Z_0 \equiv 1$ , isto é,  $\mathbb{P}[Z_0 = 1] = 1$ .

### 2.1.1 O espaço de probabilidade

Embora a maior parte do nosso estudo dos processos de ramificação será diretamente baseado nas propriedades de funções geradoras de probabilidades é importante ter uma ideia da estrutura probabilística destes. Isto é, definir explicitamente o conjunto  $\Omega$  de eventos elementares, a classe aditiva  $\mathcal{F}$  de subconjuntos de  $\Omega$ , chamados de eventos e a medida de probabilidade  $\mathbb{P}$ , para os eventos de  $\mathcal{F}$ . No que segue desta subsecção, será feita uma descrição de  $(\Omega, \mathcal{F}, \mathbb{P})$ , o espaço de probabilidade dos processos de ramificação de Galton-Watson. Mais detalhes podem ser encontrados em Otter (1949).

Seja  $\mathbb{Z}_+^* := \{1, 2, \dots\}$ , então, para cada  $n \in \mathbb{Z}_+^*$ , existe um conjunto  $V_n$  de objetos ou partículas  $v_{i_1 i_2 \dots i_n}$  chamados vértices, isto é,

$$V_n = \{v_{i_1 i_2 \dots i_n}\}; \quad (i_1, i_2, \dots, i_n) \in \{\mathbb{Z}_+^*\}^n.$$

Daqui, denote a coleção de todos os vértices, por

$$\mathbf{V} = \bigcup_{n=0}^{\infty} V_n,$$

onde o vértice  $v_0$  pode ser interpretado como a partícula inicial ou progenitor e por exemplo, o vértice  $v_{135}$ , como o quinto filho do terceiro filho do primeiro filho do progenitor. Se  $\mathcal{C}$  é um subconjunto de  $\mathbf{V}$  e se  $i_1, i_2, \dots, i_{n+m}$  são tais que, cada  $v_{i_1 i_2 \dots i_n}, v_{i_1 i_2 \dots i_n i_{n+1}}, \dots, v_{i_1 i_2 \dots i_n i_{n+1} \dots i_{n+m}}$  pertence a  $\mathcal{C}$ , então esse conjunto de vértices é chamado um caminho de  $v_{i_1 i_2 \dots i_n}$  para  $v_{i_1 i_2 \dots i_{n+m}}$  em  $\mathcal{C}$  e  $m \geq 0$  é o comprimento do caminho ou distância de  $v_{i_1 i_2 \dots i_n}$  para  $v_{i_1 i_2 \dots i_{n+m}}$ . Se  $m = 1$ , o caminho pode ser chamado de segmento. Assim, defina-se  $\mathbf{N}(\mathcal{C}, v)$ , para  $v \in \mathcal{C} \subset \mathbf{V}$ , como o número de segmentos desde  $v \in \mathcal{C}$ . Que pode ser interpretado como o número de filhos da partícula  $v$ .

**Definição 2.1.4.** Um subconjunto  $\mathcal{T}$  de  $\mathbf{V}$  é chamado de árvore se e somente se

- (1)  $v_0 \in \mathcal{T}$ ;
- (2)  $v_{i_1 i_2 \dots i_n} \in \mathcal{T}$  implica que  $v_{i_1 i_2 \dots i_{n-1} j} \in \mathcal{T}$ , para cada  $j \in \{1, \dots, i_n\}$ ;
- (3)  $\mathbf{N}(\mathcal{T}, v) < \infty$ , para  $v \in \mathcal{T}$ .

Então,  $\Omega$  é a coleção de todas as árvores. Se uma árvore contém um número finito de vértices, então é chamada de árvore finita, no outro sentido é chamada de árvore infinita. Seja  $\mathfrak{F}$  o conjunto de todas as árvores finitas e  $\mathfrak{I}$  o conjunto de todas as árvores infinitas. Para cada  $\mathcal{T}' \in \mathfrak{F}$ , seja  $f_1(\mathcal{T}'), f_2(\mathcal{T}'), \dots, f_m(\mathcal{T}')$  seus  $m$  pontos finais. (Se para cada

$v \in \mathcal{T} \in \Omega$ , temos que  $\mathbf{N}(\mathcal{T}, v) = 0$ , então  $v$  é chamado ponto final de  $\mathcal{T}$ ). Então defina-se, para  $\mathcal{T}' \in \mathfrak{F}$  e  $\mathbf{k} = (k_1, k_2, \dots, k_m) \in \{\mathbb{Z}_+\}^m$ ,

$$[\mathcal{T}', \mathbf{k}] = \{\mathcal{T} \mid \mathcal{T} \geq \mathcal{T}', \mathbf{N}(\mathcal{T}, f_i(\mathcal{T}')) = k_i, i = 1, 2, \dots, m\}$$

e chamamos  $[\mathcal{T}', \mathbf{k}]$ , uma vizinhança. Assim, para cada  $\mathcal{T} \in [\mathcal{T}', \mathbf{k}]$ , dizemos que  $[\mathcal{T}', \mathbf{k}]$  é uma vizinhança de  $\mathcal{T}$ .

Seja  $\mathfrak{C}$  uma classe de subconjuntos de  $\Omega$ , tal que,

- (i) O conjunto vazio pertence a  $\mathfrak{C}$ ;
- (ii) Qualquer conjunto  $\mathcal{S}$ , pertence a  $\mathfrak{C}$  se e somente se existe  $\mathcal{T} \in \mathfrak{F}$  e um conjunto  $R$  não vazio, chamado conjunto retângulo, isto é,  $R = \{A_1 \times A_2 \times \dots \times A_m; A_i \subset \mathbb{Z}_+\}$ , tal que

$$\mathcal{S} = \sum_{\mathbf{k} \in R} [\mathcal{T}, \mathbf{k}],$$

onde a coleção  $([\mathcal{T}, \mathbf{k}])_{\mathbf{k} \in R}$ , é chamada uma  $\mathfrak{C}$ -partição de  $\mathcal{S}$  e  $\mathcal{T}$  é chamada de geradora da  $\mathfrak{C}$ -partição.

Deve-se considerar que é possível obter apenas um número finito de  $\mathfrak{C}$ -partições para um  $S \in \mathfrak{C}$ , pois só um número finito de árvores podem ser geradoras e além disso, existe apenas uma  $\mathfrak{C}$ -partição para cada geradora. Considere também que, se  $S, S' \in \mathfrak{C}$ , então  $S \cap S' \in \mathfrak{C}$ . Agora definimos a função de medida  $\mathbb{P}$  para os conjuntos em  $\mathfrak{C}$  com as seguintes propriedades,

- (i)  $\mathbb{P}[S] \geq 0$ , para todo  $S \in \mathfrak{C}$ ;
- (ii)  $\mathbb{P}[\emptyset] = 0$ ;
- (iii)  $\mathbb{P}[S] = \sum_{\mathbf{k} \in R} \mathbb{P}([\mathcal{T}, \mathbf{k}])$

Tomando  $R = \mathbb{Z}_+$  temos  $\mathbb{P}[\Omega] = 1$ , nessas condições podemos dizer que  $\mathbb{P}$  é uma medida de probabilidade.

Agora, seja a classe  $\mathfrak{G}$ , cujos elementos são conjuntos que podem ser formados por uniões finitas de elementos disjuntos de  $\mathfrak{C}$ , isto é,  $\mathfrak{G}(\mathfrak{C}) = \{\bigcup_{i=1}^n S_i : S_1, \dots, S_n \in \mathfrak{C}\}$ , então  $\mathfrak{G}$  é uma álgebra de conjuntos, em outras palavras a álgebra gerada pela classe  $\mathfrak{C}$ . Dado que  $\mathbb{P}$  é  $\sigma$ -aditiva sobre  $\mathfrak{C}$ , pelo primeiro teorema de extensão (ver Resnick (2013)), existe uma única medida de probabilidade  $\mathbb{P}_1$  sobre  $\mathfrak{G}$  que estende  $\mathbb{P}$ . Daqui pelo segundo teorema de extensão (ver Resnick (2013)), segue que  $\mathbb{P}_1$  pode ser estendida para

uma única medida de probabilidade  $\mathbb{P}_2$  sobre a  $\sigma$ -álgebra de Borel  $\mathcal{F}$ , isto é, a menor  $\sigma$ -álgebra gerada pela classe  $\mathfrak{C}$ . Por simplicidade escreveremos  $\mathbb{P} = \mathbb{P}_2$ . Dessa forma, fica definido o espaço de probabilidade  $(\Omega, \mathcal{F}, \mathbb{P})$ .

## 2.1.2 Momentos

Uma importante ferramenta no estudo de algumas propriedades dos processos de ramificação são os momentos, obtidos pela função geradora de probabilidades.

**Definição 2.1.5.** *Seja  $X$  uma variável aleatória tomando valores no conjunto dos inteiros não negativos e com função de distribuição dada em (2.2). A função geradora de probabilidade (f.g.p)  $\phi(t)$  de  $X$  é definida por*

$$\phi(t) = E(t^X) = \sum_{i=0}^{\infty} p_i t^i, \quad |t| \leq 1 \quad (2.4)$$

e suas iterações são definidas como

$$\phi_0(t) = t, \quad \phi_1(t) = \phi(t) \quad \text{e} \quad \phi_{n+1}(t) = \phi[\phi_n(t)], \quad \text{para } n = 1, 2, \dots$$

**Proposição 2.1.2.** *Sejam  $\phi(t)$  e  $\phi_{(n)}(t)$  as funções geradoras de probabilidades das variáveis aleatórias  $X$  e  $Z_n$ , respectivamente. Então*

$$\phi_{(n)}(t) = \phi_n(t) \quad (2.5)$$

onde

$$\phi_n(t) = (\phi \circ \phi \circ \dots \circ \phi)(t).$$

*Demonstração.* Segue por indução sob  $n$ . Para  $n = 1$ ,

$$\phi_{(1)}(t) = \mathbb{E}[t^{Z_1}] = \mathbb{E}[t^X] = \phi(t).$$

Agora supomos que (2.5) é válida para  $n - 1$  e provaremos sua validade para  $n$ . Pela definição de f.g.p da variável aleatória  $Z_n$ , temos

$$\phi_{(n)}(t) = \sum_{i=0}^{\infty} \mathbb{P}[Z_n = i] t^i. \quad (2.6)$$

Por outro lado,

$$\begin{aligned}\mathbb{P}[Z_n = i] &= \sum_{j=0}^{\infty} \mathbb{P}[Z_n = i \mid Z_{n-1} = j] \mathbb{P}[Z_{n-1} = j] \\ &= \sum_{j=0}^{\infty} \mathbb{P}\left[\sum_{l=1}^j X_l = i \mid Z_{n-1} = j\right] \mathbb{P}[Z_{n-1} = j],\end{aligned}$$

então, por independência das variáveis aleatórias  $X_l$ , para  $l = 1, 2, \dots, j$  e  $Z_{n-1}$ , segue que

$$\mathbb{P}[Z_n = i] = \sum_{j=0}^{\infty} \mathbb{P}\left[\sum_{l=1}^j X_l = i\right] \mathbb{P}[Z_{n-1} = j]. \quad (2.7)$$

Inserindo (2.7) em (2.6) e como as somas são absolutamente convergentes para  $|t| \leq 1$ , podemos trocar a ordem do somatório, assim temos

$$\begin{aligned}\phi_{(n)}(t) &= \sum_{i=0}^{\infty} \left\{ \sum_{j=0}^{\infty} \mathbb{P}\left[\sum_{l=1}^j X_l = i\right] \mathbb{P}[Z_{n-1} = j] \right\} t^i \\ &= \sum_{j=0}^{\infty} \left\{ \sum_{i=0}^{\infty} \mathbb{P}\left[\sum_{l=1}^j X_l = i\right] t^i \right\} \mathbb{P}[Z_{n-1} = j].\end{aligned} \quad (2.8)$$

Usando, em seguida, a independência das variáveis aleatórias  $X_l$ , para  $l = 1, 2, \dots, j$ , segue

$$\sum_{i=0}^{\infty} \mathbb{P}\left[\sum_{l=1}^j X_l = i\right] t^i = \mathbb{E}\left[t^{\sum_{l=1}^j X_l}\right] = \prod_{l=1}^j \mathbb{E}[t^{X_l}] = (\phi(t))^j.$$

Logo, (2.8) se escreve

$$\phi_{(n)}(t) = \sum_{j=0}^{\infty} (\phi(t))^j \mathbb{P}[Z_{n-1} = j] = \phi_{n-1}(\phi(t)) = \phi_n(t),$$

o que prova (2.5). □

Consideramos agora os momentos de  $Z_n$ . Note que

$$m := \sum_{i=1}^{\infty} ip_i = \mathbb{E}[X] = \mathbb{E}[Z_1] = \phi'(1) \quad (2.9)$$

é o número médio de descendentes de uma partícula ou média de reprodução e

$$\sigma^2 = \sum_{i=1}^{\infty} i^2 p_i - m^2 = \text{Var}[X] = \text{Var}[Z_1] = \phi''(1) + \phi'(1) - \phi'(1)^2 \quad (2.10)$$

é a variância do número de descendentes.

**Proposição 2.1.3.** *A média e a variância de  $Z_n$  são dadas por*

$$\begin{aligned} E[Z_n] &= m^n, \\ \text{Var}[Z_n] &= \begin{cases} \frac{\sigma^2 m^{n-1}(m^n - 1)}{m - 1}, & m \neq 1; \\ n\sigma^2, & m = 1; \end{cases} \end{aligned} \quad (2.11)$$

para todo  $n \geq 1$ , respectivamente.

*Demonstração.* Temos por propriedades de esperança condicional, que

$$\mathbb{E}[Z_n] = \mathbb{E}[\mathbb{E}(Z_n \mid Z_{n-1})],$$

isto é,

$$\mathbb{E}[Z_n] = \sum_{i=0}^{\infty} \mathbb{E}[Z_n \mid Z_{n-1} = i] \mathbb{P}(Z_{n-1} = i),$$

mas dado que  $Z_n = \sum_{j=1}^{Z_{n-1}} X_j$ , onde  $X_1, X_2, \dots$  são variáveis aleatórias i.i.d à variável aleatória  $X$ , então

$$\begin{aligned} \mathbb{E}[Z_n] &= \sum_{i=0}^{\infty} \mathbb{E} \left[ \sum_{j=1}^i X_j \mid Z_{n-1} = i \right] \mathbb{P}(Z_{n-1} = i) \\ &= \sum_{i=0}^{\infty} \mathbb{E} \left[ \sum_{j=1}^i X_j \right] \mathbb{P}(Z_{n-1} = i) = m \mathbb{E}[Z_{n-1}]. \end{aligned}$$

Portanto,

$$\mathbb{E}[Z_n] = m \mathbb{E}[Z_{n-1}] = \dots = m^{n-1} \mathbb{E}[Z_1]$$

e como  $\mathbb{E}[Z_1] = \mathbb{E}[X] = m$ , segue que  $\mathbb{E}[Z_n] = m^n$ . Para a variância, por propriedades da função geradora de probabilidades e pela regra da cadeia, temos que

$$\begin{aligned} \phi'_n(s) &= (\phi_{n-1}(\phi(s)))' \\ &= \phi'_{n-1}(\phi(s)) \phi'(s). \\ \phi''_n(s) &= \phi''_{n-1}(\phi(s)) (\phi'(s))^2 + \phi'_{n-1}(\phi(s)) \phi''(s). \end{aligned} \quad (2.12)$$



Dado (2.10), segue

$$\begin{aligned}\phi'_n(1) &= \sigma_n^2 + m^{2n} - m^n \\ \phi''_{n-1}(1) &= \sigma_{n-1}^2 + m^{2n-2} - m^{n-1}.\end{aligned}\tag{2.13}$$

Agora, tomando  $s = 1$  em (2.12) e usando (2.10) e (2.13), temos

$$\begin{aligned}\phi''_n(1) &= \phi''_{n-1}(1)(\phi(1))^2 + \phi'_{n-1}(1)\phi''(1) \\ \sigma_n^2 + m^{2n} + m^n &= (\sigma_{n-1}^2 + m^{2n-2} + m^{n-1})m^2 + m^{n-1}(\sigma^2 + m^2 - m) \\ \sigma_n^2 &= m^2\sigma_{n-1}^2 + m^{n-1}\sigma^2.\end{aligned}$$

Isso leva a

$$\sigma_n^2 = m^{n-1}\sigma^2(1 + m + m^2 + \dots + m^{n-1}).\tag{2.14}$$

Daqui, se  $m = 1$

$$\sigma_n^2 = n\sigma^2.$$

Se  $m \neq 1$

$$\sigma_n^2 = m^{n-1}\sigma^2 \frac{1 - m^n}{1 - m} = m^{n-1}\sigma^2 \frac{m^n - 1}{m - 1}.$$

□

### 2.1.3 Extinção do processo

O evento de extinção é de muito interesse quando estudamos ou aplicamos processos de ramificação.

**Definição 2.1.6.** *Seja  $(Z_n)_{n \geq 0}$  um processo de Galton-Watson. Chamamos de extinção do processo ao evento*

$$\mathcal{E} = \bigcup_{n \geq 1} [Z_n = 0]\tag{2.15}$$

Observe que

$$\mathbb{P}(\mathcal{E}) = \lim_{n \rightarrow \infty} \mathbb{P} \left[ \bigcup_{k=1}^n [Z_k = 0] \right] = \lim_{n \rightarrow \infty} \mathbb{P}[Z_n = 0] = \lim_{n \rightarrow \infty} \phi_n(0),\tag{2.16}$$

onde na segunda igualdade temos que  $[Z_n = 0] \subset [Z_{n+1} = 0]$ , para todo  $n \geq 1$ . Isto é,  $\{Z_n = 0\}_{n \geq 1}$  é uma seqüência crescente de eventos. Denotamos a probabilidade de extinção por  $q := \mathbb{P}(\mathcal{E})$ .

O resultado a seguir é fundamental ao pleno desenvolvimento do análise da probabilidade de extinção dos processos de ramificação de Galton-Watson.

**Teorema 2.1.4.** *Seja  $p_0 + p_1 < 1$ . A probabilidade de extinção  $q$  do processo  $(Z_n)_{n \geq 0}$  é a menor raiz não negativa da equação  $t = \phi(t)$ . Além disso, se  $m = E(X) < \infty$  temos que*

(i) *se  $m \leq 1$  então  $q = 1$ ;*

(ii) *se  $m > 1$  então  $q < 1$ .*

*Demonstração.* A prova será feita em duas partes. Na primeira parte provaremos que  $q = \phi(q)$ . Então, definamos para todo  $n \geq 0$ ,

$$q_n := \mathbb{P}[Z_n = 0].$$

Daqui temos pela segunda igualdade em (2.16) que

$$\lim_{n \rightarrow \infty} q_n = q. \quad (2.17)$$

Por outro lado, condicionando sobre a primeira geração, temos que

$$\begin{aligned} q_n &= \sum_{i=0}^{\infty} \mathbb{P}[Z_n = 0 \mid Z_1 = i] \mathbb{P}[Z_1 = i] \\ &= \sum_{i=0}^{\infty} \mathbb{P} \left[ \sum_{j=1}^i Z_{n-1}^j = 0 \right] \mathbb{P}[Z_1 = i], \end{aligned}$$

onde  $Z_{n-1}^j$ , para  $j = 1, \dots, i$  na última igualdade, representa o número de descendentes da  $j$ -ésima partícula na  $(n-1)$ -ésima geração, dado que na primeira geração existiram  $i$  partículas. Daqui, por independência das variáveis aleatórias  $Z_{n-1}^j$  para  $j = 1, \dots, i$  e pela definição de  $q_{n-1}$  e  $\phi$ , segue

$$\begin{aligned} q_n &= \sum_{i=0}^{\infty} \mathbb{P} \left[ \bigcap_{j=1}^i [Z_{n-1}^j = 0] \right] \mathbb{P}[Z_1 = i] \\ &= \sum_{i=0}^{\infty} (\mathbb{P}[Z_{n-1} = 0])^i \mathbb{P}[Z_1 = i] \\ &= \phi(q_{n-1}). \end{aligned} \quad (2.18)$$

Assim, de (2.17) e (2.18) e pela continuidade de  $\phi$ , resulta que  $q = \phi(q)$ .

Na segunda parte, analisaremos as soluções de  $t = \phi(t)$ . Primeiro observemos que para  $t \geq 0$  vale que as derivadas são positivas, isto é,  $\phi'(t), \phi''(t) \geq 0$ , portanto a função

é crescente e concava para cima, assim a reta  $y = t$  encontra a curva  $y = \phi(t)$  em no máximo dois pontos do plano cartesiano, digamos que de coordenadas iguais a  $t_0$  e a  $t_1$ . Como  $\phi(1) = 1$ , as possibilidades para as raízes da equação  $t = \phi(t)$  são: a)  $t_0 < 1$  e  $t_1 = 1$ , b)  $t_0 = t_1 = 1$  c)  $t_0 = 1$  e  $t_1 > 1$ . Os tres casos podem ser vistos na figura 2.1. Ademais, pela esperança via função geradora de probabilidades, a inclinação da reta

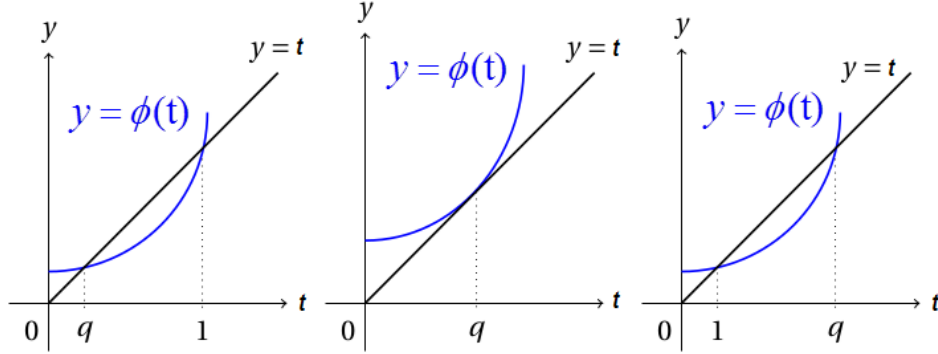


Figura 2.1

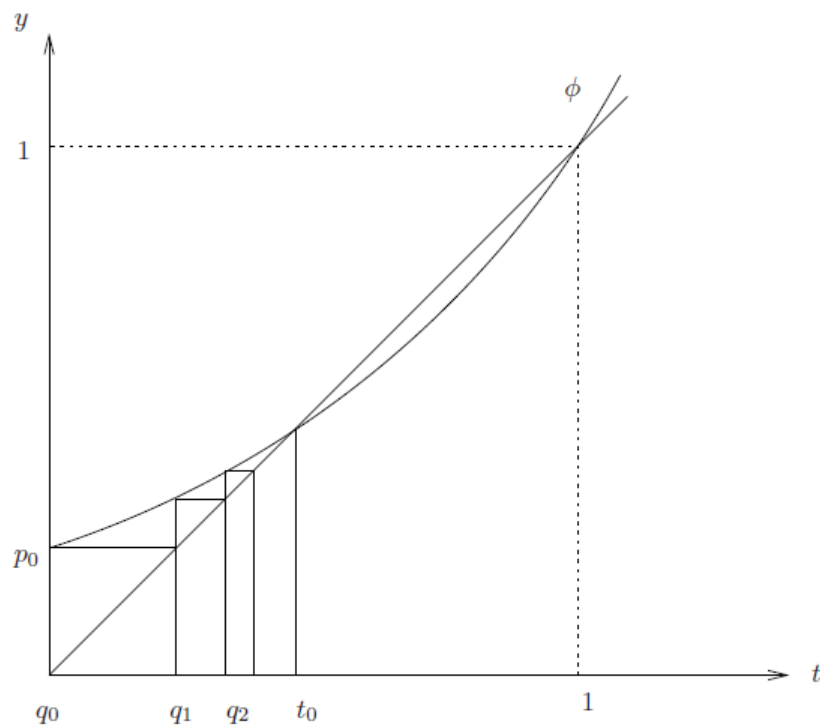
tangente em  $(1, 1)$  na curva  $y = \phi(t)$  é  $m = \mathbb{E}[X]$ , de modo que no caso a)  $m > 1$ , no caso b)  $m = 1$  e no caso c)  $m < 1$ . Dado que probabilidade não pode ser maior do que 1, dos casos b) e c) podemos concluir que  $q = 1$ . Isso prova o item (i) do Teorema.

Agora analisando o caso a), a partir da equação (2.18) podemos ver que  $q_0 = 0$ ,  $q_1 = \phi(q_0) = p_0$ ,  $q_2 = \phi(q_1) = \phi(p_0), \dots$ , segundo a construção desses valores, na figura 2.2 podemos ver que os pontos  $(q_i, \phi(q_i))$  sempre estarão acima da linha  $y = t$ . Portanto eles convergem para a primeira interseção de  $y = \phi(t)$  com  $y = t$ . Isto é, para a raiz  $q = t_0 < 1$ . Isso mostra que o item (ii) do Teorema é válido.  $\square$

**Definição 2.1.7.** Um processo de Galton-Watson é chamado *supercrítico* se  $m > 1$ , *subcrítico* se  $m < 1$  e *crítico* se  $m = 1$ .

**Observação 2.1.5.** Suponha que  $p_0 + p_1 = 1$ . Definamos  $T = \inf[n \geq 1 : Z_n = 0]$ , o tempo em que o processo atinge o estado 0. Então seja  $\mathcal{E} := [T < \infty]$  o evento de extinção do processo. Assim,  $q := \mathbb{P}[T < \infty]$  é a probabilidade de extinção. Notemos que,

$$[T = \infty] = \bigcap_{n=1}^{\infty} [T > n],$$



**Figura 2.2**

onde  $[T > n + 1] \subset [T > n]$ , para todo  $n \geq 1$ . Isto é,  $\{T > n\}_{n \geq 1}$  é uma sequência decrescente de eventos, então resulta que,

$$1 - q = [T = \infty] = \bigcap_{n=1}^{\infty} [T > n] = \lim_{n \rightarrow \infty} \mathbb{P}[T > n].$$

Dado que  $\mathbb{P}[T > n] = p_1^n$ . Logo

$$p_1 = 1 \text{ implica } q = 0;$$

$$p_1 < 1 \text{ implica } q = 1.$$

No que segue, apresentamos um exemplo que ilustra a teoria estudada, em especial o Teorema 2.1.4.

**Exemplo 2.1.1.** Os processos de ramificação de divisão binária são de muita importância em biologia. Esse tipo de processos são tais que as partículas podem representar células que dividem-se em duas, com probabilidade  $p$  ou morrem com probabilidade  $1 - p$ , para

algum  $0 < p < 1$ . Notemos que neste caso,

$$\begin{aligned}\phi(t) &= \mathbb{P}[X = 0] + \mathbb{P}[X = 2]t^2 \\ &= 1 - p + pt^2\end{aligned}$$

e portanto  $m = \phi'(1) = 2p$ . Então, pelo Teorema 2.1.4, segue que a população de células chega a extinguir-se com probabilidade 1, quando  $p \leq 1/2$ . Para  $p > 1/2$  resolvemos a equação  $t = \phi(t)$ , da qual resultam as soluções  $t_1 = 1$  e  $t_2 = (1 - p)/p < 1$ . Isto é,  $q = (1 - p)/p$  neste caso.

## 2.2 Resultados Assintóticos

Na seção anterior vimos que o processo de Galton-Watson chega a se extinguir ou cresce indefinidamente, nesse sentido, estamos interessados em saber qual é o comportamento assintótico quando o processo sobrevive e qual seria a taxa de decrescimento se este chega a se extinguir. O nosso principal objetivo nesta seção é provar teoremas limite para os casos supercrítico, subcrítico e crítico. Para maiores detalhes ver o livro de (Athreya e Ney, 1972).

### 2.2.1 Caso Supercrítico

Uma vez dada a classificação dos processos de Galton-Watson, a qual depende do número médio de descendentes  $m$  e dado que a probabilidade de sobrevivência resulta positiva unicamente no caso supercrítico, isto é, quando  $m > 1$ , uma natural pergunta é acerca do comportamento assintótico do processo, se este sobrevive. Nesta caso, temos como principal resultado que a sequência  $Z_n/m^n$  converge para uma variável aleatória não negativa  $W$ , quase certamente.

Começamos com um importante resultado de convergência de Martingales (ver Definição A.0.1 no Apêndice), que será útil na prova do teorema limite para este caso.

**Teorema 2.2.1.** *Seja  $(X_n)_{n \geq 0}$  um martingale não negativo com respeito a filtração  $(\mathcal{F}_n)_{n \geq 0}$ . Então,*

- (i)  $\lim_{n \rightarrow \infty} X_n$  existe quase certamente e é finito.

(ii) Se o martingale é  $L^2$  limitado (isto é,  $\sup_n \mathbb{E}[X_n^2] < \infty$ ), então existe  $X$ , tal que

$${}^1X_n \xrightarrow{L^2} X.$$

*Demonstração.* Ver Apêndice. □

Agora, vamos concentrar a nossa atenção na variável aleatória

$$W_n := \frac{Z_n}{m^n}.$$

Seja  $\mathcal{F}_n = \sigma(Z_0, Z_1, \dots, Z_n)$ , a  $\sigma$ -álgebra gerada por  $Z_0, Z_1, \dots, Z_n$ , isto é, a menor filtragem para a qual  $Z_n$  é adaptada, então temos que

$$\mathbb{E}[Z_{n+1} | \mathcal{F}_n] = \mathbb{E}[Z_{n+1} | Z_0, Z_1, \dots, Z_n] = \mathbb{E}\left[\sum_{k=1}^{Z_n} X_k | Z_n\right] = mZ_n. \quad (2.19)$$

Onde a segunda e terceira igualdade em (2.19) segue do fato que  $(Z_n)_{n \geq 0}$  é uma cadeia de Markov e  $Z_{n+1} = \sum_{j=1}^{Z_n} X_j$ , em que  $X_1, X_2, \dots$  são variáveis aleatórias i.i.d à variável aleatória  $X$ . Logo dividindo ambos lados de (2.19) pela constante  $m^{n+1}$ , segue que

$$\mathbb{E}[W_{n+1} | \mathcal{F}_n] = \frac{1}{m^{n+1}} \mathbb{E}[Z_{n+1} | \mathcal{F}_n] = \frac{1}{m^{n+1}} mZ_n = W_n. \quad (2.20)$$

Portanto, dado que  $\mathbb{E}[W_n] = 1$ ,  $W_n \in \mathcal{F}_n$  para todo  $n \geq 0$  e por (2.20), segue que  $(W_n)_{n \geq 0}$  é um martingale com respeito à filtragem  $(\mathcal{F}_n)_{n \geq 0}$ .

Por conseguinte, obtemos o seguinte resultado.

**Teorema 2.2.2.** *Se  $0 < m < \infty$ , então existe uma variável aleatória  $W$  tal que*

$$\lim_{n \rightarrow \infty} W_n = W,$$

*quase certamente.*

*Demonstração.* A prova segue diretamente do fato de  $W_n$  ser um martingale e pelo item (i) do Teorema 2.2.1 □

Sabemos que se  $m \leq 1$ ,  $Z_n \rightarrow 0$  quando  $n \rightarrow \infty$ , com probabilidade 1, isto é,  $\mathbb{P}[W = 0] = 1$ . Porém, para ter que  $\mathbb{P}[W = 0] < 1$ , quando  $m > 1$ , devemos estabelecer uma condição adicional de variância finita. O seguinte resultado mostra esse fato.

---

<sup>1</sup>Dizemos que a sequência de variáveis aleatórias  $(X_n)_{n \geq 0}$  converge para a variável aleatória  $X$  em  $L^2$  ( $X_n \xrightarrow{L^2} X$ ) se  $\lim_{n \rightarrow \infty} \mathbb{E}[|X_n - X|^2] = 0$ .

**Teorema 2.2.3.** Se  $m > 1$ ,  $\sigma^2 < \infty$ , e  $Z_0 \equiv 1$ , então

$$(i) \lim_{n \rightarrow \infty} \mathbb{E} [|W_n - W|^2] = 0;$$

$$(ii) \mathbb{E}[W] = 1, \quad \text{Var}[W] = \sigma^2/(m^2 - m);$$

$$(iii) \mathbb{P}[W = 0] = q.$$

*Demonstração.* (i) Pela definição dos momentos de  $Z_n$  temos que

$$\begin{aligned} \mathbb{E}[Z_n^2] &= \text{Var}[Z_n] + \mathbb{E}[Z_n]^2 \\ &= \frac{\sigma^2 m^{n-1}(m^n - 1)}{m - 1} + m^{2n}. \end{aligned}$$

Consequentemente,

$$\begin{aligned} \mathbb{E}[W_n^2] &= \frac{1}{m^{2n}} \mathbb{E}[Z_n^2] = \frac{\sigma^2 m^n (m^n - 1)}{m^{2n}(m^2 - m)} + 1 \\ &= 1 + \frac{\sigma^2}{m^2 - m} (1 - m^{-n}), \end{aligned} \quad (2.21)$$

como  $\mathbb{E}[W_n^2]$  é uma função crescente em  $n$ , então resulta que

$$\sup_n \mathbb{E}[W_n^2] = \lim_{n \rightarrow \infty} \mathbb{E}[W_n^2] = 1 + \frac{\sigma^2}{(m^2 - m)} < \infty, \quad (2.22)$$

o que significa que  $W_n$  é  $L^2$  limitado, então pelo item (ii) do Teorema (2.2.1) existe uma variável aleatória  $W'$ , tal que  $W_n \xrightarrow{L^2} W'$ , mas pelo Teorema (2.2.2) temos que  $W_n \rightarrow W$  quase certamente, então resulta que  $W = W'$  quase certamente. Pois, convergência em  $L^2$  e convergência quase certa, ambas implicam convergência em probabilidade e daí que o limite em probabilidade é único. Logo  $W_n \xrightarrow{L^2} W$  e isso mostra que (i) é válido.

(ii) Como  $W_n$  converge para  $W$  em  $L^2$ , isso implica que

$$\begin{aligned} 1 &= \lim_{n \rightarrow \infty} \mathbb{E}[W_n] = \mathbb{E}[W], \\ 1 + \frac{\sigma^2}{(m^2 - m)} &= \lim_{n \rightarrow \infty} \mathbb{E}[W_n^2] = \mathbb{E}[W^2]. \end{aligned}$$

E consequentemente

$$\begin{aligned}
\frac{\sigma^2}{(m^2 - m)} &= \lim_{n \rightarrow \infty} \text{Var}[W_n] = \lim_{n \rightarrow \infty} (\mathbb{E}[W_n^2] - (\mathbb{E}[W_n])^2) \\
&= \mathbb{E}[W^2] - \mathbb{E}[W]^2 \\
&= \text{Var}[W].
\end{aligned}$$

(iii) Dado que  $W_n$  converge em  $L^2$ ,  $\mathbb{E}[W_n] = 1$  para todo  $n$ , daqui que  $\mathbb{E}[W] = 1$  e  $W$  não pode ser constantemente zero, isto é,  $\mathbb{P}[W = 0] < 1$ . Se  $r := \mathbb{P}[W = 0]$ , temos

$$r = \mathbb{P} \left[ \lim_{n \rightarrow \infty} \frac{Z_n}{m^n} = 0 \right] = \sum_{k=0}^{\infty} \mathbb{P} \left[ \lim_{n \rightarrow \infty} \frac{Z_n}{m^n} = 0 \mid Z_1 = k \right] \mathbb{P}[Z_1 = k] \quad (2.23)$$

A distribuição da quantidade de descendentes na  $n$ -ésima geração, sabendo que na primeira geração existiram  $k$  partículas, é igual à distribuição da quantidade de descendentes na  $(n - 1)$ -ésima geração de  $k$  famílias independentes, assim

$$\begin{aligned}
\mathbb{P} \left[ \lim_{n \rightarrow \infty} \frac{Z_n}{m^n} = 0 \mid Z_1 = k \right] &= \mathbb{P} \left[ \lim_{n \rightarrow \infty} \frac{1}{m} \left( \sum_{j=1}^k \frac{Z_{n-1}^j}{m^{n-1}} \right) = 0 \right] \\
&= \mathbb{P} \left[ \lim_{n \rightarrow \infty} \left( \sum_{j=1}^k W_{n-1}^j \right) = 0 \right] \\
&= \mathbb{P} \left[ \bigcap_{j=1}^k \left( \lim_{n \rightarrow \infty} W_{n-1}^j = 0 \right) \right] \\
&= \prod_{j=1}^k \mathbb{P} \left[ \lim_{n \rightarrow \infty} W_{n-1}^j = 0 \right] = (\mathbb{P}[W = 0])^k.
\end{aligned}$$

Então, em (2.23) temos

$$r = \mathbb{P}[W = 0] = \sum_{k=0}^{\infty} (\mathbb{P}[W = 0])^k p_k = \phi(\mathbb{P}[W = 0]) = \phi(r). \quad (2.24)$$

e portanto, deve ser  $r = q$  sobre  $[0, 1)$ . Pois a equação  $\phi(t) = t$  tem uma única solução sobre  $[0, 1)$ , quando  $m > 1$ .  $\square$

Em vista do Teorema 2.2.3, podemos usar a distribuição de  $W$  para estudar a distribuição de  $Z_n$ , para  $n$  grande.



**Teorema 2.2.4.** *A transformada de Laplace da distribuição de  $W$ ,  $L_W(v) = \mathbb{E}[e^{-vW}]$ , satisfaz a equação*

$$L_W(v) = \phi_{Z_1} \left[ L_W \left( \frac{v}{m} \right) \right], \quad (2.25)$$

*chamada equação de Abel.*

*Demonstração.* Podemos escrever a transformada de Laplace

$$L_{Z_n}(u) := \mathbb{E}[e^{-uZ_n}] = \phi_{Z_n}[e^{-u}].$$

Daqui

$$\begin{aligned} L_{Z_{n+1}}(u) &= \phi_{Z_{n+1}}(e^{-u}) \\ &= \phi_{Z_1}[\phi_{Z_n}(e^{-u})] \\ &= \phi_{Z_1}[L_{Z_n}(u)]. \end{aligned} \quad (2.26)$$

Logo, a transformada de Laplace da distribuição de  $W_n$  é

$$L_{W_n}(u) = \mathbb{E}[e^{-uW_n}] = \mathbb{E} \left[ e^{\left( -\frac{u}{m^n} Z_n \right)} \right] = L_{Z_n} \left( \frac{u}{m^n} \right).$$

Consequentemente,

$$L_{W_n}(um^n) = \mathbb{E}[e^{-um^n W_n}] = \mathbb{E}[e^{-uZ_n}] = L_{Z_n}(u).$$

Agora,

$$\begin{aligned} L_{W_{n+1}}(um^{n+1}) &= \mathbb{E}[e^{-uZ_{n+1}}] \\ &= \sum_{j=0}^{\infty} \mathbb{E}[e^{-uZ_{n+1}} \mid Z_1 = j] \mathbb{P}[Z_1 = j] \\ &= \sum_{j=0}^{\infty} \mathbb{E} \left[ e^{(-u \sum_{i=1}^j Z_n^i = j)} \mid Z_1 = j \right] \mathbb{P}[Z_1 = j]. \end{aligned}$$

Por outro lado,

$$\begin{aligned}
\mathbb{E} \left[ e^{(-u \sum_{i=1}^j Z_n^i = j)} \middle| Z_1 = j \right] &= \mathbb{E} \left[ \prod_{i=1}^j e^{(-u Z_n^i)} \right] = (\mathbb{E} [e^{(-u Z_n)}])^j \\
&= (\mathbb{E} [e^{(-u m^n W_n)}])^j \\
&= (L_{W_n}(u m^n))^j.
\end{aligned}$$

Portanto,

$$\begin{aligned}
L_{W_{n+1}}(u m^{n+1}) &= \sum_{j=0}^{\infty} \mathbb{E} \left[ e^{(-u \sum_{i=1}^j Z_n^i = j)} \middle| Z_1 = j \right] \mathbb{P}[Z_1 = j] \\
&= \sum_{j=0}^{\infty} [L_{W_n}(u m^n)]^j \mathbb{P}[Z_1 = j] = \phi_{Z_1}[L_{W_n}(u m^n)]. \tag{2.27}
\end{aligned}$$

Agora fazendo a mudança de variáveis  $v = u m^{n+1}$  em (2.27), obtemos

$$L_{W_{n+1}}(v) = \phi_{Z_1} \left[ L_{W_n} \left( \frac{v}{m} \right) \right].$$

Como  $W_n \rightarrow W$  em distribuição,  $L_{W_n}(v) \rightarrow L_W(v)$  e o limite o qual é a transformada de Laplace da distribuição da variável aleatória  $W$ , satisfaz a equação (2.25).  $\square$

### 2.2.2 Caso Subcrítico

No caso subcrítico, quando  $m < 1$ , o processo chega a extinguir-se com probabilidade 1. Nesse sentido, pode-se obter uma distribuição limite não trivial de  $Z_n$  condicionando o processo para a não extinção.

O resto  $r$ , no desenvolvimento em series de Taylor em torno do ponto 1,

$$\phi(t) = 1 - m(1 - t) + r(t)(1 - t), \quad 0 \leq t < 1 \tag{2.28}$$

desempenha um papel importante na análise de processos de Galton-Watson subcríticos. Claramente,

$$r(t) = m - \frac{1 - \phi(t)}{1 - t} \tag{2.29}$$

$$r(0) = m - (1 - p_0) \geq 0$$

$$r(q) = m - 1 > 0, \quad \text{se } q < 1$$

$$r'(t) \leq 0, \quad 0 \leq t < 1.$$

Portanto,  $r$  é uma função decrescente desde  $[0, 1)$  em  $[0, m]$ . Agora, se substituirmos  $t$  por  $\phi_j(t)$  em (2.28), temos

$$\begin{aligned}\phi_{j+1}(t) &= 1 - (m - r(\phi_j(t)))(1 - \phi_j(t)) \\ \frac{1 - \phi_{j+1}(t)}{1 - \phi_j(t)} &= m \left(1 - \frac{r(\phi_j(t))}{m}\right).\end{aligned}$$

O produto da anterior igualdade, para  $0 \leq j < n$  é

$$\frac{1 - \phi_n(t)}{1 - t} = m^n \prod_{j=0}^{n-1} \left(1 - \frac{r(\phi_j(t))}{m}\right). \quad (2.30)$$

A partir de (2.29) vemos que  $0 \leq \frac{r}{m} \leq 1$ , assim  $\frac{m^{-n}(1 - \phi_n(t))}{1 - t}$  decresce para algum limite  $\varphi(t) \geq 0$ . Particularmente, temos

$$\mathbb{P}[Z_n > 0] = 1 - \phi_n(0) \rightarrow m^n \varphi(0), \quad \text{quando } n \rightarrow \infty.$$

Aqui,  $\varphi(0) > 0$  se e somente se

$$\sum_{j=1}^{\infty} r(\phi_j(0)) < \infty.$$

Pois tomando limite quando  $n \rightarrow \infty$  em (2.30), por uma conhecida relação entre convergência de somas e produtos, temos que o produto  $\prod_{j=1}^{\infty} \left(1 - \frac{r(\phi_j(0))}{m}\right)$  e a soma  $\sum_{j=1}^{\infty} \frac{r(\phi_j(0))}{m}$  convergem junto. Daqui, um resultado (ver Jagers (1975), p. 28) mostra que

$$\varphi(0) > 0 \quad \text{se e somente se} \quad \sum_{j=1}^{\infty} p_j j \log j < \infty.$$

Nessas condições, segue o importante resultado conhecido como Teorema de Yaglom.

**Teorema 2.2.5.** (*Teorema de Yaglom*) *Se  $m < 1$  e  $p_0 < 1$ , temos que*

$$\lim_{n \rightarrow \infty} \mathbb{P}[Z_n = j \mid Z_n > 0] = b_j,$$

*existe para  $j \in \mathbb{N}$ , com*

$$\sum_{j=1}^{\infty} b_j = 1 \quad e \quad \sum_{j=1}^{\infty} j b_j = \frac{1}{\varphi(0)} < \infty \quad \text{se e somente se} \quad \sum_{j=1}^{\infty} p_j j \log j < \infty$$

e se  $g(t) = \sum_{j=1}^{\infty} b_j t^j$ , então  $g$  satisfaz a equação

$$g(\phi(t)) = mg(t) + 1 - m.$$

**Observação 2.2.6.** A função geradora de probabilidade condicional é dada por

$$\mathbb{E}[t^{Z_n} \mid Z_n > 0] = \sum_{j=0}^{\infty} \mathbb{P}[Z_n = j \mid Z_n > 0] t^j = \sum_{j=1}^{\infty} \frac{\mathbb{P}[Z_n = j]}{\mathbb{P}[Z_n > 0]} t^j = \frac{\phi_n(t) - \phi_n(0)}{1 - \phi_n(0)}. \quad (2.31)$$

*Demonstração.* Teorema 2.2.5 Supondo que  $m < 1$ , seja

$$g_n(t) := \mathbb{E}[t^{Z_n} \mid Z_n > 0] = \frac{\phi_n(t) - \phi_n(0)}{1 - \phi_n(0)} = 1 - \frac{1 - \phi_n(t)}{1 - \phi_n(0)}. \quad (2.32)$$

Daqui, usando a relação (2.30) na última igualdade de (2.32), temos que

$$g_n(t) = 1 - (1 - t) \prod_{j=0}^{n-1} (1 - r(\phi_j(t))/m)(1 - r(\phi_j(0))/m)^{-1}. \quad (2.33)$$

Uma vez que  $\phi_j(t) \geq \phi_j(0)$  e  $r$  é não crescente, os termos do produto em (2.33) são maiores do que um e assim existe uma função  $g$  tal que  $g_n \downarrow g$ . Pelo fato que  $g_n$  são f.g.p, podemos escrever  $g$  como

$$g(t) = \sum_{j=1}^{\infty} b_j t^j.$$

Note que  $g(0) = 0$ . Então, quando  $n \rightarrow \infty$

$$g_n(\phi_j(0)) = 1 - \frac{1 - \phi_j(\phi_n(0))}{1 - \phi_n(0)} \rightarrow 1 - m^j,$$

que pode ser perto de um, para uma escolha arbitrariamente grande de  $j$ . Portanto  $g(1) = 1$  e além disso

$$\begin{aligned} \sum_{j=1}^{\infty} j b_j &= g'(1) = \lim_{j \rightarrow \infty} \frac{1 - g(\phi_j(0))}{1 - \phi_j(0)} \\ &= \lim_{j \rightarrow \infty} \frac{m^j}{1 - \phi_j(0)} = \frac{1}{\varphi(0)} \end{aligned}$$

e

$$g_n(\phi) = 1 - \frac{1 - \phi_{n+1}}{1 - \phi_{n+1}(0)} \frac{1 - \phi(\phi_n(0))}{1 - \phi_n(0)} \rightarrow 1 - (1 - g)m.$$

□

### 2.2.3 Caso Crítico

No caso crítico temos que  $Z_n \rightarrow 0$  com probabilidade 1, quando  $n \rightarrow \infty$ . Estamos interessados em analisar a taxa de convergência para zero. Como resultado temos que o decréscimo do processo é exponencial.

**Lema 2.2.7.** *Se  $m = 1$  e  $\sigma^2 < \infty$ , então*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \left[ \frac{1}{1 - \phi_n(t)} - \frac{1}{1 - t} \right] = \frac{\sigma^2}{2} \quad (2.34)$$

*uniformemente, para  $0 \leq t \leq 1$ .*

*Demonstração.* Vamos usar o desenvolvimento em série de Taylor só para os três primeiros termos, isto é

$$\phi(t) = \phi(1) + \phi'(1)(t - 1) + \frac{\phi''(1)}{2}(t - 1)^2 + r(t)(t - 1)^2,$$

para alguma função  $r$  tal que  $\lim_{t \uparrow 1} r(t) = 0$ , considerando  $\phi(1) = 1$ ;  $\phi'(1) = 1$  e  $\phi''(1) = \sigma^2$ , temos

$$\begin{aligned} \phi(t) &= 1 + (t - 1) + \frac{\sigma^2}{2}(t - 1)^2 + r(t)(t - 1)^2 \\ \phi(t) - t &= \frac{\sigma^2}{2}(t - 1)^2 + r(t)(t - 1)^2 \\ \frac{\phi(t) - t}{(t - 1)^2} &= \frac{\sigma^2}{2} + r(t)(t - 1)^2. \end{aligned}$$

Agora tomando limite quando  $t \uparrow 1$ , resulta

$$\lim_{t \uparrow 1} \frac{\phi(t) - t}{(t - 1)^2} = \frac{\sigma^2}{2} =: a.$$

Seja

$$\varepsilon(t) = a - \frac{\phi(t) - t}{(1 - t)^2}$$

e note que

$$\varepsilon(t) > 0 \quad e \quad \varepsilon(t) \downarrow 0 \quad \text{quando} \quad t \uparrow 1. \quad (2.35)$$

Portanto

$$\frac{\phi(t) - t}{(t - 1)^2} \leq a. \quad (2.36)$$

Agora seja

$$\delta(t) = a - \left[ \frac{1}{1 - \phi(t)} - \frac{1}{1 - t} \right], \quad (2.37)$$

mas  $t \leq \phi(t)$  implica que

$$\frac{1}{1 - \phi(t)} - \frac{1}{1 - t} \geq \frac{\phi(t) - t}{(t - 1)^2},$$

assim

$$\delta(t) \leq \varepsilon(t). \quad (2.38)$$

Sustituindo  $t$  por  $\phi_i(t)$  em (2.38), somando sobre  $i$ , e usando (2.35), vemos que

$$an - \left[ \frac{1}{1 - \phi_n(t)} - \frac{1}{1 - t} \right] = \sum_{i=0}^{n-1} \delta(\phi_i(t)) \leq \sum_{i=0}^{n-1} \varepsilon[\phi_i(t)] \leq \sum_{i=0}^{n-1} \varepsilon[\phi_i(0)] = r(n), \quad (2.39)$$

onde a última igualdade se deve ao fato que  $\varepsilon[\phi_i(0)] \rightarrow 0$ , quando  $i \rightarrow \infty$ . Agora para obter uma cota na outra direção nós vemos que

$$\delta(s) = \left( a \frac{1 - \phi(s)}{1 - s} - \frac{\phi(s) - s}{(1 - s)^2} \right) \left( \frac{1 - \phi(s)}{1 - s} \right)^{-1} \geq a \frac{s - \phi(s)}{1 - \phi(s)},$$

por (2.35). Logo pela de definição de  $\varepsilon(s)$ , segue que

$$\frac{\phi(s) - s}{(1 - s)} = (1 - s)(a - \varepsilon(s)).$$

Daqui

$$\begin{aligned} \delta(s) &\geq -a \frac{\phi(s) - s}{1 - s} \frac{1 - s}{1 - \phi(s)} = -a(1 - s)(a - \varepsilon(s)) \frac{1 - s}{1 - \phi(s)} \geq -a^2(1 - s) \left[ \frac{1 - s}{1 - \phi(s)} \right] \\ &\geq -a^2(1 - s) \frac{1}{1 - \phi(0)}. \end{aligned}$$

A última desigualdade é pelo fato que  $\frac{1 - s}{1 - \phi(s)}$  é decrescente e positiva em  $[0, 1]$ . Portanto temos

$$\sum_{k=0}^{n-1} \delta[\phi_k(t)] \geq -\frac{a^2}{1 - \phi(0)} \sum_{k=0}^{n-1} [1 - \phi_k(t)] \geq \frac{-a^2}{1 - \phi(0)} \sum_{k=0}^{n-1} [1 - \phi_k(0)] = r(n). \quad (2.40)$$

Dividindo (2.39) e (2.40) por  $n$  e fazendo  $n \rightarrow \infty$ , obtemos (2.34).  $\square$

Baseado no Lema 2.2.7, a taxa na qual o processo crítico se torna extinto pode ser estimada. O comportamento limite da probabilidade de não extinção,  $\mathbb{P}[Z_n > 0]$ , é encontrado fixando  $t = 0$  no Lema 2.2.7.

**Teorema 2.2.8.** *Se  $m = 1$  e  $\sigma^2 < \infty$ , então*

$$\lim_{n \rightarrow \infty} n\mathbb{P}[Z_n > 0] = \frac{2}{\sigma^2}. \quad (2.41)$$

*Demonstração.* Tomando  $t = 0$  no Lema 2.2.7, temos

$$\lim_{n \rightarrow \infty} \frac{1}{n} \left( \frac{1}{1 - \phi_n(0)} - 1 \right) = \frac{\sigma^2}{2}.$$

Dado que  $\mathbb{P}[Z_n > 0] = 1 - \phi_n(0)$ , segue que

$$n\mathbb{P}[Z_n > 0] = n(1 - \phi_n(0)) = \left( \frac{1}{n} \left( \frac{1}{1 - \phi_n(0)} - 1 \right) + \frac{1}{n} \right)^{-1}. \quad (2.42)$$

Agora tomando limite quando  $n \rightarrow \infty$  em (2.42), temos

$$\lim_{n \rightarrow \infty} n\mathbb{P}[Z_n > 0] = \frac{2}{\sigma^2}.$$

□

Usando o fato

$$1 = \mathbb{E}[Z_n] = \mathbb{E}[Z_n | Z_n > 0]\mathbb{P}[Z_n > 0] + 0\mathbb{P}[Z_n = 0],$$

temos que o valor esperado do processo, condicionado para não extinção é dado por

$$\mathbb{E}[Z_n | Z_n > 0] = \frac{1}{\mathbb{P}[Z_n > 0]}$$

e portanto do Teorema 2.2.8, segue que

$$\mathbb{E}[Z_n | Z_n > 0] \sim \frac{n\sigma^2}{2}, \quad \text{quando } n \rightarrow \infty.$$

Isso sugere que a distribuição limite poderia existir para o processo normalizado e condicional,  $\left[ \frac{Z_n}{n} \middle| Z_n > 0 \right]$ . Isto leva ao seguinte resultado.

**Teorema 2.2.9.** *Se  $m = 1$  e  $\sigma^2 < \infty$ , então*

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[ \frac{Z_n}{n} > z \middle| Z_n > 0 \right] = \exp \left( -\frac{2z}{\sigma^2} \right). \quad (2.43)$$

*Demonstração.* Podemos provar que

$$\lim_{n \rightarrow \infty} \mathbb{E}[e^{-\alpha(Z_n/n)} \mid Z_n > 0] = \frac{1}{1 + \frac{\alpha\sigma^2}{2}}, \quad (2.44)$$

isso é suficiente pois o lado direito de (2.44) é a transformada de Laplace do lado direito de (2.43), mas a esperança condicional de (2.44) é igual a

$$\frac{\phi_n(e^{-\alpha/n}) - \phi_n(0)}{1 - \phi_n(0)} = 1 - \frac{\{n[1 - \phi_n(0)]\}^{-1}}{\{n[1 - \phi_n(e^{-\alpha/n})]\}^{-1}}. \quad (2.45)$$

Logo, pelo Lema 2.2.7, segue

$$\lim_{n \rightarrow \infty} \frac{1}{n[1 - \phi_n(0)]} = \frac{\sigma^2}{2}$$

e usando a convergencia uniforme

$$\lim_{n \rightarrow \infty} \frac{1}{n[1 - \phi_n(e^{-\alpha/n})]} = \frac{\sigma^2}{2} + \lim_{n \rightarrow \infty} \frac{1}{n[1 - e^{-\alpha/n}]} = \frac{\sigma^2}{2} + \frac{1}{\alpha}.$$

Substituindo em (2.45) implica (2.44) e assim, obtemos (2.43).  $\square$





---

# Processos de Galton-Watson em meios variáveis

---

No capítulo anterior foram consideradas algumas características específicas acerca dos processos de Galton-Watson, isto é, assumimos que todas as partículas são iguais e que a distribuição do número de descendentes diretos de uma partícula é sempre a mesma. Neste capítulo apresentamos uma generalização dessas características, baseada em processos de Galton-Watson em meios variáveis.

## 3.1 Noções preliminares

Um processo de Galton-Watson em meio variável é um processo de ramificação para o qual a distribuição do número de descendentes diretos de uma partícula depende da geração à qual a partícula pertence. Formalmente, se  $Z_n$  é o número de descendentes na  $n$ -ésima geração, então

$$Z_{n+1} = \sum_{i=1}^{Z_n} X_{ni},$$

onde  $X_{n1}, X_{n2}, \dots$  são cópias independentes e identicamente distribuídas de uma variável aleatória  $X_n$ , cuja distribuição depende da geração  $n$ . Supomos que  $\mathbb{P}[X_n = k] = p_{nk}$ , para  $k, n = 0, 1, 2, \dots$  e que  $\mu_n := \mathbb{E}[X_n]$ . Denotamos a função geradora de probabilidade

desta variável aleatória por  $\varphi_n$ . Isto é

$$\varphi_n(t) = \mathbb{E}[t^{X_n}] = \sum_{j=0}^{\infty} p_{nj} t^j, \quad |t| \leq 1.$$

Consequentemente, pode-se verificar que  $Z_n$  tem função geradora de probabilidade

$$f_n = \varphi_0 \circ \varphi_1 \circ \dots \circ \varphi_{n-1} \tag{3.1}$$

Em efeito, por indução sob  $n$ , temos que para  $n = 1$ ,

$$f_1(t) = \mathbb{E}[t^{Z_1}]$$

Agora, suponha que (3.1) é válida para  $n - 1$  e provemos a veracidade para  $n$ , assim pela definição de f.g.p de  $Z_n$  e por propriedades de esperança condicional, temos

$$f_n(t) = \mathbb{E}[t^{Z_n}] = \sum_{k=0}^{\infty} \mathbb{E}[t^{Z_n} \mid Z_{n-1} = k] \mathbb{P}[Z_{n-1} = k].$$

Daqui, dado que  $Z_n = \sum_{i=1}^{Z_{n-1}} X_{(n-1)i}$ , segue que

$$\mathbb{E}[t^{Z_n}] = \sum_{k=0}^{\infty} \mathbb{E}[t^{\sum_{i=1}^{Z_{n-1}} X_{(n-1)i}} \mid Z_{n-1} = k] \mathbb{P}[Z_{n-1} = k],$$

então pela independência entre as variáveis aleatórias  $X_{(n-1)i}$  e  $Z_{n-1}$ , temos

$$\begin{aligned} \mathbb{E}[t^{Z_n}] &= \sum_{k=0}^{\infty} \mathbb{E}[t^{\sum_{i=1}^k X_{(n-1)i}}] \mathbb{P}[Z_{n-1} = k] \\ &= \sum_{k=0}^{\infty} \mathbb{E}\left[\prod_{i=1}^k t^{X_{(n-1)i}}\right] \mathbb{P}[Z_{n-1} = k]. \end{aligned}$$

Agora, usando a independência das variáveis aleatórias  $X_{(n-1)i}$ , segue

$$\begin{aligned} f_n(t) = \mathbb{E}[t^{Z_n}] &= \sum_{k=0}^{\infty} \mathbb{E}[t^{X_{n-1}}]^k \mathbb{P}[Z_{n-1} = k] \\ &= \sum_{k=0}^{\infty} (\varphi_{n-1}(t))^k \mathbb{P}[Z_{n-1} = k] = f_{n-1}(\varphi_{n-1}(t)), \end{aligned}$$

ou seja (3.1) é válida.

Por outro lado,  $m_n := \mathbb{E}[Z_n]$  a reprodução média na geração  $n$ , é dada por

$$m_n = \prod_{j=0}^{n-1} \mu_j. \quad (3.2)$$

Isto é, por propriedades de esperança condicional, temos

$$\mathbb{E}[Z_n] = \sum_{k=0}^{\infty} \mathbb{E}[Z_n \mid Z_{n-1} = k] \mathbb{P}[Z_{n-1} = k].$$

Daqui, segue que

$$\begin{aligned} \mathbb{E}[Z_n] &= \sum_{k=0}^{\infty} \mathbb{E} \left[ \sum_{i=1}^{Z_{n-1}} X_{(n-1)i} \mid Z_{n-1} = k \right] \mathbb{P}[Z_{n-1} = k] \\ &= \sum_{k=0}^{\infty} \mathbb{E} \left[ \sum_{i=1}^k X_{(n-1)i} \right] \mathbb{P}[Z_{n-1} = k] \\ &= \sum_{k=0}^{\infty} k \mathbb{E}[X_{n-1}] \mathbb{P}[Z_{n-1} = k] \\ &= \mu_{n-1} \sum_{k=0}^{\infty} k \mathbb{P}[Z_{n-1} = k] = \mu_{n-1} \mathbb{E}[Z_{n-1}]. \end{aligned}$$

Portanto,

$$\begin{aligned} \mathbb{E}[Z_n] &= \mu_{n-1} \mathbb{E}[Z_{n-1}] \\ &= \mu_{n-1} \mu_{n-2} \mathbb{E}[Z_{n-2}] \\ &\vdots \\ &= \mu_{n-1} \mu_{n-2} \cdots \mu_1 \mathbb{E}[Z_1] \\ &= \mu_{n-1} \mu_{n-2} \cdots \mu_1 \mu_0 = \prod_{j=0}^{n-1} \mu_j. \end{aligned}$$

Isso mostra que (3.2) é verdadeira.

## 3.2 Resultados

Nesta secção serão apresentados e mostrados teoremas limites dos processos de Galton-Watson em meios variáveis, isto é, será mostrada a existência de uma variável aleatória

(possivelmente infinita), para a qual  $Z_n$  converge quase certamente. Além disso, será mostrado que a diferença com o caso clássico de processos de Galton-Watson pode acontecer só se o tamanho da população resulta relativamente constante.

**Teorema 3.2.1.** *Existe uma variável aleatória  $Z_\infty$  tomando valores em  $\mathbb{Z}_+ \cup \{\infty\}$ , tal que  $Z_n$  converge para  $Z_\infty$  quase certamente.*

**Teorema 3.2.2.** *Dado  $p_{n0} < 1$ , para qualquer  $n \geq 0$ , a variável aleatória  $Z_\infty$  satisfaz  $\mathbb{P}[0 < Z_\infty < \infty] > 0$  se e somente se*

- (i)  $\sum_{n=1}^{\infty} [1 - p_{n1}] < \infty$ ;
- (ii)  $\mathbb{E}[Z_\infty] \leq \liminf_{n \rightarrow \infty} m_n$ .

**Corolário 3.2.3.** *Se*

$$\sum_{n=1}^{\infty} [1 - p_{n1}] = \infty$$

e

$$\liminf_{n \rightarrow \infty} m_n < \infty.$$

Então,  $\mathbb{P}[Z_\infty = 0] = 1$ .

Os lemas seguintes são fundamentais para o desenvolvimento das provas dos Teoremas anteriormente apresentados.

**Lema 3.2.4.** *Seja  $f_n$  a f.g.p de  $Z_n$ . Se  $g := \lim_{n \rightarrow \infty} f_n$  existe e é estritamente crescente, então  $\varphi_n \rightarrow \varepsilon$  uniformemente, quando  $n \rightarrow \infty$ . Aqui  $\varepsilon(t) = t$ .*

*Demonstração.* Fixemos  $0 \leq t < 1$  e seja  $\delta > 0$  arbitrário, tal que  $t + \delta < 1$ . Agora, assumamos por contradição que  $\limsup_{n \rightarrow \infty} \varphi_n(t) > t$ . Daqui que existe  $\delta > 0$  e uma subsequência  $\{\varphi_{n_k} : k \geq 0\}$ , tal que  $\varphi_{n_k}(t) \geq t + \delta$ , assim

$$f_{n_{k+1}}(t) = \varphi_0 \circ \varphi_1 \circ \dots \circ \varphi_{n_k-1} \circ \varphi_{n_k} = f_{n_k} \circ \varphi_{n_k}(t) \geq f_{n_k}(t + \delta)$$

e fazendo  $k \rightarrow \infty$ , temos

$$g(t) \geq g(t + \delta),$$

que é uma contradição, pois  $g$  é estritamente crescente. Portanto

$$\limsup_{n \rightarrow \infty} \varphi_n(t) \leq t.$$

Similarmente suponha que  $\liminf_{n \rightarrow \infty} \varphi_n(t) < t$ . Daqui que existe  $\delta > 0$  e uma subsequência  $\{\varphi_{n_k} : k \geq 0\}$ , tal que  $\varphi_{n_k}(t) \leq t - \delta$ , assim

$$f_{n_k+1}(t) = \varphi_0 \circ \varphi_1 \circ \dots \circ \varphi_{n_k-1} \circ \varphi_{n_k} = f_{n_k} \circ \varphi_{n_k}(t) \leq f_{n_k}(t - \delta)$$

e fazendo  $k \rightarrow \infty$ , resulta

$$g(t) \leq g(t - \delta).$$

Isso é uma contradição. Logo

$$\liminf_{n \rightarrow \infty} \varphi_n(t) \geq t.$$

Por conseguinte

$$\limsup_{n \rightarrow \infty} \varphi_n(t) \leq \lim_{n \rightarrow \infty} \varphi_n(t) \leq \liminf_{n \rightarrow \infty} \varphi_n(t)$$

e assim, a convergência de funções geradoras de probabilidades, para uma função (como  $\varepsilon$ ), é uniforme.  $\square$

**Lema 3.2.5.** *Seja  $\{g_n\}$  e  $\{h_n\}$  duas seqüências de funções geradoras. Suponha que  $g_n \circ h_n \rightarrow \varepsilon$ , quando  $n \rightarrow \infty$ . Então ambas,  $g_n \rightarrow \varepsilon$  e  $h_n \rightarrow \varepsilon$ , quando  $n \rightarrow \infty$ .*

*Demonstração.* Sejam  $X_{nk}$ ,  $Y_n$ , para  $n, k$  duas variáveis aleatórias independentes com funções geradoras  $h_n$  e  $g_n$ , respectivamente. Então

$$\sum_{k=1}^{Y_n} X_{nk}$$

tem função geradora  $g_n \circ h_n$ , isto é,

$$\begin{aligned} \mathbb{E} \left[ t^{\sum_{k=1}^{Y_n} X_{nk}} \right] &= \sum_{l=0}^{\infty} \mathbb{E} \left( t^{\sum_{k=1}^{Y_n} X_{nk}} \mid Y_n = l \right) \mathbb{P}(Y_n = l) \\ &= \sum_{l=0}^{\infty} \mathbb{E} \left( t^{\sum_{k=1}^l X_{nk}} \right) \mathbb{P}(Y_n = l) \\ &= \sum_{l=0}^{\infty} \mathbb{E} \left( \prod_{k=1}^l t^{X_{nk}} \right) \mathbb{P}(Y_n = l) = \sum_{l=0}^{\infty} \mathbb{E} \left( t^{X_{nk}} \right)^l \mathbb{P}(Y_n = l) = g_n(h_n(t)). \end{aligned}$$

Então a convergência de  $g_n \circ h_n$  para  $\varepsilon$  implica que  $\sum_{k=1}^{Y_n} X_{nk}$  converge em distribuição para

1. Dado que  $Y_n$  e  $X_{nk}$  são variáveis aleatórias não negativas de valor inteiro, isso requer que ambas,  $Y_n$  e  $X_{nk}$  convirjam em distribuição para 1, quando  $n \rightarrow \infty$ .

Isso completa a prova do lema.  $\square$

**Lema 3.2.6.** *O Teorema 3.2.1 é verdadeiro para convergência em distribuição.*

*Demonstração.* Primeiro temos que  $f_n(0) = \varphi_0 \circ \dots \circ \varphi_{n-1}(0)$  é não decrescente em  $n$

$$q = \mathbb{P}[Z_n \rightarrow 0] = \lim_{n \rightarrow \infty} f_n(0).$$

Se para todo  $0 < t < 1$ , temos que

$$\limsup_{n \rightarrow \infty} f_n(t) = q,$$

segue que  $f_n$  converge para  $q$  sobre  $[0, 1)$ . Agora, assumimos que existe  $0 < t < 1$  tal que

$$\limsup_{n \rightarrow \infty} f_n(t) = a > q.$$

Então  $f_{n_k}(t)$  converge para  $a$ , para alguma subsequência  $\{n_k\}$ . Esta subsequência pode ser escolhida tal que  $f_{n_k}$  converge para alguma função  $g$  uniformemente sobre subconjuntos compactos de  $[0, 1)$ . Dado que as  $f_{n_k}$  são funções geradoras, segue que  $g$  também é função geradora e como  $g(s) = a > q = g(0)$ , resulta que  $g$  é estritamente crescente. Daqui,

$$f_{n_k} = (\varphi_0 \circ \dots \circ \varphi_{n_1-1}) \circ \dots \circ (\varphi_{n_{k-1}} \circ \dots \circ \varphi_{n_k-1}),$$

e pelo Lema 3.2.4, temos que

$$\lim_{k \rightarrow \infty} (\varphi_{n_{k-1}} \circ \dots \circ \varphi_{n_k-1}) = \varepsilon.$$

Agora, para qualquer  $n \in \mathbb{N}$ , definamos  $k(n)$  por  $n_{k(n)} \leq n < n_{k(n)+1}$ . Deste modo pode-se ver que

$$f_n = f_{n_{k(n)}} \circ (\varphi_{n_{k(n)}} \circ \dots \circ \varphi_{n-1}),$$

mas  $(\varphi_{n_{k(n)}} \circ \dots \circ \varphi_{n-1}) \circ (\varphi_n \circ \dots \circ \varphi_{n_{k(n)+1}} \circ \varphi_{n-1})$  converge para  $\varepsilon$ . Então pelo Lema 3.2.5, temos que  $\varphi_{n_{k(n)}} \circ \dots \circ \varphi_{n-1}$  converge para  $\varepsilon$  e portanto, dado que  $f_{n_{k(n)}}$  converge para  $g$  uniformemente sobre subconjuntos compactos de  $[0, 1)$ , segue que  $f_n$  converge para  $g$ . Isso completa a prova do lema.  $\square$

Agora, podemos passar a provar os Teoremas 3.2.1 e 3.2.2 os quais são os resultados principais desta seção.

*Demonstração.* Teorema 3.2.1 A prova desse teorema será feita considerando dois casos:

**Caso 1:** Suponha  $\sum_{n=0}^{\infty} (1 - p_{n1}) < \infty$ .

Seja,

$$\begin{aligned}
|\varphi_n(s) - s| &= \left| \sum_{k \neq 1} p_{nk} s^k + p_{n1} s - s \right| = \left| \sum_{k \neq 1} p_{nk} s^k - (1 - p_{n1}) s \right| \\
&\leq \left| \sum_{k \neq 1} p_{nk} s^k \right| + |(1 - p_{n1}) s|, \quad \text{por desigualdade triangular} \\
&\leq \sum_{k \neq 1} |p_{nk}| |s^k| + |(1 - p_{n1})| |s|, \\
&\leq \sum_{k \neq 1} p_{nk} + 1 - p_{n1}, \quad \text{para } |s| \leq 1 \\
&= 2(1 - p_{n1}). \tag{3.3}
\end{aligned}$$

Daqui que,  $2(1 - p_{n1})$  é um limitante superior para  $|\varphi_n(s) - s|$  e portanto,  $\sup_{0 < s < 1} |\varphi_n(s) - s| \leq 2(1 - p_{n1})$ , assim, somando para um  $j$  suficientemente grande e usando a hipótese, segue

$$\sum_{n > j} \sup_{0 < s < 1} |\varphi_n(s) - s| \leq 2 \sum_{n \geq j} (1 - p_{n1}) < 1/4.$$

Dado que  $1 - p_{n1} < 1/8$  para  $n > j$ , as funções  $\varphi_n$  são crescentes em  $n$  e têm inversas. Então, definamos  $a_j := 1/2$  e para  $n > j$ , definamos  $a_{n+1} := \varphi^{-1}(a_n)$ . Além disso, por (3.3), temos que

$$|a_{n+1} - a_n| = |a_{n+1} - \varphi_n(a_{n+1})| \leq 2(1 - p_{n1}),$$

daqui, somando para algum  $i \geq n$ , segue

$$|a_{n+k} - a_n| \leq 2 \sum_{i \geq n} (1 - p_{i1}) < 1/4. \tag{3.4}$$

Nessas condições, resulta que  $\lim_{n \rightarrow \infty} a_n =: a$  existe. Isto é, tomando  $n = j$  e fazendo  $k \rightarrow \infty$  em (3.4), pode-se ver que  $|a - a_j| \leq 1/4$  ou  $1/4 \leq a \leq 3/4$ .

Por outro lado, definamos no intervalo  $[0, 1]$ , a variável aleatória  $Y_n := a_n^{Z_n}$ , a qual é adaptada para a filtragem  $\mathcal{F}_n = \sigma(Z_1, Z_2, \dots, Z_n)$  e

$$\mathbb{E}[Y_{n+1} | \mathcal{F}_n] = \mathbb{E}[a_{n+1}^{Z_{n+1}} | Z_n] = \{\varphi_n(a_{n+1})\}^{Z_n} = a_n^{Z_n}, \tag{3.5}$$



onde a terceira igualdade em (3.5) resulta do fato que  $Z_{n+1} = \sum_{k=1}^{Z_n} X_{nk}$  e pela independência das variáveis  $Z_n$  e  $X_{nk}$ , isto é,

$$\begin{aligned} \mathbb{E} \left[ a_{n+1}^{Z_{n+1}} \mid Z_n \right] &= \mathbb{E} \left[ a_{n+1}^{\sum_{k=1}^{Z_n} X_{nk}} \mid Z_n \right] = \mathbb{E} \left[ a_{n+1}^{\sum_{k=1}^{Z_n} X_{nk}} \right] \\ &= \prod_{k=1}^{Z_n} \mathbb{E} \left[ a_{n+1}^{X_{nk}} \right] \\ &= \{\varphi_n(a_{n+1})\}^{Z_n}. \end{aligned} \quad (3.6)$$

Portanto, de (3.5) resulta que  $Y_n$  é um martingale não negativo com respeito a filtragem  $(\mathcal{F}_n)_{n \geq 0}$  e pelo Teorema (2.2.1), segue que  $\lim_{n \rightarrow \infty} Y_n =: Y$  existe quase certamente. Logo tomando logaritmo, temos que

$$\lim_{n \rightarrow \infty} Z_n \log a_n = \log Y, \quad \text{q.c}$$

e dado que  $\lim_{n \rightarrow \infty} a_n = a$ , segue que,  $Z_n$  converge para  $\frac{\log Y}{\log a} =: Z_\infty$ , quase certamente.

**Caso 2:** Suponha  $\sum_{n=0}^{\infty} (1 - p_{n1}) = \infty$ . Se tivermos que  $q := \lim_{n \rightarrow \infty} f_n(0) = \mathbb{P}[Z_n \rightarrow 0] = 1$ , resulta que  $Z_n$  converge para zero quase certamente. Então, assumiremos que  $q < 1$ . Agora, lembremos o Teorema (3.2.6) o qual diz que existe uma variável aleatória  $Z_\infty$  e que  $Z_n$  converge para  $Z_\infty$  em distribuição, dessa forma,  $f_n$  converge para a função  $g$ , que é função geradora da variável aleatória  $Z_\infty$ . Por outro lado, se tivermos que para todo  $s \in (0, 1)$ ,  $\limsup_{n \rightarrow \infty} f_n(s) = q$ , então,  $\lim_{n \rightarrow \infty} f_n(s) = q$ , para  $s \in [0, 1)$ , pois  $\lim_{n \rightarrow \infty} f_n(0) = q$ . Então por unicidade de limite temos que  $g(s) = q$ , para  $s \in [0, 1)$ . Assim, podemos definir  $t_n$ , para um  $n$  suficientemente grande, como segue

$$\sup_{k \geq n} \sup_{0 \leq s < 1/2} |f_k(s) - q| =: t_n \downarrow 0. \quad (3.7)$$

Daqui, tomemos um  $j$  suficientemente grande tal que  $q + 2t_n \leq 1$ , para  $n \geq j$ . Para esse  $n$ , definamos  $a_n := f_n^{-1}(q + 2t_n)$ . Consequentemente  $f_n(a_n) - q = 2t_n$ , daqui que  $a_n > 1/2$ , para  $t_n > 0$ . Pois  $t_n = 0$ , quando  $f_k$  é constante, isto é, se algum  $\varphi_k = 1$  e nesse caso teríamos  $Z_n = 0$  para  $n \geq k$ . Então, trabalharemos restringindo para  $t_n > 0$ , para todo  $n$ . Além disso, temos que

$$\varphi_n(a_{n+1}) = \varphi_n \circ f_{n+1}^{-1}(q + 2t_{n+1}) = f_n^{-1}(q + 2t_{n+1}) \leq f_n^{-1}(q + 2t_n) = a_n, \quad (3.8)$$

onde a segunda igualdade em (3.8) é porque  $f_{n+1} = f_n \circ \varphi_n$  e a desigualdade pelo fato de  $t_n$  ser decrescente. Agora definamos a variável aleatória  $W_n := a_n^{Z_n} - 1_{\{Z_n=0\}}$ , para  $n \geq k$ , a qual é adaptada para a filtragem  $\mathcal{F}_n = \sigma(Z_1, \dots, Z_n)$  e

$$\begin{aligned} \mathbb{E}[W_{n+1} \mid \mathcal{F}_n] &= \mathbb{E}\left[a_{n+1}^{Z_{n+1}} - 1_{\{Z_{n+1}=0\}} \mid Z_n\right] \\ &= \{\varphi_n(a_{n+1})\}^{Z_n} - \{\varphi_n(0)\}^{Z_n} \\ &\leq a_n^{Z_n} - \{\varphi_n(0)\}^{Z_n} \\ &\leq a_n^{Z_n} - 1_{\{Z_n=0\}} = W_n. \end{aligned} \tag{3.9}$$

Onde o primeiro termo na segunda igualdade de (3.9), foi provado no caso 1 na equação (3.6), então resta provar a validade do segundo termo. Com efeito,

$$\begin{aligned} \mathbb{E}\left[1_{\{Z_{n+1}=0\}} \mid Z_n\right] &= \mathbb{P}[Z_{n+1} = 0 \mid Z_n] = \mathbb{P}\left[\sum_{k=1}^{Z_n} X_{nk} = 0 \mid Z_n\right] \\ &= \mathbb{P}\left[\sum_{k=1}^{Z_n} X_{nk} = 0\right] \\ &= \mathbb{P}\left[\bigcap_{k=1}^{Z_n} \{X_{nk} = 0\}\right] \\ &= \{\mathbb{P}[X_n = 0]\}^{Z_n} \\ &= \{\varphi_n(0)\}^{Z_n}. \end{aligned}$$

Assim, de (3.9) segue que  $W_n$  é um supermartingal não negativo com respeito a filtragem  $(\mathcal{F}_{n \geq 0})$  e pelo Teorema 2.2.1,  $X_n$  converge para alguma variável  $W$  quase certamente, porém,

$$\mathbb{E}[W_n] = f_n(a_n) - f_n(0) = q + 2t_n - f_n(0) \leq 3t_n \downarrow 0. \tag{3.10}$$

Daqui, dado que  $\lim_{n \rightarrow \infty} W_n =: W$  existe quase certamente, então por (3.10) e pelo Lema de Fatou (ver Resnick (2013)), segue que

$$\mathbb{E}[W] = \mathbb{E}[\liminf_{n \rightarrow \infty} W_n] \leq \liminf_{n \rightarrow \infty} \mathbb{E}[W_n] = \lim_{n \rightarrow \infty} \mathbb{E}[W_n] = 0.$$

Isto é,  $W = 0$  quase certamente e portanto  $W_n$  converge para 0 quase certamente.

Agora, se tivermos que  $Z_n$  converge para 0, então

$$\lim_{n \rightarrow \infty} W_n = \lim_{n \rightarrow \infty} (a_n^{Z_n} - 1_{\{0\}}(Z_n)) = 1 - 1 = 0,$$

mas, se tivermos que  $Z_n$  não converge para 0, então  $a_n^{Z_n}$  converge para 0 quase certamente, e dado que  $a_n \geq 1/2$ , daí resulta que  $Z_n \rightarrow \infty$ .  $\square$

*Demonstração.* Teorema 3.2.2

Para provar o item (i), suponhamos primeiro que  $\mathbb{P}[0 < Z_\infty < \infty] > 0$ . Analiticamente, comparamos

$$\prod_{k=n}^{\infty} \varphi'_k(0)$$

para um  $n$  grande, com o produto

$$\prod_{k=n}^{\infty} \varphi'_k \circ g_{k+1}(0).$$

Aqui,

$$g_k = \lim_{n \rightarrow \infty} \varphi_k \circ \dots \circ \varphi_n,$$

existe pelo Lema 3.2.6. Dado que  $\mathbb{P}[0 < Z_\infty < \infty]$  é estritamente positivo, resulta que  $g$  é estritamente crescente e escrevemos  $g := f_k \circ g_k$ , assim,  $g_k \rightarrow \varepsilon$ . Portanto

$$\prod_{k=n}^{\infty} \varphi'_k \circ g_{k+1}(0) = g'_n(0) \rightarrow 1,$$

quando  $n \rightarrow \infty$ . Além disso,

$$\begin{aligned} \frac{\prod_{k=n}^{\infty} \varphi'_k(0)}{\prod_{k=n}^{\infty} \varphi'_k \circ g_{k+1}(0)} &= \exp \left( \log \left( \prod_{k=n}^{\infty} \frac{\varphi'_k(0)}{\varphi'_k \circ g_{k+1}(0)} \right) \right) \\ &= \exp \left( \sum_{k=n}^{\infty} \log \left( \frac{\varphi'_k(0)}{\varphi'_k \circ g_{k+1}(0)} \right) \right) \\ &= \exp \left( - \sum_{k=n}^{\infty} \log \left( \frac{\varphi'_k \circ g_{k+1}(0)}{\varphi'_k(0)} \right) \right) \\ &\geq \exp \left( - \sum_{k=n}^{\infty} \left( \frac{\varphi'_k \circ g_{k+1}(0)}{\varphi'_k(0)} - 1 \right) \right) \\ &\geq \exp \left( - \sum_{k=n}^{\infty} \frac{\varphi''_k \circ g_{k+1}(0) g_{k+1}(0)}{\varphi'_k(0)} \right). \end{aligned}$$

Onde a primeira desigualdade sai do resultado,

$$\log x \leq x; \quad x > 0,$$

com igualdade em  $x = 1$ . A segunda desigualdade é pelo seguinte fato

$$\begin{aligned} \varphi_k(t) &= \sum_{i=0}^{\infty} t^i \mathbb{P}[X_k = i]; \\ \varphi'_k(t) &= \sum_{i=1}^{\infty} i t^{i-1} \mathbb{P}[X_k = i]; \\ \varphi'_k(g_{k+1}(0)) &= \sum_{i=1}^{\infty} i (g_{k+1}(0))^{i-1} \mathbb{P}[X_k = i]; \\ \varphi''_k(t) &= \sum_{i=2}^{\infty} i(i-1) t^{i-2} \mathbb{P}[X_k = i]; \\ \varphi''_k(g_{k+1}(0)) &= \sum_{i=2}^{\infty} i(i-1) (g_{k+1}(0))^{i-2} \mathbb{P}[X_k = i]. \end{aligned}$$

Daqui, temos

$$\begin{aligned} \varphi'_k \circ g_{k+1}(0) - \varphi'_k(0) &= \sum_{i=1}^{\infty} i (g_{k+1}(0))^{i-1} \mathbb{P}[X_k = i] - \mathbb{P}[X_k = 1] \\ &= \sum_{i=2}^{\infty} i (g_{k+1}(0))^{i-1} \mathbb{P}[X_k = i]. \end{aligned} \quad (3.11)$$

Por outro lado

$$\varphi''_k \circ g_{k+1}(0) g_{k+1}(0) = \sum_{i=2}^{\infty} i(i-1) (g_{k+1}(0))^{i-1} \mathbb{P}[X_k = i] \quad (3.12)$$

Comparando (3.11) e (3.12), temos que

$$\varphi'_k \circ g_{k+1}(0) - \varphi'_k(0) \leq \varphi''_k \circ g_{k+1}(0) g_{k+1}(0).$$

Isso implica a segunda desigualdade. Agora, dado que  $\varphi'_k(0) \rightarrow 1$  podemos elegir  $n$  tal que  $\varphi'_k(0) \geq 1/2$ , para  $k \geq n$ . Então para tal  $n$ , segue que

$$\prod_{k=n}^{\infty} \varphi'_k(0) \geq \prod_{k=n}^{\infty} \varphi'_k \circ g_{k+1}(0) \exp \left( -2 \sum_{k=n}^{\infty} \varphi''_k \circ g_{k+1}(0) \right).$$

No entanto desde que  $\varphi'_k \circ g_{k+1}(0) \rightarrow 1$ ,  $g'_n(0) \rightarrow 1$  e  $\varphi''_k \geq 0$ , para um  $n$  grande, temos que

$$0 \leq \sum_{k=n}^{\infty} \varphi''_k \circ g_{k+1}(0) \leq 2g'_n(0) \sum_{k=n}^{\infty} \frac{\varphi''_k \circ g_{k+1}(0)}{\varphi'_k \circ g_{k+1}(0)} = 2g''_n(0) \rightarrow 0,$$

quando  $n \rightarrow \infty$ . Portanto,

$$\prod_{k=n}^{\infty} \varphi'_k(0) > 0,$$

para um  $n$  grande e dado que  $\varphi'_n(0) = p_{n1} \leq 1$ ,

$$\sum_{n=0}^{\infty} \{1 - p_{n1}\} < \infty.$$

Reciprocamente, se  $\sum_{n=0}^{\infty} \{1 - p_{n1}\} < \infty$ , então existem inteiros positivos  $i, j$ , tais que  $\mathbb{P}[Z_i = j] > 0$  e  $p_{n1} > 0$ , para todo  $n > i$ .

Com efeito, suponhamos que para todo  $i, j$ ,  $\mathbb{P}[Z_i = j] = 0$  ou  $p_{n1} = 0$ . Daqui, se  $p_{n1} = 0$ , temos que  $\sum_{n=0}^{\infty} \{1 - p_{n1}\} = \infty$ , o que é uma contradição. Agora, se para todo  $i, j$ ,  $\mathbb{P}[Z_i = j] = 0$  então temos que  $\mathbb{P}[Z_i = \infty]$ , isso implica que  $\sum_{n=0}^{\infty} \{1 - p_{n1}\} = \infty$ , o que é uma contradição. Além disso, temos que

$$\begin{aligned} \mathbb{P}[Z_{\infty} = j] &\geq \mathbb{P}[Z_i = j, Z_{i+k} = j; k \geq 1] \\ &= \mathbb{P}[Z_i = j, Z_{i+1} = j, Z_{i+2} = j, Z_{i+3} = j \cdots] \\ &= \mathbb{P}[Z_{i+1} = j, Z_{i+2} = j, Z_{i+3} = j, \cdots \mid Z_i = j] \mathbb{P}[Z_i = j] \\ &= \mathbb{P}[Z_{i+2} = j, Z_{i+3} = j, Z_{i+4} = j, \cdots \mid Z_{i+1} = j] \mathbb{P}[Z_{i+1} = j \mid Z_i = j] \mathbb{P}[Z_i = j], \end{aligned}$$

onde  $\mathbb{P}[Z_{i+1} = j \mid Z_i = j] = p_{i1}^j$ . Logo,

$$\begin{aligned} \mathbb{P}[Z_{\infty} = j] &\geq \mathbb{P}[Z_{i+2} = j, Z_{i+3} = j, Z_{i+4} = j, \cdots \mid Z_{i+1} = j] p_{i1}^j \mathbb{P}[Z_i = j] \\ &= \mathbb{P}[Z_{i+3} = j, Z_{i+4} = j, \cdots \mid Z_{i+2} = j] \mathbb{P}[Z_{i+2} = j \mid Z_{i+1} = j] p_{i1}^j \mathbb{P}[Z_i = j] \\ &= \mathbb{P}[Z_{i+3} = j, Z_{i+4} = j, \cdots \mid Z_{i+2} = j] p_{(i+1)1}^j p_{i1}^j \mathbb{P}[Z_i = j] \\ &\vdots \\ &= \mathbb{P}[Z_i = j] \prod_{n=i}^{\infty} p_{n1}^j > 0. \end{aligned}$$

Isso implica que  $\mathbb{P}[0 < Z_{\infty} < \infty] > 0$ . O que completa a prova do item (i).

A prova do item (ii) segue diretamente do Lema de Fatou (ver Resnick (2013)). Isto é, dado que  $Z_n$  converge para a variável aleatória  $Z_{\infty}$ , temos que

$$Z_{\infty} = \liminf_{n \rightarrow \infty} Z_n,$$

dessa forma, pelo Lema de Fatou, segue

$$\mathbb{E}[Z_\infty] = \mathbb{E}\left[\liminf_{n \rightarrow \infty} Z_n\right] \leq \liminf_{n \rightarrow \infty} \mathbb{E}[Z_n] = \liminf_{n \rightarrow \infty} m_n.$$

□



---

## Aplicações

---

Neste capítulo apresentam-se dois modelos de processos de ramificação binária com aplicações biológicas. No modelo 1, Schinazi (2006) avalia o risco da emergência da resistência durante um tratamento. No modelo 2, apresenta-se uma extensão do modelo 1, isto é, avalia-se também a probabilidade de emergência da resistência, só que desta vez no tratamento são ministradas duas drogas.

### 4.1 Probabilidade de resistência aos medicamentos

A resistência aos medicamentos é uma ameaça constante para a saúde de indivíduos que estejam sendo tratados por diversas doenças como por exemplo: tuberculose, câncer, AIDS, etc., uma vez que existe um risco que uma doença tratável seja substituída por uma não tratável. Resistência é a capacidade que certos patógenos têm de continuar-se multiplicando apesar da presença de medicamentos que normalmente os aniquilam. Patógenos que têm DNA resistente a medicamentos podem transferir uma cópia destes genes com outros patógenos e assim patógenos não resistentes ou sensíveis aos medicamentos recebem o novo DNA e tornam-se resistentes (mutação). Na presença de medicamentos, somente os patógenos resistentes sobrevivem, depois estes multiplicam-se e prosperam. No que segue, descreve-se um modelo proposto por Schinazi (2006) para calcular a probabilidade de erradicação de patógenos antes que a resistência apareça.



### 4.1.1 Descrição do modelo

Considera-se um processo de Galton-Watson de divisão binária a tempo discreto. Suponha que, neste caso, as partículas do processo representam patógenos sensíveis, os quais, dão origem a novos patógenos de acordo a uma variável aleatória  $X$ , com distribuição de probabilidades dada por

$$X = \begin{cases} 2, & \text{com probabilidade } p, \\ 0, & \text{com probabilidade } 1 - p. \end{cases} \quad (4.1)$$

Denota-se por  $(Z_n)_{n \geq 0}$  o processo de ramificação associado a  $X$ , onde  $Z_n$  representa o número de patógenos na  $n$ -ésima geração. Assuma-se que os descendentes diretos de uma partícula são i.i.d à variável aleatória  $X$ . Agora, considera-se  $\phi$  como a função geradora de probabilidades da variável aleatória  $X$ , (ver Exemplo 2.1.1 no Capítulo 2), isto é, para todo  $t$ ,

$$\phi(t) = pt^2 - p + 1, \quad (4.2)$$

daqui que o valor médio de descendentes por patógeno é  $2p$  e segundo o Teorema 2.1.4 do Capítulo 2, podem-se considerar dois casos:

- (i) Se  $2p > 1$ , então há uma probabilidade positiva para a árvore genealógica do patógeno inicial de sobreviver, isto é,  $\mathbb{P}[Z = \infty] > 0$ .
- (ii) Se  $2p \leq 1$ , então a erradicação é quase certa, isto é,  $\mathbb{P}[Z < \infty] = 1$ .

Nesse sentido, o parâmetro  $p$  é uma medida da eficácia do tratamento. Em outras palavras, quanto menor for o  $p$ , mais eficiente será o medicamento e mais provável a erradicação dos patógenos sensíveis. Seja  $N \in \mathbb{N}$  o número de patógenos sensíveis, no início do tratamento. Começamos com o caso  $N = 1$ , isto é, supomos  $Z_0 = 1$ . Assim, seja  $Z$  a variável aleatória que conta o número total de patógenos na árvore genealógica do patógeno inicial, ou seja,

$$Z = \sum_{n=0}^{\infty} Z_n.$$

Dado que o interesse é calcular a probabilidade de erradicação de patógenos sensíveis antes que a resistência apareça, dita probabilidade será denotada por  $q$ , iremos calcular primeiro a distribuição de  $Z$  e usaremos este resultado para calcular  $q$ .

Uma vez que o número de descendentes é sempre par, por se tratar de um processo de divisão binária e dado que o tratamento foi iniciado com um único patógeno, resulta

que  $Z$  é sempre ímpar. Então, seja  $F$  a função geradora de probabilidades da variável aleatória  $Z$ , dada por

$$F(t) = \sum_{n=1}^{\infty} \mathbb{P}[Z = 2n - 1 \mid Z_0 = 1] t^{2n-1}, \quad \text{para } |t| < 1. \quad (4.3)$$

Note que  $F$  é solução da equação

$$F(t) = t\phi(F(t)). \quad (4.4)$$

Com efeito,

$$\begin{aligned} \mathbb{P}[Z = 2n - 1 \mid Z_0 = 1] &= \sum_{j=0,2} \mathbb{P}[Z = 2n - 1, Z_1 = j \mid Z_0 = 1] \\ &= \sum_{j=0,2} \mathbb{P}[Z = 2n - 1 \mid Z_1 = j, Z_0 = 1] \mathbb{P}[Z_1 = j \mid Z_0 = 1] \\ &= \sum_{j=0,2} \mathbb{P} \left[ \sum_{i=1}^j Z^i = 2n - 2 \right] \mathbb{P}[Z_1 = j], \end{aligned} \quad (4.5)$$

e (4.3) pode-se também escrever como,

$$F(t) = t \sum_{n=1}^{\infty} \mathbb{P}[Z = 2n - 1 \mid Z_0 = 1] t^{2n-2}, \quad (4.6)$$

então, inserindo (4.5) em (4.6), resulta que

$$\begin{aligned} F(t) &= t \sum_{n=1}^{\infty} \sum_{j=0,2} \mathbb{P} \left[ \sum_{i=1}^j Z^i = 2n - 2 \right] \mathbb{P}[Z_1 = j] t^{2n-2} \\ &= t \sum_{j=0,2} \sum_{n=1}^{\infty} \mathbb{P} \left[ \sum_{i=1}^j Z^i = 2n - 2 \right] t^{2n-2} \mathbb{P}[Z_1 = j] \\ &= t \sum_{j=0,2} \mathbb{P}[Z_1 = j] (F(t))^j \\ &= t\phi(F(t)). \end{aligned}$$

Ou seja, (4.3) é verdadeira.

Logo, desde (4.2) e (4.3), temos que

$$F(t) = pt(F(t))^2 - pt + t, \quad \text{para todo } |t| < 1. \quad (4.7)$$

Daqui, resolvendo a equação quadrática (4.7) em relação a  $F(t)$ , obtemos as soluções

$$F(t) = \frac{1 \pm \sqrt{1 - 4pt(t - pt)}}{2pt}.$$

Particularmente para  $n \geq 1$ , toma-se a solução

$$F(t) = \frac{1}{2pt}(1 - \sqrt{1 - 4p(1-p)t^2}), \quad (4.8)$$

pois,  $F$  deve ser finito para  $t$  perto de zero. Agora usando expansão binomial, temos que

$$1 - \sqrt{1 - x} = \sum_{n=1}^{\infty} c_n x^n, \quad \text{para } |x| < 1,$$

onde

$$c_n = \frac{(2n-2)!}{2^{2n-1}n!(n-1)!}, \quad \text{para } n \geq 1.$$

Por conseguinte, usando esse fato em (4.8), tem-se que

$$F(t) = \frac{1}{2pt} \sum_{n=1}^{\infty} c_n (4p(1-p)t^2)^n = \sum_{n=1}^{\infty} \frac{(2n-2)!}{n!(n-1)!} (1-p)^n p^{n-1} t^{2n-1}. \quad (4.9)$$

Assim, comparando a última igualdade de (4.9) com (4.3), resulta que

$$\mathbb{P}[Z = 2n - 1 \mid Z_0 = 1] = \frac{(2n-2)!}{n!(n-1)!} (1-p)^n p^{n-1}, \quad \text{para } n \geq 1. \quad (4.10)$$

Uma vez obtida a distribuição de  $Z$  em (4.10), antes de calcular  $q$ , precisaremos definir mais algumas notações. Isto é, assumamos que em cada divisão de um patógeno sensível, existe uma probabilidade  $\mu$ , de algum descendente obter independentemente dos outros, mutação de resistência. Assim, se  $\mathcal{Z}$  denota o conjunto de patógenos na árvore genealógica do patógeno inicial (note que  $Z = |\mathcal{Z}|$ ) e se  $M_v$ , para todo  $v \in \mathcal{Z}$ , é uma variável aleatória indicadora do evento “O patógeno  $v$  tem uma mutação”, então,  $M_v \sim \text{Bernoulli}(\mu)$ . Daqui, pode-se ver que

$$\mathbb{P}^1 \left[ \bigcap_{v \in \mathcal{Z}} \{M_v = 0\} \mid Z = k \right] = (1 - \mu)^{k-1}, \quad (4.11)$$

onde o índice 1 na probabilidade indica que  $N = 1$ , isto é,  $\mathbb{P}^1[\cdot] := \mathbb{P}[\cdot \mid Z_0 = 1]$ .

A partir daqui, podemos calcular  $q(1, \mu, p)$ , a probabilidade de erradicação dos patógenos sensíveis antes que a mutação para a resistência apareça, ou em outras palavras, a

probabilidade de não ter resistência, em que  $1$ ,  $\mu$  e  $p$  indicam a existência de um único patógeno inicial no tempo  $n = 0$ , a probabilidade de mutação de cada descendente e a probabilidade divisão, respectivamente. Logo, usando (4.10) e (4.12) temos que

$$\begin{aligned} q(1, \mu, p) &= \mathbb{P}^1 \left[ \bigcap_{v \in \mathcal{Z}} \{M_v = 0\} \right] = \sum_{n=1}^{\infty} (1 - \mu)^{(2n-1)-1} \mathbb{P}[Z = 2n - 1 \mid Z_0 = 1] \\ &= \sum_{n=1}^{\infty} (1 - \mu)^{2n-2} \frac{(2n - 2)!}{n!(n - 1)!} (1 - p)^n p^{n-1}. \end{aligned}$$

Fazendo alguns cálculos algébricos resulta

$$\begin{aligned} q(1, \mu, p) &= \sum_{n=1}^{\infty} \frac{(2n - 2)!}{n!(n - 1)!} \frac{(1 - p)^n p^n (1 - \mu)^{2n}}{p(1 - \mu)^2} \frac{2^{2n-1}}{2^{2n-1}} \\ &= \sum_{n=1}^{\infty} \frac{(2n - 2)!}{n!(n - 1)!} \frac{(1 - p)^n p^n (1 - \mu)^{2n} 2^{2n}}{2p(1 - \mu)^2} \\ &= \frac{1}{2p(1 - \mu)^2} \sum_{n=1}^{\infty} \frac{(2n - 2)!}{2^{2n-1} n!(n - 1)!} (4(1 - \mu)^2 p(1 - p))^n \\ &= \frac{1}{2p(1 - \mu)^2} \left( 1 - \sqrt{1 - 4p(1 - p)(1 - \mu)^2} \right). \end{aligned}$$

Daqui, usando a independência das variáveis aleatórias envolvidas, segue que

$$q(N, \mu, p) = (q(1, \mu, p))^N = \left( \frac{1}{2p(1 - \mu)^2} \left( 1 - \sqrt{1 - 4p(1 - p)(1 - \mu)^2} \right) \right)^N. \quad (4.12)$$

Uma forma de obter uma expressão mais simples para (4.12), pode ser usando aproximação linear.

**Definição 4.1.1.** *Dada uma função  $f(x)$  contínua, diferenciável e com uma variável real  $x$ ,*

$$f(x) = f(a) + f'(a)(x - a) + o(x).$$

Onde  $o(x)$  é uma função que representa o erro ( $\lim_{x \rightarrow a} o(x)/x = 0$ ). Uma aproximação linear de  $f(x)$  é obtida desconsiderando a função erro, isto é,

$$f(x) \approx f(a) + f'(a)(x - a),$$

para valores próximos de  $a$ , a curva descrita pela função  $f(x)$  se aproxima de uma reta.

No que segue, aproximaremos (4.12), por partes. Isto é, tomemos primeiro a função  $f_1(\mu) = (1 - \mu)^{-2}$  em (4.12), daqui que  $f'_1(\mu) = 2(1 - \mu)^{-3}$ ,  $f_1(0) = 1$  e  $f'_1(0) = 2$ . Logo,

pela Definição 4.1.1, resulta a seguinte aproximação linear para  $f_1$ ,

$$f_1(\mu) \approx 1 + 2\mu.$$

Agora tomemos  $f_2(\mu) = \sqrt{1 - 4p(1-p)(1-\mu)^2}$  em (4.12) e notemos que

$$f_2'(\mu) = \frac{8p(1-p)(1-\mu)}{2\sqrt{1-4p(1-p)(1-\mu)^2}},$$

$$f_2(0) = \sqrt{1-4p(1-p)} \quad \text{e} \quad f_2'(0) = \frac{8p(1-p)}{2\sqrt{1-4p(1-p)}},$$

assim, desde a Definição 4.1.1, uma aproximação linear para  $f_2$  é

$$f_2(\mu) \approx |1-2p| \left( 1 + \frac{4p(1-p)}{(1-2p)^2} \mu \right). \quad (4.13)$$

Daqui, usando as aproximações de  $f_1$  e  $f_2$ , obtemos para  $p < 1/2$  a seguinte aproximação para (4.12),

$$\begin{aligned} q(N, \mu, p) &\approx \left( \frac{(1+2\mu)}{2p} \left( 1 - (1-2p) \left( 1 + \frac{4p(1-p)}{(1-2p)^2} \mu \right) \right) \right)^N \\ &= \left( \frac{(1+2\mu)}{2p} \left( 2p - \frac{4p(1-p)}{(1-2p)} \mu \right) \right)^N \\ &= \left( (1+2\mu) \left( 1 - \frac{2(1-p)}{(1-2p)} \mu \right) \right)^N. \end{aligned} \quad (4.14)$$

Agora, tomando a função  $f_3(\mu) = (1+2\mu) \left( 1 - \frac{2(1-p)}{(1-2p)} \mu \right)$  em (4.14), tal que

$$f_3'(\mu) = (1+2\mu) \left( -\frac{2(1-p)}{(1-2p)} \right) + 2 \left( 1 - \frac{2(1-p)}{(1-2p)} \mu \right),$$

$f_3(0) = 1$  e  $f_3'(0) = -\frac{2(1-p)}{(1-2p)} + 2$ , novamente pela Definição 4.1.1, resulta a seguinte aproximação para  $f_3$ ,

$$f_3(\mu) \approx 1 - \frac{2p}{1-2p} \mu.$$

Por conseguinte,

$$q(N, \mu, p) \approx \left( 1 - \frac{2p}{1-2p} \mu \right)^N, \quad \text{para } p < 1/2. \quad (4.15)$$

Daqui, usando o fato  $e^{-x} > 1 - x$ , para todo  $x > 0$ , podemos obter um limitante para a aproximação em (4.15), isto é,

$$q(N, \mu, p) \approx \left(1 - \frac{2p}{1-2p}\mu\right)^N < \exp\left(-\frac{2p}{1-2p}N\mu\right).$$

Assim, pode-se ver que para  $p < 1/2$  fixo, a probabilidade de não ter resistência depende do parâmetro  $m \equiv N\mu$  e decresce exponencialmente quando  $m$  aumenta.

Para  $p > 1/2$ , pode ser obtida uma expressão semelhante, isto é, usando novamente as aproximações de  $f_1$  e  $f_2$ , resulta a seguinte aproximação para (4.12),

$$\begin{aligned} q(N, \mu, p) &\approx \left(\frac{(1+2\mu)}{2p} \left(1 + (1-2p) \left(1 + \frac{4p(1-p)}{(1-2p)^2}\mu\right)\right)\right)^N \\ &= \left(\frac{(1+2\mu)}{p} \left((1-p) + \frac{2p(1-p)}{(1-2p)}\mu\right)\right)^N \\ &= \left((1+2\mu)\frac{(1-p)}{p} \left(1 + \frac{2p}{(1-2p)}\mu\right)\right)^N. \end{aligned} \quad (4.16)$$

Daqui, tomando a função  $f_4(\mu) = (1+2\mu)\frac{(1-p)}{p} \left(1 + \frac{2p}{(1-2p)}\mu\right)$  em (4.16), segue que

$$f_4'(\mu) = \left((1+2\mu)\frac{(1-p)}{p}\right) \left(\frac{2p}{(1-2p)}\right) + \left(\frac{2(1-p)}{p}\right) \left(1 + \frac{2p}{(1-2p)}\mu\right),$$

$f_4(0) = \frac{(1-p)}{p}$  e  $f_4'(0) = \frac{(1-p)}{p} \left(\frac{2p}{(1-2p)}\right) + \frac{2(1-p)}{p}$ , assim, pela Definição 4.1.1, temos que

$$\begin{aligned} f_4(\mu) &\approx \frac{(1-p)}{p} + \left(\frac{(1-p)}{p} \left(\frac{2p}{(1-2p)}\right) + \frac{2(1-p)}{p}\right)\mu \\ &= \frac{1-p}{p} + \left(\frac{(1-p)2(1-p)}{p(1-2p)}\right)\mu \\ &= \frac{(1-p)}{p} \left(1 - \frac{2(1-p)}{(2p-1)}\mu\right). \end{aligned} \quad (4.17)$$

Logo, inserindo (4.17) em (4.16), segue

$$q(N, \mu, p) \approx \left(\frac{(1-p)}{p} \left(1 - \frac{2(1-p)}{(2p-1)}\mu\right)\right)^N, \quad \text{para } p > 1/2. \quad (4.18)$$

Daqui, usando novamente o fato  $e^{-x} > 1 - x$ , para todo  $x > 0$ , podemos obter um limitante para a aproximação em (4.18),

$$q(N, \mu, p) \approx \left( \frac{(1-p)}{p} \left( 1 - \frac{2(1-p)}{(2p-1)} \mu \right) \right)^N < \left( \frac{1-p}{p} \right)^N \exp \left( -\frac{2(1-p)}{(2p-1)} N \mu \right),$$

mas principalmente, temos que

$$q(N, \mu, p) < \left( \frac{1-p}{p} \right)^N. \quad (4.19)$$

Isto é, se  $N$  for suficientemente grande, é difícil evitar que a resistência aos medicamentos aconteça.

## 4.2 Probabilidade de resistência com duas drogas

Com o intuito de estender o modelo proposto por Schinazi (2006), descreve-se agora, um modelo para calcular a probabilidade de resistência quando no tratamento são ministradas duas drogas.

### 4.2.1 Descrição do modelo

Considera-se um processo de Galton-Watson a tempo discreto de quatro tipos para modelar a probabilidade de resistência num tratamento com duas drogas. Sejam 0, 1, 2, 3 os quatro tipos de partículas, em que cada tipo, representa um patógeno que pode ser sensível ou resistente para uma ou ambas drogas. Isto é, o tipo 0 representa um patógeno sensível para as duas drogas, o tipo 1 representa um patógeno resistente para a primeira droga, o tipo 2 representa um patógeno resistente para a segunda droga e o tipo 3 um resistente para as duas drogas. Assuma-se que em cada passo de tempo, as partículas dos tipos 0, 1, 2 podem dividir-se de acordo a uma variável aleatória  $X'$  i.i.d à variável aleatória  $X$  definida no modelo da secção 4.1, o que significa que a média de descendentes para cada tipo  $i$ ,  $i = 0, 1, 2$ , é  $2p$ . Além disso suponha que em cada divisão das partículas 0, 1, 2; existem probabilidades  $\mu_1$  e  $\mu_2$  de obter resistência para a primeira e segunda droga respectivamente. Cada descendente de uma partícula do tipo 0 pode mutar para o tipo 1 com probabilidade  $\mu_{01} := \mu_1(1 - \mu_2)$ , pode também mutar para o tipo 2 com probabilidade  $\mu_{02} := \mu_2(1 - \mu_1)$  e pode mutar para o tipo 3 com probabilidade  $\mu_{03} := \mu_1\mu_2$ . Os descendentes de uma partícula do tipo 1 podem unicamente mutar para o tipo 3 com probabilidade  $\mu_{13} := \mu_2$  e por último os descendentes de uma partícula do tipo 2 podem

unicamente mutar para o tipo 3 com probabilidade  $\mu_{23} := \mu_1$ . As partículas do tipo 3 somente podem dividir-se, isto é, elas não morrem e além disso elas somente produzem partículas do tipo 3, assim, se alguma das partículas dos tipos 0, 1, 2 obtém uma mutação do tipo 3, existirá uma probabilidade positiva do processo continuar indefinidamente, o que em outras palavras significa a falha do tratamento.

Formalmente, seja  $(\mathbf{Z}_n)_{n \geq 0}$  o processo de Galton-Watson de quatro tipos tal que  $\mathbf{Z}_n = (Z_n^0, Z_n^1, Z_n^2, Z_n^3)$ , onde  $Z_n^i$  representa o número de patógenos do tipo  $i$ ,  $i = 0, 1, 2, 3$  na  $n$ -ésima geração. A f.g.p da distribuição de descendentes do tipo  $i$ , será definida por  $g^i$ ,  $i = 0, 1, 2, 3$ . Assim a f.g.p de  $\mathbf{Z}_n$  dado que  $\mathbf{Z}_0 = \mathbf{e}_i$ , (onde  $\mathbf{e}_i$ ,  $i = 0, 1, 2, 3$ , denota o vetor cuja  $i$ -ésima componente é 1 e cujas outras componentes são 0) é dada por  $g_n^i(s_0, s_1, s_2, s_3) = g_n^i(\mathbf{s})$ ,  $n \geq 0$ ;  $i = 0, 1, 2, 3$ . Consequentemente a f.g.p de  $\mathbf{Z}_1$  dado que  $\mathbf{Z}_0 = \mathbf{e}_0$  é dada por

$$\begin{aligned} g_1^0(\mathbf{s}) &= g^0(\mu_{00}s_0 + \mu_{01}s_1 + \mu_{02}s_2 + \mu_{03}s_3) \\ &= g^1((1 - \mu_1(1 - \mu_2) - \mu_2(1 - \mu_1) - \mu_1\mu_2)s_0 + \mu_1(1 - \mu_2)s_1 + \mu_2(1 - \mu_1)s_2 + \mu_1\mu_2s_3). \end{aligned}$$

A f.g.p de  $\mathbf{Z}_1$  dado que  $\mathbf{Z}_0 = \mathbf{e}_1$  é dada por

$$\begin{aligned} g_1^1(\mathbf{s}) &= g^2(\mu_{11}s_1 + \mu_{13}s_3) \\ &= g^1((1 - \mu_2)s_1 + \mu_2s_3). \end{aligned}$$

E por último, a f.g.p de  $\mathbf{Z}_1$  dado que  $\mathbf{Z}_0 = \mathbf{e}_2$  é dada por

$$\begin{aligned} g_1^2(\mathbf{s}) &= g^3(\mu_{22}s_2 + \mu_{23}s_3) \\ &= g^2((1 - \mu_1)s_2 + \mu_1s_3). \end{aligned}$$

Seja  $q_i$  a probabilidade de extinção dado que o processo foi iniciado por uma partícula do tipo  $i$ ,  $i = 0, 1, 2, 3$ , isto é,  $q_i = \mathbb{P}[\mathcal{E} \mid \mathbf{Z}_0 = \mathbf{e}_i]$ , onde  $\mathcal{E} := \bigcup_{n \geq 1} \{\mathbf{Z}_n = \mathbf{0}\}$  e seja o vetor  $\mathbf{q} = (q_0, q_1, q_2, q_3)$ . A probabilidade de obter uma mutação para o tipo 3 (patógeno resistente para as duas drogas), dado que o processo foi iniciado por uma partícula do tipo  $i$ ,  $i = 0, 1, 2, 3$  é definida por

$$r_i = 1 - q_i. \quad (4.20)$$

No que segue, obteremos uma aproximação para cada  $r_i$ . Note que

$$q_i = g^i(\mu_{i0}q_0 + \mu_{i1}q_1 + \mu_{i2}q_2 + \mu_{i3}q_3). \quad (4.21)$$



Com efeito, dado que

$$\begin{aligned}
q_i &= \mathbb{P}[\mathcal{E} \mid \mathbf{Z}_0 = e_i] \\
&= \sum_{m=0,2} \mathbb{P}[\mathcal{E}, \mathbf{Z}_1 = m \mid \mathbf{Z}_0 = e_i] \\
&= \sum_{m=0,2} \mathbb{P}[\mathcal{E} \mid \mathbf{Z}_1 = m, \mathbf{Z}_0 = e_i] \mathbb{P}[\mathbf{Z}_1 = m \mid \mathbf{Z}_0 = e_i] \\
&= \sum_{m=0,2} \mathbb{P} \left[ \sum_{i=1}^m Z_1^i = \mathbf{0} \right] \mathbb{P}[\mathbf{Z}_1 = m] \\
&= \sum_{m=0,2} \mathbb{P}[\mathcal{E}]^m \mathbb{P}[\mathbf{Z}_1 = m]. \tag{4.22}
\end{aligned}$$

Agora, se  $A_j$  denota o evento “Uma partícula do tipo  $i$  obtem uma mutação para o tipo  $j$ ”, para  $i \neq j$ , temos que

$$\mathbb{P}[\mathcal{E}] = \sum_{j=0}^3 \mathbb{P}[\mathcal{E} \mid A_j] \mathbb{P}[A_j] = \sum_{j=0}^3 q_j \mu_{ij}. \tag{4.23}$$

Substituindo (4.23) em (4.22) segue (4.21)

$$q_i = \sum_{m=0,2} \left( \sum_{j=0}^3 \mu_{ij} q_j \right)^m \mathbb{P}[\mathbf{Z}_1 = m] = g^i \left( \sum_{j=0}^3 \mu_{ij} q_j \right). \tag{4.24}$$

Então usando (4.21), temos

$$q_0 = g^0(\mu_{00}q_0 + \mu_{01}q_1 + \mu_{02}q_2 + \mu_{03}q_3),$$

usando (4.20), segue

$$\begin{aligned}
1 - r_0 &= g^0((1 - \mu_1(1 - \mu_2) - \mu_2(1 - \mu_1) - \mu_1\mu_2)(1 - r_0) + \mu_1(1 - \mu_2)(1 - r_1) + \\
&\quad \mu_2(1 - \mu_1)(1 - r_2) + \mu_1\mu_2(1 - r_3))
\end{aligned}$$

$$\begin{aligned}
r_0 &= 1 - g^0(1 - [(1 - \mu_1(1 - \mu_2) - \mu_2(1 - \mu_1) - \mu_1\mu_2)r_0 + \mu_1(1 - \mu_2)r_1 + \mu_2(1 - \mu_1)r_2 + \\
&\quad \mu_1\mu_2r_3]).
\end{aligned}$$

Agora usando a fórmula de Taylor com resto de Lagrange para  $g^0$  em torno do ponto 1, para algum  $0 \leq \theta \leq 1$ , segue

$$r_0 = 1 - g^{0'}(1 - [(1 - \mu_1(1 - \mu_2) - \mu_2(1 - \mu_1) - \mu_1\mu_2)r_0 + \mu_1(1 - \mu_2)r_1 + \mu_2(1 - \mu_1)r_2 + \mu_1\mu_2r_3] - 1) + g^{0''}(1 - \theta[(1 - \mu_1(1 - \mu_2) - \mu_2(1 - \mu_1) - \mu_1\mu_2)r_0 + \mu_1(1 - \mu_2)r_1 + \mu_2(1 - \mu_1)r_2 + \mu_1\mu_2r_3])(1 - [(1 - \mu_1(1 - \mu_2) - \mu_2(1 - \mu_1) - \mu_1\mu_2)r_0 + \mu_1(1 - \mu_2)r_1 + \mu_2(1 - \mu_1)r_2 + \mu_1\mu_2r_3] - 1)^2.$$

Dado que  $g^0(1) = 1$ ,  $g^{0'}(1) = 2p$  e escrevendo  $(1 - \mu_1(1 - \mu_2) - \mu_2(1 - \mu_1) - \mu_1\mu_2)r_0 + \mu_1(1 - \mu_2)r_1 + \mu_2(1 - \mu_1)r_2 + \mu_1\mu_2r_3 =: x$ , temos

$$r_0 = 2p(1 - \mu_1(1 - \mu_2) - \mu_2(1 - \mu_1) - \mu_1\mu_2)r_0 + 2p(\mu_1(1 - \mu_2)r_1 + \mu_2(1 - \mu_1)r_2 + \mu_1\mu_2r_3) - \frac{1}{2}g^{0''}(1 - \theta x)x^2.$$

Daqui, segue

$$r_0 = \frac{2p}{1 - 2p\mu_{00}}(\mu_{01}r_1 + \mu_{02}r_2 + \mu_{03}r_3) - \frac{g^{0''}(1 - \theta x)x^2}{2(1 - 2p\mu_{00})}. \quad (4.25)$$

Agora, notemos que

$$\frac{g^{0''}(1 - \theta x)x^2}{2(1 - 2p\mu_{00})} \leq \frac{g^{0''}(1)x^2}{1 - 2p\mu_{00}} \leq \frac{\sigma^2 x^2}{1 - 2p}. \quad (4.26)$$

Onde  $\sigma^2 = Var(Z_1^0)$ . Na primeira desigualdade usamos o fato que  $g^{0''}$  é uma função não decrescente e na segunda desigualdade consideramos  $2p < 1$ . Observe que se consideramos  $\sigma^2 < \infty$  e dado que  $x^2$  é de ordem  $\mu^2$ , então resulta que a última expressão em (4.26) também é de ordem  $\mu^2$ , com  $\mu$  pequeno.

Agora, usando novamente a fórmula de Taylor com resto de Lagrange para  $\frac{2p}{1 - 2p\mu_{00}}$ , em torno do ponto 1, para algum  $0 \leq \beta \leq 1$ , obtemos

$$\frac{2p}{1 - 2p\mu_{00}} = \frac{2p}{1 - 2p} + \frac{(2p)^2}{(1 - 2p)^2}(\mu_{01} + \mu_{02} + \mu_{03}) + \frac{(2p)^3(\mu_{01} + \mu_{02} + \mu_{03})^2}{[1 - (1 - \beta(\mu_{01} + \mu_{02} + \mu_{03}))2p]^3}. \quad (4.27)$$

Onde

$$\frac{(2p)^3(\mu_{01} + \mu_{02} + \mu_{03})^2}{[1 - (1 - \beta(\mu_{01} + \mu_{02} + \mu_{03}))2p]^3} \leq \frac{(2p)^3(\mu_{01} + \mu_{02} + \mu_{03})^2}{(1 - 2p)^3}.$$

Substituindo (4.27) na equação (4.25) e considerando que  $\mu_{01} + \mu_{02} + \mu_{03}$  é de ordem  $\mu$ , para  $\mu$  pequeno e desconsiderando todos os termos de ordem maior ou igual do que  $\mu^2$ , obtemos uma aproximação da probabilidade de mutação para o tipo 3 (resistência para as duas drogas) dado que o processo foi iniciado com uma partícula do tipo 0, que é dada

por

$$r_0 \approx \frac{2p}{1-2p}(\mu_1(1-\mu_2)r_1 + \mu_2(1-\mu_1)r_2 + \mu_1\mu_2r_3). \quad (4.28)$$

Da mesma forma podemos obter

$$r_1 \approx \frac{2p}{1-2p}\mu_2r_3 \quad (4.29)$$

$$r_2 \approx \frac{2p}{1-2p}\mu_1r_3 \quad (4.30)$$

Dado que as partículas do tipo 3 só se dividem, isto é, elas nunca morrem e além disso elas não mutam, ou seja, não existe mutação desde o tipo 3 para qualquer outra partícula do tipo  $i$ ,  $i = 0, 1, 2$ , isso significa que a probabilidade de mutação para o tipo 3 dado que processo foi iniciado com uma partícula do tipo 3,  $r_3$ , é sempre 1, ou seja  $r_3 = 1$ .

Por outro lado, mostraremos que o vetor  $\mathbf{r}$ , das aproximações para a probabilidade de mutação para o tipo 3 dado que o processo foi iniciado por alguma partícula do tipo  $i$ ,  $i = 0, 1, 2$ , ou seja,  $\mathbf{r} = (r_0, r_1, r_2)^T$ , é a solução de um sistema linear de equações. Então, seja  $D$  uma matriz diagonal da seguinte forma

$$D = \begin{pmatrix} \frac{2p}{1-2p\mu_{00}} & 0 & 0 \\ 0 & \frac{2p}{1-2p\mu_{11}} & 0 \\ 0 & 0 & \frac{2p}{1-2p\mu_{22}} \end{pmatrix}$$

Agora, seja  $M$  a matriz das probabilidades de mutação desde uma partícula do tipo  $i$ ,  $i = 0, 1, 2$  para uma do tipo  $j$ , com  $i \neq j$ , mas dado que apenas existem as probabilidades de mutação  $\mu_{01}$  e  $\mu_{02}$ , a matriz  $M$  fica da seguinte forma

$$M = \begin{pmatrix} 0 & \mu_{01} & \mu_{02} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Também, seja  $\mathbf{u}_3$  o vetor das probabilidades de mutação desde uma partícula do tipo  $i$ ,  $i = 0, 1, 2$  para o tipo 3, isto é,  $\mathbf{u}_3 = (\mu_{03}, \mu_{13}, \mu_{23})^T$ . Então desde as equações (4.28), (4.29), (4.30) temos que  $\mathbf{r}$  satisfaz aproximadamente o sistema de equações

$$\mathbf{r} = D\mathbf{M}\mathbf{r} + D\mathbf{u}_3r_3,$$

que tem a seguinte solução

$$\mathbf{r} = [I - DM]^{-1} D\mathbf{u}_3 r_3,$$

onde  $I$  é uma matriz identidade de ordem  $3 \times 3$ . Em efeito, temos que

$$[I - DM] = \begin{pmatrix} 1 & -\frac{2p\mu_{01}}{1-2p\mu_{00}} & -\frac{2p\mu_{02}}{1-2p\mu_{00}} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Cuja inversa é

$$[I - DM]^{-1} = \begin{pmatrix} 1 & \frac{2p\mu_{01}}{1-2p\mu_{00}} & \frac{2p\mu_{02}}{1-2p\mu_{00}} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Por outro lado,

$$[I - DM]^{-1} D = \begin{pmatrix} \frac{2p}{1-2p\mu_{00}} & \frac{2p\mu_{01}}{1-2p\mu_{00}} & \frac{2p}{1-2p\mu_{11}} & \frac{2p\mu_{02}}{1-2p\mu_{00}} & \frac{2p}{1-2p\mu_{22}} \\ 0 & \frac{2p}{1-2p\mu_{11}} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{2p}{1-2p\mu_{22}} & 0 \end{pmatrix}$$

Assim, temos que

$$\mathbf{r} = [I - DM]^{-1} D\mathbf{u}_3 r_3 = \begin{pmatrix} \frac{2p\mu_{03}}{1-2p\mu_{00}} + \frac{2p\mu_{01}}{1-2p\mu_{00}} \frac{2p\mu_{13}}{1-2p\mu_{11}} + \frac{2p\mu_{02}}{1-2p\mu_{00}} \frac{2p\mu_{23}}{1-2p\mu_{22}} \\ \frac{2p\mu_{13}}{1-2p\mu_{11}} \\ \frac{2p\mu_{23}}{1-2p\mu_{22}} \end{pmatrix}$$

Assim, foi obtida uma aproximação para a probabilidade da emergência da resistência quando no tratamento são ministradas duas drogas.

**Observação 4.2.1.** *Se consideramos apenas dois tipos de partículas 0, 1, sendo o tipo 0 um patógeno sensível e o tipo 1 um patógeno resistente e além disso, assumimos que  $\mu_1 = \mu$  e  $\mu_2 = 0$ . Então, temos que a aproximação da probabilidade de mutação para o tipo 1 dado que o processo foi iniciado por uma partícula do tipo 0, é dada por*

$$r_0 \approx \frac{2p}{1-2p} \mu,$$

---

*pois  $r_1 = 1$  neste caso. Deste modo, particularmente o presente modelo pode ser também visto como o modelo da secção 4.1.*

## Martingale

Estudamos algumas definições e resultados elementares de convergência de Martingales. O presente apêndice está baseado na secção 5.2 do livro de Durrett (2010).

**Definição A.0.1.** *Seja  $(\mathcal{F}_n)_{n \geq 0}$  uma filtragem, isto é, uma sequência crescente de  $\sigma$ -álgebras num espaço de probabilidade  $(\Omega, \mathcal{F}, P)$ . A sequência  $(X_n)_{n \geq 0}$  de variáveis aleatórias é dita de ser adaptada para  $\mathcal{F}_n$  se  $X_n \in \mathcal{F}_n$ , para todo  $n \geq 0$ . Se  $(X_n)_{n \geq 0}$  é uma sequência com*

- (i)  $\mathbb{E}[|X_n|] < \infty$ ,
- (ii)  $X_n$  é adaptada para  $\mathcal{F}_n$ ,
- (iii)  $\mathbb{E}[X_{n+1} | \mathcal{F}_n] = X_n$ , para todo  $n$ ,

então  $(X_n)_{n \geq 0}$  é dita de ser um martingale com respeito a  $(\mathcal{F}_n)_{n \geq 0}$ . Da mesma forma  $(X_n)_{n \geq 0}$  diz-se um supermartingale (respectivamente submartingale), se para todo  $n$ ,

$$\mathbb{E}[X_{n+1} | \mathcal{F}_n] \leq (\geq) X_n.$$

**Teorema A.0.2.** *Se  $X_n$  é um submartingale com respeito a  $(\mathcal{F}_n)_{n \geq 0}$  e  $\varphi$  é uma função convexa crescente com  $\mathbb{E}[|\varphi(X_n)|] < \infty$ , então  $\varphi(X_n)$  é um submartingale com respeito a  $(\mathcal{F}_n)_{n \geq 0}$ . Consequentemente, se  $X_n$  é um submartingale, então  $(X_n - a)^+$  é um submartingale.*

*Demonstração.* Temos que

$$\mathbb{E}[\varphi(X_{n+1}) \mid \mathcal{F}_n] \geq \varphi(\mathbb{E}[X_{n+1} \mid \mathcal{F}_n]) \geq \varphi(X_n),$$

onde a primeira desigualdade se cumpre por  $X_{n+1}$  ser uma variável aleatória e  $\varphi$  uma função convexa (desigualdade de Jensen) e a segunda desigualdade por ser  $X_n$  um submartingale.  $\square$

Seja  $(\mathcal{F}_n)_{n \geq 0}$  uma filtragem e seja  $(H_n)_{n \geq 0}$  uma sequência que é chamada de previsível, se  $H_n \in \mathcal{F}_{n-1}$ , para todo  $n \neq 1$ . Em palavras, o valor de  $H_n$  pode ser previsível a partir da informação disponível no tempo  $n - 1$ . Então

$$(H.X)_n = \sum_{k=1}^n H_k(X_k - X_{k-1}), \quad (\text{A.1})$$

representa os ganhos de um jogador até o instante  $n$ . Daqui, temos o seguinte resultado.

**Teorema A.0.3.** *Seja  $(X_n)_{n \geq 0}$  um submartingale. Se  $H_n > 0$  é previsível e cada  $H_n$  é acotada, então  $(H.X)_n$  é um submartingale.*

*Demonstração.* Seja,

$$\begin{aligned} \mathbb{E}[(H.X)_{n+1} \mid \mathcal{F}_n] &= \mathbb{E}[(H.X)_n + H_{n+1}(X_{n+1} - X_n) \mid \mathcal{F}_n] \\ &= (H.X)_n + H_{n+1}\mathbb{E}[(X_{n+1} - X_n) \mid \mathcal{F}_n] \geq (H.X)_n, \end{aligned}$$

onde a segunda igualdade segue pela linearidade da esperança condicional,  $(H.X)_n \in \mathcal{F}_n$ ,  $H_n \in \mathcal{F}_{n-1}$ . A desigualdade se cumpre por  $\mathbb{E}[(X_{n+1} - X_n) \mid \mathcal{F}_n] \geq 0$  e  $H_{n+1} \geq 0$ .  $\square$

Por outro lado, uma variável aleatória  $T$  é dita de ser um tempo de parada se  $[T = n] \in \mathcal{F}_n$ , para todo  $n < \infty$ . Agora, defina  $T_0 = -1$  e para  $r \geq 1$ , seja

$$\begin{aligned} T_{2r-1} &= \inf\{k > T_{2r-2} : X_k \leq a\}, \\ T_{2r} &= \inf\{k > T_{2r-1} : X_k \geq b\}. \end{aligned}$$

O número de passagens de baixo para cima (“Upcrossings”) de um intervalo  $[a, b]$ , isto é, o número de vezes que um processo cruza vindo de baixo de  $a$  para cima de  $b$ , no instante  $n$ , é dado por

$$U_n = \sup\{r : T_{2r} \leq n\}.$$

**Teorema A.0.4.** *Se  $(X_k)_{k \geq 0}$  é um submartingale, então*

$$(b - a)\mathbb{E}[U_n] \leq \mathbb{E}[(X_n - a)^+] - \mathbb{E}[(X_0 - a)^+].$$

*Demonstração.* Tome  $Y_k = a + (X_k - a)^+$ , daqui pelo Teorema A.0.2 segue que  $Y_k$  é um submartingale. Podemos ver que  $Y_k$  cruza  $[a, b]$  o mesmo número de vezes que  $X_k$  e dado que cada “Upcrossing” resulta em um lucro  $\geq (b - a)$  segue que  $(b - a)U_n \leq (H.Y)_n$  e se existir um upcrossing final incompleto então este faz uma contribuição não negativa para o lado direito é por isso que  $X_k$  foi substituído por  $Y_k$ . Seja  $M_k = 1 - H_k$ . Então, segue que

$$Y_n - Y_0 = (H.Y)_n + (M.Y)_n. \quad (\text{A.2})$$

Com efeito,

$$\begin{aligned} (H.Y)_n + (M.Y)_n &= \sum_{k=1}^n H_k(Y_k - Y_{k-1}) + \sum_{k=1}^n M_k(Y_k - Y_{k-1}) \\ &= \sum_{k=1}^n H_k(Y_k - Y_{k-1}) + \sum_{k=1}^n (Y_k - Y_{k-1}) - \sum_{k=1}^n H_k(Y_k - Y_{k-1}) \\ &= \sum_{k=1}^n (Y_k - Y_{k-1}) = Y_n - Y_0. \end{aligned}$$

E isso mostra (A.2). Logo pelo Teorema A.0.3 temos que  $\mathbb{E}[(k.Y)_n] \geq \mathbb{E}[(K.Y)_0] = 0$ . Dessa forma, desde (A.2), temos

$$\begin{aligned} \mathbb{E}[Y_n - Y_0] &= \mathbb{E}[(H.Y)_n + (M.Y)_n] \\ &= \mathbb{E}[(H.Y)_n] + \mathbb{E}[(M.Y)_n] \geq \mathbb{E}[(H.Y)_n]. \end{aligned}$$

Por conseguinte,

$$(b - a)\mathbb{E}[U_n] \leq \mathbb{E}[(H.Y)_n] \leq \mathbb{E}[(X_n - a)^+] - \mathbb{E}[(X_0 - a)^+].$$

□

O próximo teorema é uma consequência do teorema acima.

**Teorema A.0.5.** *Se  $(X_n)_{n \geq 0}$  é um submartingale com  $\sup \mathbb{E}[X_n^+] < \infty$ , então  $\lim_{n \rightarrow \infty} X_n = X$ , quase certamente, com  $\mathbb{E}[|X|] < \infty$ .*



*Demonstração.* Dado que  $(X_n - a)^+ \leq X^+ + |a|$ , pelo Teorema A.0.4, segue que

$$\mathbb{E}[U_n] \geq \frac{(|a| + \mathbb{E}[X_n^+])}{(b - a)}.$$

Agora, seja  $U = \lim_{n \rightarrow \infty} U_n$ . Logo, por convergência monótona e dado que  $\sup \mathbb{E}[X_n^+] < \infty$ , temos

$$\mathbb{E}[U] = \lim_{n \rightarrow \infty} \mathbb{E}[U_n] < \infty,$$

e portanto,  $U < \infty$  quase certamente. Dado que a última conclusão se cumpre para cada  $a, b$  racionais, tomando a união sobre todos os pares de racionais  $a, b$ , temos que quase certamente a sequência  $X_n$  não poderá ter  $\limsup X_n > \liminf X_n$ , isto é,

$$\bigcup_{a, b \in \mathbb{Q}} \{\liminf X_n < a < b < \limsup X_n\},$$

tem probabilidade zero. Logo,  $\limsup X_n = \liminf X_n$  quase certamente, isto é,  $\lim_{n \rightarrow \infty} X_n$  existe quase certamente. Por outro lado, já que  $X_n$  é um martingal, resulta que  $\mathbb{E}[X_n] \geq \mathbb{E}[X_0]$  e além disso,

$$\begin{aligned} \mathbb{E}[|X_n|] &= \mathbb{E}[X_n^+] + \mathbb{E}[X_n^-] \\ &= 2\mathbb{E}[X_n^+] - \mathbb{E}[X_n] \\ &\leq 2\mathbb{E}[X_n^+] - \mathbb{E}[X_0]. \end{aligned}$$

Logo, pelo Lema de Fatou, segue

$$\mathbb{E}[|X|] = \mathbb{E}[\liminf |X_n|] \leq \liminf_n \mathbb{E}[|X_n|] \leq \sup_n \mathbb{E}[|X_n|] < \infty. \quad (\text{A.3})$$

Isto é, o limite é finito.

**Teorema A.0.6.** *Se  $(X_n)_{n \geq 0}$  é um supermartingale não negativo, então  $(X_n)_{n \geq 0}$  converge para um limite finito, quase certamente.*

*Demonstração.* A prova pode ser feita considerando  $Y_n = -X_n \leq 0$ , desde que  $X_n$  é um supermartingale não negativo, então resulta que  $Y_n$  é um submartingale limitado por cima, por zero. Logo, pelo Teorema A.0.5,  $Y_n$  converge quase certamente para um limite finito.  $\square$

**Corolário A.0.7.** *Todo martingale não negativo, converge quase certamente.*

A prova do item (i) do Teorema 2.2.1 no Capítulo 2, segue diretamente do Corolário A.0.7.

**Teorema A.0.8.** *Se  $X_n$  é um martingale com  $\sup_n \mathbb{E}[|X_n|^p] < \infty$ , onde  $p > 1$ , então  $X_n$  converge para  $X$  quase certamente e em  $L^p$ .*

*Demonstração.* Como  $\sup_n \mathbb{E}[|X_n|^p] < \infty$ , pelo Teorema A.0.5, segue que  $X_n$  converge para  $X$  quase certamente. Agora, para a convergência em  $L^p$ , usamos a  $L^p$ -desigualdade (ver Durrett (2010)). Isto é,

$$\mathbb{E} \left[ \left( \sup_{0 \leq m \leq n} |X_m| \right)^p \right] \leq \left( \frac{p}{p-1} \right)^p \mathbb{E}[|X_n|^p]. \quad (\text{A.4})$$

Se  $\sup_{0 \leq m \leq n} |X_m| \uparrow \sup_{0 \leq m} |X_m|$ , tomando limite quando  $n \rightarrow \infty$  em (A.4), pelo Teorema da convergência monótona, temos

$$\mathbb{E} \left[ \left( \sup_{0 \leq m} |X_m| \right)^p \right] \leq \left( \frac{p}{p-1} \right)^p \sup_n \mathbb{E}[|X_n|^p] < \infty \quad (\text{A.5})$$

Deste modo, resulta que

$$\sup_{0 \leq m} |X_m| \in L^p.$$

Dado que  $|X_n - X| \leq 2 \sup |X_n|$ . Pelo Teorema da convergência dominada, temos que

$$\lim_{n \rightarrow \infty} \mathbb{E}[|X_n - X|^p] = 0.$$

□

O item (ii) do Teorema 2.2.1 é um caso especial do Teorema A.0.8, para  $p = 2$ .



## Referências Bibliográficas

---

---

- ALLEN, L. J. *An introduction to stochastic processes with applications to biology*. Pearson Education Upper Saddle River, NJ, 2003.
- ATHREYA, K. B.; NEY, P. E. *Branching processes*. Springer, 1972.
- DURRETT, R. *Probability: theory and examples*. Cambridge university press, 2010.
- GUT, A. *An intermediate course in probability*. 2 ed. Springer, 2009.
- HARRIS, T. E. *The theory of branching processes*. Courier Dover Publications, 2002.
- JAGERS, P. *Branching processes with biological applications*. John Wiley Sons Ltd, 1975.
- KARLIN, S.; TAYLOR, H. M. *An introduction to stochastic modeling*. Academic Press, 1998.
- OTTER, R. *The multiplicative process*. JSTOR, 206–224 p., 1949.
- RESNICK, S. I. *A probability path*. Springer Science & Business Media, 2013.
- SCHINAZI, R. B. *Classical and spatial stochastic processes*. Springer, 1999.
- SCHINAZI, R. B. *The probability of treatment induced drug resistance*, v. 54. Springer, 13–19 p., 2006.