

# Processos de ramificação: teoria e aplicações

Pablo Martín Rodríguez  
ICMC-USP, São Carlos  
pablor@icmc.usp.br

II Colóquio de Matemática da Região Sul  
Universidade Estadual de Londrina (UEL), 24 ao 28 de abril de 2012

# Sumário

<b>Introdução</b>	<b>3</b>
<b>1 Definições preliminares</b>	<b>4</b>
<b>2 Processos de Bienaymé-Galton-Watson</b>	<b>7</b>
2.1 Um pouco de história . . . . .	7
2.2 Definição . . . . .	8
2.3 Funções geradoras de probabilidade . . . . .	11
2.4 Probabilidade de extinção . . . . .	13
<b>3 Aplicações e generalizações</b>	<b>18</b>
3.1 Outros exemplos e aplicações . . . . .	18
3.2 Generalizações . . . . .	20
3.2.1 Processos de ramificação em meios variáveis . . . . .	20
3.2.2 Processos de ramificação multitempo . . . . .	21
3.2.3 Processos de ramificação a tempo contínuo . . . . .	21
<b>Referências</b>	<b>24</b>

## Introdução

A noção matemática de processo de ramificação foi introduzida independentemente por Bienaymé e por Galton e Watson com o intuito de estudar a probabilidade de extinção do sobrenome de uma família. Desde então, esta teoria também é conhecida por suas aplicações em campos como a física e a biologia, entre outros.

De modo geral podemos pensar um processo de ramificação como partículas que geram partículas do mesmo tipo e em número distribuído de acordo a uma variável aleatória discreta.

Neste minicurso pretende-se apresentar as ideias básicas detrás da formulação matemática destes processos e suas aplicações. Estas notas contêm vários exemplos que aparecem em diferentes livros que abordam o tema e as respectivas referências são mencionadas ao longo do texto.

# 1 Definições preliminares

Antes de começar com nossa introdução aos processos de ramificação precisamos introduzir a noção de processo estocástico e de cadeia de Markov. Os livros de Karlin e Taylor (1998) [7] e Schinazi (1999) são ótimas referências de processos estocásticos.

**Definição 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. Nestas notas vamos assumir  $S = \mathbb{Z}$ .

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 ao longo desse tempo. Assim, se lançamos sucessivamente uma moeda e  $X_n$  representa o resultado do  $n$ -ésimo lançamento (cara ou coroa), a sequência  $(X_n)_{n \geq 1}$  é um processo estocástico a tempo discreto. Da mesma forma, 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 o primeiro registro da infecção, a sequência  $(Y_n)_{n \geq 1}$  é um processo estocástico a tempo discreto.

No que segue vemos outro exemplo de processo estocástico bem conhecido na teoria de probabilidade.

**Exemplo 1.2.** *Passeio aleatório em  $\mathbb{Z}$ .* Suponha que queremos modelar o seguinte fenômeno. Em um instante de tempo  $n = 0$  temos uma partícula no vértice  $0 \in \mathbb{Z}$ . A cada instante discreto de tempo a partícula pula um vértice à direita com probabilidade  $p$  ou um vértice à esquerda com probabilidade  $1 - p$  (Figura 1). Se  $X_n$  representa a posição da partícula no  $n$ -ésimo instante de tempo. Então  $X_n$  toma valores no conjunto  $\{0, 1, 2, \dots\}$  e a sequência  $(X_n)_{n \geq 0}$  é um processo estocástico a tempo discreto. Esse processo é chamado de passeio aleatório em  $\mathbb{Z}$ .  $\square$

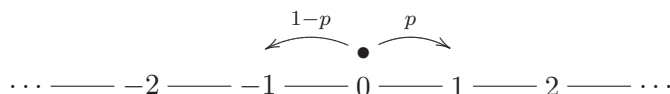


Figura 1: Passeio aleatório em  $\mathbb{Z}$ .

No exemplo anterior, notemos que para saber a posição da partícula no instante de tempo  $n$  é suficiente saber sua posição no tempo  $n - 1$ , independente das possíveis posições da partícula em instantes anteriores. Esta propriedade é conhecida como propriedade de Markov e nos permite definir uma classe muito importante de processos estocásticos, isto é, as cadeias de Markov.

**Definição 1.3.** Uma cadeia de Markov a tempo discreto é um processo estocástico a tempo discreto  $(X_n)_{n \geq 0}$  tal que

$$P(X_{n+1} = j | X_0 = i_0, X_1 = i_1, X_2 = i_2, \dots, X_{n-1} = i_{n-1}, X_n = i) = P(X_{n+1} = j | X_n = i),$$

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 o futuro depende somente do presente do processo.

**Observação 1.4.** Da Definição 1.3 vemos que, para todo  $i, j \in \mathbb{Z}$  e para todo  $n \geq 0$ , as probabilidades

$$p(i, j) := P(X_{n+1} = j | X_n = i),$$

chamadas probabilidades de transição, são fundamentais para definir a cadeia de Markov. Com efeito, podemos verificar que a cadeia de Markov está completamente determinada pelas probabilidades de transição e pela distribuição do estado inicial  $X_0$ . Isto é, se  $P(X_0 = i) = \alpha_i$  podemos calcular

$$P(X_0 = i_0, X_1 = i_1, \dots, X_n = i_n) \tag{1}$$

em função das probabilidades  $p(i, j)$  e dos valores  $\alpha_i$ . Para isto notemos que, por definição da probabilidade condicional, (1) é igual a

$$P(X_0 = i_0, X_1 = i_1, \dots, X_{n-1} = i_{n-1})P(X_n = i_n | X_0 = i_0, X_1 = i_1, \dots, X_{n-1} = i_{n-1}) \tag{2}$$

e da propriedade de Markov temos que

$$P(X_n = i_n | X_0 = i_0, X_1 = i_1, \dots, X_{n-1} = i_{n-1}) = P(X_n = i_n | X_{n-1} = i_{n-1}) = p(i_{n-1}, i_n). \tag{3}$$

Substituindo (3) em (2) temos que

$$P(X_0 = i_0, X_1 = i_1, \dots, X_n = i_n) = P(X_0 = i_0, X_1 = i_1, \dots, X_{n-1} = i_{n-1})p(i_{n-1}, i_n)$$

e repetindo esse argumento  $n - 1$  vezes concluímos que

$$P(X_0 = i_0, X_1 = i_1, \dots, X_n = i_n) = \alpha_{i_0} p(i_0, i_1) \dots p(i_{n-1}, i_n)$$

mostrando assim que a cadeia de Markov está definida por estas quantidades.

*Exercício 1.1.* Considere o passeio aleatório em  $\mathbb{Z}$  e escreva as probabilidades de transição  $p(i, j)$  para todo  $i, j \in \mathbb{Z}$ .

As definições anteriores podem ser estendidas de maneira natural para o caso contínuo. Isto é, podemos definir um processo estocástico a tempo contínuo como sendo uma família de variáveis aleatórias  $(X_t)_{t \geq 0}$ , com valores em  $\mathbb{Z}$ , e tais que  $t \in [0, \infty)$ . De maneira análoga, dizemos que um processo estocástico a tempo contínuo é uma cadeia de Markov a tempo contínuo se para todo  $s, t \geq 0$ , e valores  $i, j, i_r \in \mathbb{Z}$ , com  $0 \leq r < s$ , temos que

$$P(X_{t+s} = j | X_s = i, X_r = i_r, 0 \leq r < s) = P(X_{t+s} = j | X_s = i).$$

**Exemplo 1.5.** *Passeio aleatório a tempo contínuo em  $\mathbb{Z}$ .* Considere o fenômeno descrito no Exemplo 1.2 com a diferença que a partícula espera um tempo exponencialmente distribuído antes de pular. Uma vez que isto acontece, ela pula para algum dos sítios vizinhos de acordo às probabilidades  $p$  ou  $1-p$ . Se  $X_t$  representa a posição da partícula no instante de tempo  $t \in [0, \infty)$  então o processo estocástico a tempo contínuo  $(X_t)_{t \geq 0}$  é conhecido como passeio aleatório a tempo contínuo. Em particular, é uma cadeia de Markov a tempo contínuo.  $\square$

## 2 Processos de Bienaymé-Galton-Watson

### 2.1 Um pouco de história

A modo de motivação vamos mencionar alguns fatos sobre os começos da teoria de processos de ramificação. A seguinte descrição está baseada no artigo de Kendall (1966) [9] e na resenha histórica apresentada por Grinstead e Snell (1997) [4]. Outros detalhes dos começos e da evolução histórica desta teoria podem ser encontrados no livro de Harris (1963) [5].

Parece ser que o nome “processo de ramificação” foi introduzido por Kolmogorov e Dmitriev em 1947<sup>1</sup> para descrever processos estocásticos que aparecem na modelagem matemática de populações. No entanto, o conceito apareceu muito antes e foi pensado independentemente por Bienaymé, e por Galton e Watson.

Até algumas décadas atrás pensava-se que a ideia de processo de ramificação apareceu como resposta ao seguinte problema proposto por Francis Galton na *Educational Times* em 1873.

*Problem 4001: A large nation, of whom we will only concern ourselves with the adult males,  $N$  in number, and who each bear separate surnames, colonise a district. Their law of population is such that, in each generation,  $a_0$  per cent of the adult males have no male children who reach adult life;  $a_1$  have one such male child;  $a_2$  have two; and so on up to  $a_5$  who have five.*

*Find (1) what proportion of the surnames will have become extinct after  $r$  generations; and (2) how many instances there will be of the same surname being held by  $m$  persons.*

Em outras palavras, Galton propõe estudar a sobrevivência do sobrenome de uma família quando esse é passado de pai para filho e quando é assumido que cada homem têm  $i$  filhos com probabilidade  $a_i$ ,  $i = 0, 1, \dots, 5$ .

Henry William Watson encontrou, incorretamente, uma solução para esse problema usando funções geradoras. Watson concluiu que a probabilidade de extinção do sobrenome é sempre 1. Mas por sorte para a teoria, o erro que ele cometeu foi algébrico e seus argumentos de funções geradoras são os usados até hoje para encontrar a solução correta ao problema proposto por Galton e muitas das suas variantes.

Como esta era a história conhecida até anos atrás, muitos livros chamam esses processos de processos de ramificação de Galton-Watson. Na resenha de Grinstead e Snell os autores mencionam a descoberta, de Heyde e Seneta (1977)<sup>2</sup>, de uma comunicação de Bienaymé (1845) que antecipa-se a Galton e Watson. Em dita comunicação, Bienaymé mostra que conhece a solução correta do problema de Galton. Uma tradução ao inglês do trabalho original de Bienaymé, dada por Heyde e Seneta é a seguinte:

---

<sup>1</sup>Kolmogorov, A. N. e Dmitriev, N. A. *Branching stochastic processes*, Doklady Akad. Nauk U.S.S.R., 56 (1947), 5-8.

<sup>2</sup>Heyde, C. C. e Seneta, E. I. J. *Bienaymé: Statistical Theory Anticipated*, New York: Springer-Verlag (1977).

*If . . . the mean of the number of male children who replace the number of males of the preceding generation were less than unity, it would be easily realized that families are dying out due to the disappearance of the members of which they are composed. However, the analysis shows further that when this mean is equal to unity families tend to disappear, although less rapidly . . .*

*The analysis also shows clearly that if the mean ratio is greater than unity, the probability of the extinction of families with the passing of time no longer reduces to certainty. It only approaches a finite limit, which is fairly simple to calculate and which has the singular characteristic of being given by one of the roots of the equation (in which the number of generations is made infinite) which is not relevant to the question when the mean ratio is less than unity.*

Esta é a ideia central do Teorema 2.8 que garante condições sobre as quais a descendência dos homens de uma família se extingue ou não, em geral, dependendo da média de filhos de um determinado homem. No entanto, Bienaymé não apresenta os argumentos que o levaram para estas conclusões. Na sua comunicação ele manifesta interesse em publicar esses resultados mas até hoje esta suposta publicação nunca foi encontrada.

Esse é o motivo pelo qual recentes referências chamam aos processos de ramificação de processos de ramificação de Bienaymé-Galton-Watson.

## 2.2 Definição

De modo geral, podemos pensar um processo de ramificação de Bienaymé-Galton-Watson (BGW) da seguinte maneira. Suponha que temos partículas que dão nascimento a novas partículas de acordo a uma variável aleatória discreta  $X$  com valores no conjunto  $\{0, 1, 2, \dots\}$  e função de distribuição de probabilidades dada por

$$P(X = k) = p_k \tag{4}$$

para  $k = 0, 1, 2, \dots$ . Inicialmente suponha que, digamos no tempo  $n = 0$ , temos uma única partícula. Esta partícula dá nascimento, no tempo  $n = 1$ , a novas partículas segundo a variável aleatória  $X$ . Dizemos que estas novas partículas são descendentes diretas da anterior. Isto é, ela dá nascimento a  $k$  descendentes diretas com probabilidade  $p_k$ . Em geral, se temos um certo número de partículas no tempo  $n$ , cada uma delas dá nascimento a novas partículas no tempo  $n + 1$ . Em cada caso, isto é feito de acordo a uma variável aleatória independente e identicamente distribuída (i.i.d.) à variável aleatória  $X$ . Na Figura 2 ilustramos uma possível realização deste processo.

Dizemos que a partícula do tempo  $n = 0$  constitui a geração 0 e as partículas que nascem no tempo  $n$  constituem a  $n$ -ésima geração do processo,  $n \geq 1$ . Denotamos por  $Z_n$  a variável aleatória que conta o número de partículas da  $n$ -ésima geração. Notemos que  $(Z_n)_{n \geq 0}$  é uma cadeia de Markov. Com estas ideias em mente podemos formalizar a definição de processo de ramificação de BGW.



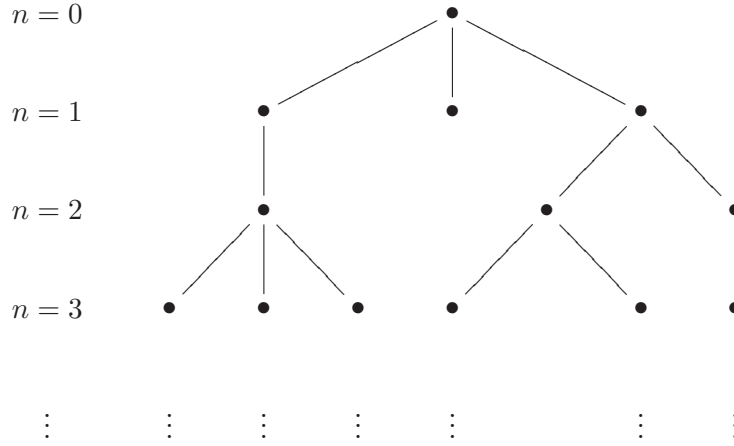


Figura 2: Possível realização de um processo de ramificação de BGW.

**Definição 2.1.** Seja  $X$  uma variável aleatória discreta com distribuição de probabilidades dada por (4). Chamamos processo de ramificação de Bienaymé-Galton-Watson, ou simplesmente processo de ramificação, à cadeia de Markov  $(Z_n)_{n \geq 0}$  com valores no conjunto  $\{0, 1, 2, \dots\}$  e probabilidades de transição dadas por:

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

onde  $X_1, \dots, X_i$  são i.i.d. à variável aleatória  $X$ .

Temos definido o processo de ramificação de BGW em função das suas probabilidades de transição. Da Observação 1.4 sabemos que ele está completamente definido a partir destas probabilidades e da distribuição do seu estado inicial  $Z_0$ . Para simplificar a exposição nós assumimos que  $Z_0 \equiv 1$ .

Notemos que se não há partículas na  $n$ -ésima geração então não haverá partículas nas gerações seguintes. Daí que  $p(0, j) = 0$  para  $j > 0$  e que  $p(0, 0) = 1$ . Nesse sentido, dizemos que o estado 0 é um estado absorvente e representa o evento que chamamos de extinção do processo.

**Definição 2.2.** Seja  $(Z_n)_{n \geq 0}$  um processo de ramificação. Chamamos de extinção do processo ao evento

$$\mathcal{E} = \bigcup_{n \geq 1} \{Z_n = 0\} \quad (5)$$

e denotamos a probabilidade de extinção como  $q := P(\mathcal{E})$ .

**Observação 2.3.** A extinção do processo pode ser interpretada como a existência de uma geração a partir da qual não nascem mais partículas. Esse evento pode ser escrito da seguinte maneira

$$\bigcup_{n \geq 1} \bigcap_{k \geq n} \{Z_k = 0\}. \quad (6)$$

*Exercício 2.1.* Mostre que os eventos (5) e (6) são equivalentes.

O evento de extinção é de muito interesse quando estudamos ou aplicamos processos de ramificação. No problema de estudar a descendência do sobrenome de uma família, a não extinção do processo é equivalente à sobrevivência do sobrenome. O ponto central destas notas é descrever o teorema fundamental dos processos de ramificação que garante condições sobre as quais teremos ou não a extinção do processo. Antes disto, vamos ver algumas propriedades da variável aleatória  $Z_n$ .

**Proposição 2.4.** *Sejam  $m = E(X)$  e  $v = Var(X)$ . Então  $E(Z_n) = m^n$  e*

$$Var(Z_n) = \begin{cases} v^2 m^{n-1} \left( \frac{m^n - 1}{m - 1} \right), & \text{se } m \neq 1 \\ nv^2, & \text{se } m = 1 \end{cases}$$

para todo  $n \geq 1$ .

**Prova.** Vamos mostrar que  $E(Z_n) = m^n$  usando a seguinte propriedade das esperanças condicionais:

$$E(Z_n) = E(E(Z_n | Z_{n-1})).$$

Em particular, como  $Z_{n-1}$  é uma variável aleatória discreta, a equação anterior diz que

$$E(Z_n) = \sum_{i=1}^{\infty} E(Z_n | Z_{n-1} = i) P(Z_{n-1} = i),$$

mas  $Z_n = \sum_{j=1}^{Z_{n-1}} X_j$ , onde  $X_j$  é i.i.d. à variável aleatória  $X$ , para todo  $j$ . Então

$$E(Z_n) = \sum_{i=1}^{\infty} E \left( \sum_{j=1}^{Z_{n-1}} X_j | Z_{n-1} = i \right) P(Z_{n-1} = i).$$

Por outro lado, dado que as variáveis aleatórias  $X_i$  e  $Z_{n-1}$  são independentes temos que

$$E(Z_n) = \sum_{i=1}^{\infty} E \left( \sum_{j=1}^i X_j \right) P(Z_{n-1} = i) = m \sum_{i=1}^{\infty} i P(Z_{n-1} = i)$$

Logo,

$$\begin{aligned} E(Z_n) &= mE(Z_{n-1}) \\ &= m^2E(Z_{n-2}) \\ &\vdots \\ &= m^{n-1}E(Z_1). \end{aligned}$$

e como  $E(Z_1) = E(X) = m$  concluímos que  $E(Z_n) = m^n$ . A prova da expressão para  $Var(Z_n)$  resulta de maneira análoga e é deixada como exercício. ■

*Exercício 2.2.* Prove a expressão para  $Var(Z_n)$  dada na Proposição 2.4. Dica: usar a fórmula da variância condicional  $Var(Z_n) = E(Var(Z_n|Z_{n-1})) + Var(E(Z_n|Z_{n-1}))$ .

### 2.3 Funções geradoras de probabilidade

Para analisar a extinção ou não de um processo de ramificação vamos usar a função geradora de probabilidade (f.g.p.) da variável aleatória  $X$ . Isto é, consideramos a função  $\phi(t)$  definida por

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

Observemos que o conhecimento da função  $\phi(t)$  nós permite obter os valores  $p_k$  a partir da fórmula de Taylor. Isto é, como  $p_k$  é o coeficiente de  $t^k$  em  $\phi(t)$  então

$$p_k = \frac{((d^{(k)}\phi(t)/dt)|_0)}{k!}. \quad (7)$$

Daí o nome de função geradora de probabilidades.

No seguinte resultado vemos uma interessante relação entre as iterações desta função  $\phi$  e a f.g.p. do tamanho da  $n$ -ésima geração do processo. Na sequência denotamos  $(\phi \circ \phi)(t)$  para representar a função  $\phi(\phi(t))$  resultante de iterar duas vezes a função  $\phi$ .

**Proposição 2.5.** *Sejam  $\phi(t)$  e  $\phi_n(t)$  as f.g.p. das variáveis aleatórias  $X$  e  $Z_n$ ,  $n \geq 1$ , respectivamente. Então*

$$\phi_n(t) = \phi^n(t) \quad (8)$$

onde

$$\phi^n(t) = \underbrace{(\phi \circ \phi \circ \dots \circ \phi)}_{n \text{ vezes}}(t).$$

**Prova.** Vamos provar (8) por indução sob  $n$ . O caso  $n = 1$  é claro pois

$$\phi_1(t) = E(t^{Z_1}) = E(t^X) = \phi(t).$$

Supomos que (8) é válida para  $n - 1$  e verificamos sua validade para  $n$ . Da definição de f.g.p.

$$\phi_n(t) = \sum_{i=1}^{\infty} P(Z_n = i)t^i. \quad (9)$$

Por outro lado,

$$\begin{aligned} P(Z_n = i) &= \sum_{j=1}^{\infty} P\left(\sum_{k=1}^{Z_{n-1}} X_k = i \mid Z_{n-1} = j\right) P(Z_{n-1} = j) \\ &= \sum_{j=1}^{\infty} P\left(\sum_{k=1}^j X_k = i \mid Z_{n-1} = j\right) P(Z_{n-1} = j) \\ &= \sum_{j=1}^{\infty} P\left(\sum_{k=1}^j X_k = i\right) P(Z_{n-1} = j) \end{aligned} \quad (10)$$

onde as sucessivas igualdades foram obtidas usando a definição de  $Z_n$  a partir de  $Z_{n-1}$  e a independência entre as variáveis aleatórias  $X_i$  e  $Z_{n-1}$ . De (9) e (10) temos que

$$\phi_n(t) = \sum_{j=1}^{\infty} \left( \sum_{i=1}^{\infty} P\left(\sum_{k=1}^j X_k = i\right) t^i \right) P(Z_{n-1} = j)$$

mas

$$\sum_{i=1}^{\infty} P\left(\sum_{k=1}^j X_k = i\right) t^i = E(t^{\sum_{k=1}^j X_k}) = (\phi(t))^j.$$

Então

$$\phi_n(t) = \sum_{j=1}^{\infty} (\phi(t))^j P(Z_{n-1} = j) = \phi_{n-1}((\phi(t))) = (\phi_{n-1} \circ \phi)(t)$$

e da hipótese indutiva temos que

$$\phi_{n-1}(t) = \phi^{n-1}(t) = \underbrace{(\phi \circ \phi \circ \dots \circ \phi)}_{n-1 \text{ vezes}}(t).$$

Isto completa nossa prova. ■

No que segue enumeramos algumas propriedades chaves para nossa análise da probabilidade de extinção do processo de ramificação. A prova, que é deixada como exercício, usa argumentos básicos de análise e a definição da f.g.p.

**Proposição 2.6.** *Se  $p_0 + p_1 < 1$ , então  $\phi$  satisfaz as seguintes propriedades:*

- i.  $\phi$  é estritamente convexa e crescente em  $[0, 1]$ ;*
- ii.  $\phi(0) = p_0$  e  $\phi(1) = 1$ ;*
- iii. se  $\phi'(1) \leq 1$  então  $\phi(t) > t$  para  $t \in [0, 1)$ ;*
- iv. se  $\phi'(1) > 1$  então  $\phi(t) = t$  tem uma única raiz em  $[0, 1)$ .*

*Exercício 2.3.* Prove os itens *i* até *iv* da Proposição 2.6.

**Observação 2.7.** Em particular, temos que  $\phi'(1) = m$ .

## 2.4 Probabilidade de extinção

Estamos em condições de provar o resultado fundamental dos processos de ramificação.

**Teorema 2.8.** *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,*

*i. se  $m \leq 1$  então  $q = 1$ ;*

*ii. se  $m > 1$  então  $q < 1$ .*

**Prova.** Vamos dividir a prova em duas partes. Primeiro vamos provar que a probabilidade de extinção  $q$  satisfaz a equação  $q = \phi(q)$ . Para isto vamos usar propriedades do processo. Na segunda parte vamos verificar que  $q$  é a menor das raízes da equação em  $[0, 1]$ . Isto será feito analisando o comportamento da f.g.p. da variável aleatória  $X$ .

**Primeira parte:** Vamos provar que  $q = \phi(q)$ . Definimos, para todo  $n \geq 0$ ,

$$q_n := P(Z_n = 0)$$

e notamos que

$$0 = q_0 \leq q_1 \leq q_2 \leq \dots \leq q_n \leq \dots \leq 1.$$

Com efeito, notemos que  $P(Z_0 = 0) = 0$  e que  $\{Z_n = 0\} \subset \{Z_{n+1} = 0\}$  para todo  $n \geq 1$ . Logo  $\{Z_n = 0\}$  é uma sequência crescente de eventos<sup>3</sup> e portanto

$$\lim_{n \rightarrow \infty} q_n = q. \tag{11}$$

Por outro lado, temos que

$$\begin{aligned} q_n &= \sum_{i=1}^{\infty} P(Z_n = 0 | Z_1 = i) P(Z_1 = i), && \text{condicionando sobre a primeira geração} \\ &= \sum_{i=1}^{\infty} (P(Z_{n-1} = 0))^i p_i, && \text{porque? (exercício)} \\ &= \sum_{i=1}^{\infty} (q_{n-1})^i p_i, && \text{definição de } q_{n-1} \\ &= \phi(q_{n-1}), && \text{definição de } \phi. \end{aligned}$$

Isto é,

$$q_n = \phi(q_{n-1}). \tag{12}$$

---

<sup>3</sup>Dizemos que uma sequência de eventos  $A_1, A_2, \dots$  é crescente se  $A_n \subset A_{n+1}$  para todo  $n \geq 1$ . Neste caso  $\lim_{n \rightarrow \infty} P(A_n) = P\left(\bigcup_{i=1}^{\infty} A_i\right)$ .

Como  $\phi$  é uma função contínua, resulta de (11) e (12) que  $q = \phi(q)$ .

**Segunda parte:** Vamos analisar as soluções de  $t = \phi(t)$  usando propriedades de  $\phi$ . Em particular, vamos analisar estas soluções a partir do gráfico da função  $\phi(t)$  (Figura 3). Da Proposição 2.6 temos três possíveis comportamentos da função  $\phi$ , dependendo de que (a)  $\phi'(1) > 1$ , (b)  $\phi'(1) = 1$  ou  $\phi'(1) < 1$ .

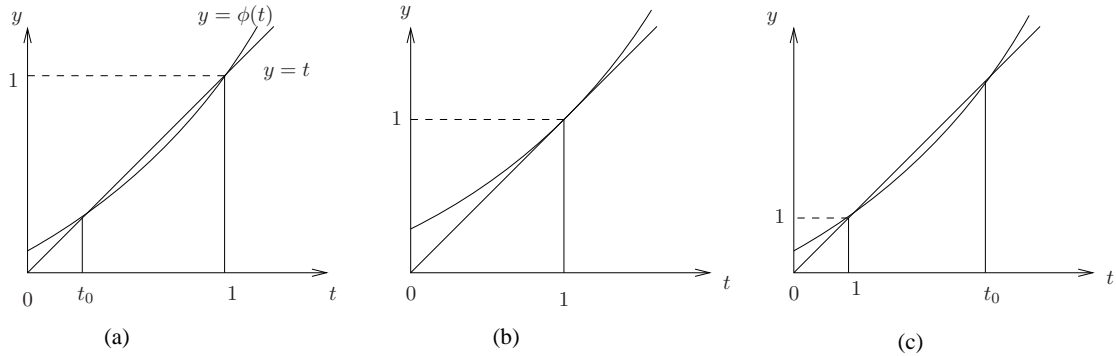


Figura 3: Comportamento de  $\phi(t)$  quando (a)  $\phi'(1) > 1$ , (b)  $\phi'(1) = 1$  e (c)  $\phi'(1) < 1$ .

Notemos que em todos os casos existem no máximo dois valores,  $t_0$  e 1, tais que  $t = \phi(t)$ . Como  $q$  é raiz desta equação temos que:

- de (a) ou  $q < 1$  ou  $q = 1$ ,
- de (b)  $q = 1$ , e
- de (c)  $q = 1$  ou  $q > 1$ .

Como  $\phi'(1) = m$  concluímos de (b) e (c) que  $q$  é a menor raiz não negativa da equação  $t = \phi(t)$  e que se  $m \leq 1$  então  $q = 1$  (não pode ser  $q > 1$  pois  $q$  é uma probabilidade). Com isto provamos *ii*.

Só falta analisar o caso (a). Vamos usar (12) para construir os valores  $q_i$  a partir do gráfico da função  $\phi$ . Notemos que

$$\begin{aligned} q_0 &= 0 \\ q_1 &= \phi(q_0) = \phi(0) = p_0 \\ q_2 &= \phi(q_1) = \phi(p_0) \\ &\vdots \end{aligned}$$

Na Figura 4 temos a construção geométrica desses valores a partir do gráfico de  $\phi$ . Daí podemos concluir que os valores  $q_i$  convergem à primeira interseção de  $y = \phi(t)$  com  $y = t$ . Disto e (11), temos que deve ser  $q = t_0$  e portanto  $q$  é a menor raiz de  $t = \phi(t)$ . Em particular  $q < 1$  e provamos *i*. ■

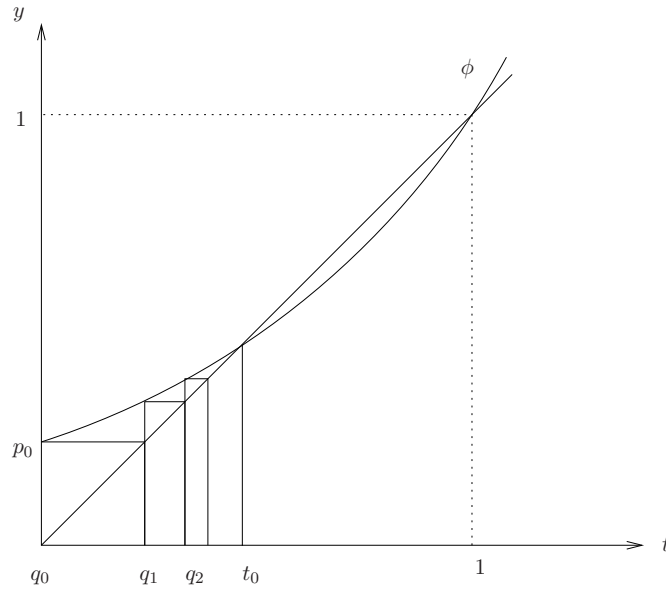


Figura 4: Construção geométrica dos valores  $q_i$  e  $q$ .

*Exercício 2.4.* Mostre que  $P(Z_n = 0 | Z_1 = i) = P(Z_{n-1} = 0)^i$ .

*Exercício 2.5.* Analise o comportamento de  $q$  quando  $p_0 + p_1 = 1$ .

**Exemplo 2.9.** Suponha que  $p_0 = 1/4$ ,  $p_1 = 1/4$  e  $p_2 = 1/2$ . Neste caso  $m > 1$  e do Teorema 2.8 concluímos que  $q < 1$ . Em particular,

$$\phi(t) = \frac{1}{4} + \frac{1}{4}t + \frac{1}{2}t^2$$

e  $q$  é a menor raiz da equação  $t = \phi(t)$ . Como  $t = \phi(t)$  se, e somente se,

$$2t^2 - 3t + 1 = 0$$

temos que  $q = 1/2$ . □

**Exemplo 2.10.** Os processos de ramificação de divisão binária são de muita importância na biologia. Esse tipo de processos são tais que

$$p_0 = 1 - p \quad \text{e} \quad p_2 = p,$$

para algum  $0 \leq p \leq 1$ . Notemos que neste caso

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

e portanto  $m = \phi'(1) = 2p$ . Do Teorema 2.8 temos que  $q = 1$  se  $p \leq 1/2$  e  $q = (1 - p)/p$  se  $p > 1/2$ .  $\square$

O seguinte exemplo é dado no livro de Grinstead e Snell [4] e refere-se a dados reais compilados e analisados por Keyfitz em 1977<sup>4</sup>.

**Exemplo 2.11.** Voltando à motivação inicial dos processos de ramificação vamos analisar a continuação da linhagem familiar de sexo feminino entre mulheres japonesas. Keyfitz estimou que a distribuição do número de filhas de mulheres japonesas, de idades entre 45 e 49 anos em 1960, é dada pela seguinte tabela:

$p_0$	=	0.2092
$p_1$	=	0.2584
$p_2$	=	0.2360
$p_3$	=	0.1593
$p_4$	=	0.0828
$p_5$	=	0.0357
$p_6$	=	0.0133
$p_7$	=	0.0042
$p_8$	=	0.0011
$p_9$	=	0.0002
$p_{10}$	=	0.0000

Figura 5: Distribuição do número de filhas.

Notemos que o número esperado de filhas em uma família é dado por 1.837 e portanto a probabilidade de extinção  $q < 1$ . Do Teorema 2.8 podemos estimar que na verdade  $q \approx 0.324$ .  $\square$

**Exemplo 2.12.** Suponha que  $p_k = bp^{k-1}$ , para  $k = 1, 2, \dots$ , e que

$$p_0 = 1 - \sum_{i=1}^{\infty} p_i = \frac{1 - b - p}{1 - p}.$$

Nesse caso, temos que a f.g.p. da variável aleatória  $X$  é dada por

$$\phi(t) = \frac{1 - b - p}{1 - p} + \sum_{i=1}^{\infty} bp^{i-1}t^i = \frac{1 - b - p}{1 - p} + \frac{b}{p} \sum_{i=1}^{\infty} (pt)^i = 1 - \frac{b}{1 - p} + \frac{bt}{1 - pt}.$$

Em particular,

$$m = \frac{b}{(1 - p)^2}$$

---

<sup>4</sup>Keyfitz, N. *Introduction to the Mathematics of Population*, rev. ed. (Reading, PA: Addison Wesley), (1977).



e podemos aplicar o Teorema 2.8 para analisar a probabilidade de extinção  $q$ . Se  $m > 1$ , temos que  $q < 1$  e é a menor raiz da equação

$$t = 1 - \frac{b}{1-p} + \frac{bt}{1-pt}.$$

Isto é,

$$q = \frac{1-b-p}{p(1-p)}.$$

Por outro lado, podemos aplicar a Proposição 8 para obter a distribuição de  $Z_n$  (ver I.4 de [2]). Quando  $m \neq 1$  temos que

$$\phi_n(t) = 1 - m^n \left( \frac{1-d}{m^n-d} + \frac{m^n \left( \frac{1-d}{m^n-d} \right)^2 t}{1 - \left( \frac{m^n-1}{m^n-d} \right) t} \right), \quad (13)$$

e quando  $m = 1$

$$\phi_n(t) = \frac{np - (np + p - 1)t}{1 - p + np - npt}. \quad (14)$$

Logo, podemos calcular as probabilidades  $P(Z_n = k)$  a partir da f.g.p. e de (7). Se  $m \neq 1$  temos que

$$P(Z_n = 0) = 1 - m^n \left( \frac{1-d}{m^n-d} \right)$$

e para  $i \geq 1$

$$P(Z_n = i) = m^n \left( \frac{1-d}{m^n-d} \right)^2 \left( \frac{m^n-1}{m^n-d} \right)^{i-1}.$$

□

*Exercício 2.6.* Considere a distribuição do Exemplo 2.12 e obtenha, para  $m = 1$ , as probabilidades  $P(Z_n = i)$  a partir de (14) e (7).

**Observação 2.13.** Processos de ramificação com o tipo de distribuição apresentada no Exemplo 2.12 podem ser usados para estudar a evolução da descendência de uma família. Um exemplo muito citado ([3, 5, 8]) é dado por Lotka (1939)<sup>5</sup> quem mostrou que a distribuição  $p_0 = 0.4825$  e  $p_k = (0.2126)(0.5893)^{k-1}$ , para  $k \geq 1$ , é apropriada para descrever a descendência direta de homens americanos (os valores numéricos foram baseados em um censo de 1920). Lotka aplicou o Teorema 2.8 para determinar que a probabilidade de extinção neste caso é  $q = 0.819$ .

Outro exemplo interessante surge de analisar os dados do Exemplo 2.11. Neste caso podemos mostrar que a distribuição  $p_k = (0.3666)(0.5533)^{k-1}$  pode ser apropriada para descrever os dados obtidos (ver Exemplo 10.12, pág. 385, [4]).

---

<sup>5</sup>Lotka, A. J. *Theorie analytique des associations biologiques*, Actualites scientifiques et industrielles 780, Paris, Hermann (1939), 123-136.

### 3 Aplicações e generalizações

Com o intuito de ilustrar algumas aplicações dos processos de ramificação apresentamos exemplos bem conhecidos na literatura.

#### 3.1 Outros exemplos e aplicações

**Reações nucleares em cadeia.** Segundo Feller [3] a seguinte descrição é devida a Schroedinger (1945)<sup>6</sup>. Suponha que as partículas em nosso processo representam nêutrons que estão sujeitos a possíveis colisões com outras partículas. Suponha que em cada colisão um nêutron se divide em  $k$  nêutrons. Desta maneira, se denotamos por  $p$  a probabilidade de colisão temos que o número de nêutrons cresce como um processo de ramificação onde

$$p_0 = 1 - p, \quad p_k = p \quad \text{e} \quad p_i = 0,$$

para  $i \neq 0, k$ . No pior dos casos, nenhuma colisão acontece. No melhor dos casos, teremos  $k$  nêutrons na primeira geração,  $k^2$  na segunda, e assim sucessivamente. Logo, se  $p$  é suficientemente próximo de 1 teremos que o número de nêutrons cresce muito rapidamente. Neste sentido, a sobrevivência do processo pode ser pensada como “explosão”.

**Filas de clientes.** Uma interessante aplicação de processos de ramificação em teoria das filas é motivada por Kendall (1951)<sup>7</sup> (ver Feller [3]). A seguinte descrição é baseada no livro de Schinazi (1999) [8]. Suponha que queremos analisar o comportamento de uma fila formada por clientes que chegam a um determinado servidor. Suponha que em cada instante de tempo  $n$  existem duas possibilidades:

- ou chega um cliente, com probabilidade  $p$ ,
- ou não chega nenhum cliente, com probabilidade  $1 - p$ .

Por outro lado, quando um cliente chega ao servidor temos que:

- se o servidor está livre, o serviço para esse cliente começa imediatamente,
- se o servidor está ocupado, o cliente se junta a uma fila.

Supondo que todos os tempos de serviço são independentes e têm a mesma distribuição, podemos usar um processo de ramificação para analisar esta fila. Para isto, assumamos que a geração 0 está formada por um único cliente. Digamos o primeiro em chegar. A primeira geração consiste de todos os clientes que chegam durante o tempo de serviço do cliente da geração 0. Em geral, a  $n$ -ésima geração consiste dos clientes que chegam durante tempos

<sup>6</sup>Schroedinger, E. *Probability problems in nuclear chemistry*, Proceedings of the Royal Irish Academy, vol. 51, sect. A, No 1 (1945).

<sup>7</sup>Kendall, D. G. *Some problems in the theory of queues*, J. Roy. Statist. Soc. (Séries B). vol. 13 (1951), 151-173.

de serviço de algum cliente da geração anterior  $n - 1$ . Desta forma temos um processo de ramificação  $(Z_n)_{n \geq 1}$  que pode ser construído a partir da variável aleatória

$$X = X_1 + X_2 + \dots + X_T$$

onde  $T$  é o tempo de serviço de um cliente e  $X_i$  são variáveis aleatórias i.i.d. com distribuição Bernoulli de parâmetro  $p$ . Assumindo que as variáveis aleatórias  $X_i$  ( $i = 1, 2, \dots$ ) e  $T$  são independentes, podemos mostrar que  $E(X) = pE(T)$ . Logo, temos que

- i. se  $pE(T) > 1$ , o processo de ramificação sobrevive com probabilidade positiva. Em termos da fila isto quer dizer que sempre teremos fila com probabilidade positiva.
- ii. se  $pE(T) \leq 1$  o processo de ramificação se extingue com probabilidade 1, isto é, com probabilidade 1 a fila ficará vazia infinitas vezes.

**Propagação de infecções e rumores.** Os processos de ramificação tem sido usados com êxito, desde há mais de 50 anos, como uma aproximação da evolução dos primeiros estágios de uma epidemia. Um modelo teórico da literatura de modelagem matemática de epidemias é o modelo SIR. Esse modelo foi formulado com o intuito de descrever a difusão de uma infecção em uma população. Na versão estocástica básica podemos assumir uma população formada por  $N$  indivíduos. O modelo SIR é chamado desta forma para ressaltar as classes de indivíduos em que a população está subdividida: susceptíveis, infectados e removidos. Podemos pensar que em cada encontro de um infectado com um susceptível o susceptível vira infectado e que depois de um tempo aleatório um infectado vira removido. Nos primeiros estágios da difusão de uma infecção, para uma população suficientemente grande, é intuitivamente claro que com alta probabilidade um infectado encontra-se com um susceptível antes que outro infectado ou removido. Logo, o crescimento do número de indivíduos infectados segue um comportamento parecido ao de um processo de ramificação. Neste sentido se cada partícula representa um infectado e os descendentes diretos desta partícula são aqueles indivíduos que ele/ela infecta, temos um processo de ramificação para o qual a sobrevivência é equivalente à epidemia na população. Para uma formalização desta aproximação, precisamos da noção de acoplamento de processos estocásticos. Dita abordagem pode ser encontrada no livro de Andersson e Britton (2000) [1].

As mesmas ideias podem ser usadas para estudar a propagação de um rumor em uma população. Nesta direção assume-se que uma população de tamanho  $N$  é subdividida em ignorantes, informantes e contidos e o rumor se propaga através da população por contato direto entre informantes e outros indivíduos. Neste sentido temos as seguintes possibilidades:

- se um informante encontra um ignorante o ignorante vira informante,
- se o informante encontra outro informante ou um contido, esse informante inicial vira contido,

- um informante esquece o rumor depois de um tempo aleatório.

A segunda possibilidade representa a perda de interesse de um informante em continuar contando o rumor quando ele percebe que o rumor já é conhecido. Novamente, se uma partícula representa um informante e seus descendentes diretos são aqueles indivíduos que ele/ela informa antes de esquecer do rumor, temos um comportamento parecido ao processo de ramificação.

## 3.2 Generalizações

No processo de ramificação da Definição 2.1 são assumidas as seguintes características. O processo evolui a tempo discreto, isto é, entre uma geração e a seguinte ocorre um passo unitário de tempo. Por outro lado supomos que todas as partículas são iguais e que a distribuição do número de descendentes diretos de cada partícula é sempre a mesma.

No que segue, vamos mencionar possíveis generalizações destas características. Na prática, estas variantes fornecem uma descrição mais aproximada para diversos fenômenos da vida real.

### 3.2.1 Processos de ramificação em meios variáveis

Um processo de ramificação em meio variável é um processo para o qual a distribuição do número de descendentes diretos de uma partícula depende da geração à qual a partícula pertence. Formalmente, podemos definir esses processos da seguinte maneira.

**Definição 3.1.** Seja  $X_1, X_2, X_3, \dots$  uma sequência de variáveis aleatórias discretas independentes. Chamamos processo de ramificação em meio variável à cadeia de Markov  $(Z_n)_{n \geq 0}$  com valores no conjunto  $\{0, 1, 2, \dots\}$  e probabilidades de transição dadas por:

$$p(i, j) = \begin{cases} P(Z_{n+1} = j | Z_n = i) = P\left(\sum_{r=1}^i X_{n,r} = j\right), & \text{para } i \geq 1 \text{ e } j \geq 0 \\ 0, & \text{para } i = 0 \text{ e } j > 0 \\ 1, & \text{para } i = 0 \text{ e } j = 0 \end{cases}$$

onde  $X_{n,1}, \dots, X_{n,i}$  são i.i.d. à variável aleatória  $X_n$ .

**Observação 3.2.** Neste caso a f.g.p. de  $Z_n$  é dada por

$$\phi_n(t) = (\phi^1 \circ \phi^2 \circ \dots \circ \phi^{n-1})(t),$$

onde  $\phi^i(t)$  é a f.g.p. da variável aleatória  $X_i$ ,  $i \geq 1$ . Como nos processos de ramificação de BGW, as f.g.p. também têm um papel importante na procura de condições para a extinção ou não de um processo de ramificação em meio variável. Uma reunião interessante de resultados e exemplos pode ser encontrada no livro de Jagers [6].

### 3.2.2 Processos de ramificação multitypo

Outra variante natural do processo de ramificação é permitir diferentes tipos de partículas. Suponha que existem  $k$  tipos diferentes de partículas ( $k < \infty$ ) e que cada partícula de tipo  $r$  está associada a um vetor aleatório

$$\mathbf{X}_r = (X_r^1, X_r^2, \dots, X_r^k),$$

onde  $X_r^i$  é uma variável aleatória que representa o número de partículas de tipo  $i$ ,  $i = 1, 2, \dots, k$ , que nasce a partir de uma partícula de tipo  $r$ , para  $r = 1, 2, \dots, k$ . Supomos que

$$P(X_r^i = j) = p_r^i(j) \quad (15)$$

A partir desta variáveis aleatórias podemos construir o processo de ramificação multitypo.

**Definição 3.3.** Chamamos processo de ramificação multitypo à cadeia de Markov  $(\mathbf{Z}_n)_{n \geq 0}$  com valores em  $\mathbb{Z}^k$ , onde

$$\mathbf{Z}_n = (Z_n(1), Z_n(2), \dots, Z_n(k))$$

e  $Z_n(r)$  representa o número de partículas de tipo  $r$  na geração  $n$ ,  $r = 1, 2, \dots, k$ .

Vimos que no processo de ramificação de BGW o valor esperado  $m$  joga um papel importante na hora de analisar a probabilidade de extinção do processo. Neste caso, o interesse é na matriz

$$M = \begin{bmatrix} m_{11} & m_{12} & \dots & m_{1k} \\ m_{21} & m_{22} & \dots & m_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ m_{k1} & m_{k2} & \dots & m_{kk} \end{bmatrix}$$

onde  $m_{ri} = E(X_r^i)$ , para  $i, r = 1, 2, \dots, k$ . Se  $M$  é uma matriz estritamente positiva então tem um autovalor  $\rho$  máximo, positivo e simples. Esse valor  $\rho$  joga o mesmo papel que  $m$  na hora de procurar condições para a extinção ou não do processo de ramificação multitypo. Mais detalhes podem ser encontrados no livro de Athreya (1977) [2].

### 3.2.3 Processos de ramificação a tempo contínuo

A última variante que discutiremos está relacionada com o tempo transcorrido entre os nascimentos de uma partícula e suas partículas descendentes. Até agora assumimos tempo discreto, isto é, entre cada geração e a seguinte assumimos que transcorre uma unidade de tempo. Neste caso vamos supor que o processo evolui a tempo contínuo. Inicialmente, digamos no instante de tempo  $t = 0$  existe uma única partícula. Esta partícula viverá por um tempo  $T$ , onde  $T$  é uma variável aleatória contínua com função de distribuição  $G$ . Isto é

$$G(a) = P(T \leq a).$$

No final da sua vida, esta partícula dará nascimento (sendo substituída) a um número de partículas distribuído segundo uma variável aleatória  $X$  com distribuição dada por (4) (ver Figura 6). O processo continua desta maneira, sendo que cada partícula vive por um tempo i.i.d. à variável aleatória  $T$  e tem descendentes diretos segundo uma variável aleatória i.i.d. à variável aleatória  $X$ . Seja  $Z_t$  o número de partículas no instante de tempo  $t$ ,  $t \geq 0$ .

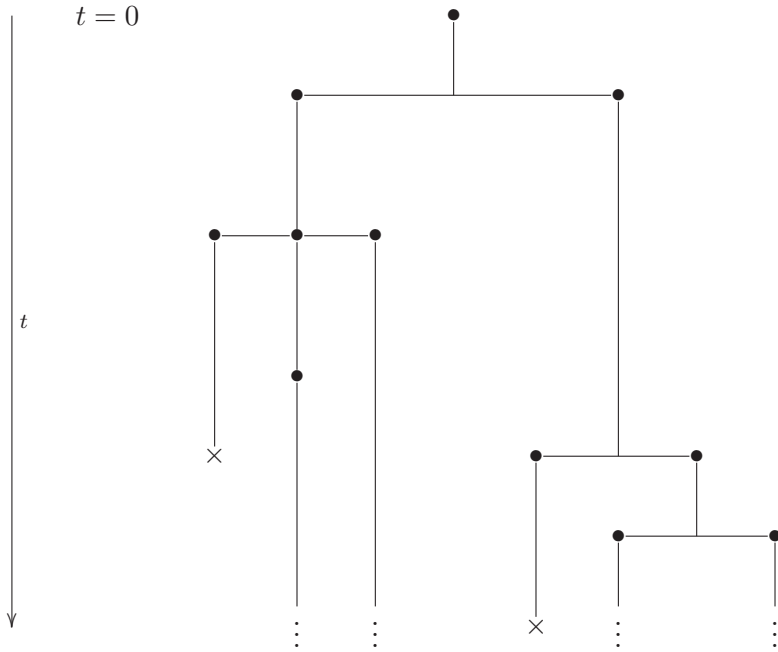


Figura 6: Possível realização de um processo de ramificação a tempo contínuo.

**Definição 3.4.** Chamamos processo de ramificação a tempo contínuo ao processo estocástico  $(Z_t)_{t \geq 0}$  com valores no conjunto  $\{0, 1, 2, \dots\}$ , construído a partir das variáveis aleatórias  $T$  e  $X$  definidas anteriormente.

**Observação 3.5.** Em geral, o processo  $(Z_t)_{t \geq 0}$  não é uma cadeia de Markov. No entanto, se a variável aleatória  $T$  tem distribuição exponencial então o processo é uma cadeia de Markov.

**Observação 3.6.** Novamente as funções geradoras jogam um papel importante no estudo da probabilidade de extinção do processo. Os detalhes neste caso requerem de maior conhecimento de processos estocásticos a tempo contínuo. Para uma análise completa da construção, definição e resultados destes processos ver o livro de Athreya (1977) [2].

**Exemplo 3.7.** A proliferação celular pode ser pensada como um processo de ramificação a tempo contínuo. Neste caso, uma célula vive durante um tempo aleatório chamado ciclo

celular, depois do qual divide-se em duas células. Muitas vezes a célula pode morrer antes desta divisão. Para descrever esta proliferação podemos assumir que a probabilidade de cada célula morrer antes da divisão é  $1 - p$ . Por outro lado, supomos que a distribuição da duração do ciclo celular é dada por  $G$  e dado que uma célula morre antes deste tempo, ela morre em um tempo com distribuição dada por  $F$ . Desta maneira, pensando células como partículas no processo de ramificação, temos que cada partícula vive por um tempo distribuído de acordo a

$$H = (1 - p)F + pG.$$

Depois deste tempo aleatório, cada partícula dá nascimento a 2 partículas com probabilidade  $p$  ou a nenhuma com probabilidade  $1 - p$ . Esse é um exemplo de processo de ramificação de divisão binária na biologia.  $\square$

## Referências

- [1] H. Andersson e T. Britton. Stochastic Epidemic Models and their Statistical Analysis. Springer-Verlag, New York, 2000.
- [2] K. B. Athreya e P. E. Ney. Branching Processes. Springer-Verlag, New York, 1972.
- [3] W. Feller. An Introduction to Probability Theory and its Applications, vol. I, Third edition, John Wiley & Sons, 1968.
- [4] C. M. Grinstead e J. L. Snell. Introduction to Probability. 2nd Revised Ed. AMS, 1997.
- [5] T. Harris. The Theory of Branching Processes. Springer-Verlag, Berlin, 1963.
- [6] P. Jagers. Branching processes with biological applications. John Wiley and Sons, 1975.
- [7] S. Karlin e H. M. Taylor. An Introduction To Stochastic Modeling. Third Edition, Academic Press, 1998.
- [8] R. B. Schinazi. Classical and Spatial Stochastic Processes. Birkhäuser, Boston, 1999.
- [9] Kendall, D. G. Branching processes since 1873. *Journal of the London Mathematical Society* **41**, 385-406, 1966.