

UNIVERSIDADE FEDERAL DE MINAS GERAIS
Instituto de Ciências Exatas
Programa de Pós-Graduação em Estatística

Carlos Henrique Santos

Um passeio aleatório com catástrofes

Belo Horizonte
2024

Carlos Henrique Santos

Um passeio aleatório com catástrofes

Versão Final

Dissertação apresentada ao Programa de Pós-Graduação em Estatística da Universidade Federal de Minas Gerais, como requisito parcial à obtenção do título de Mestre em Estatística.

Orientador: Roger William Camara Silva

Belo Horizonte
2024

Santos, Carlos Henrique.

S237p Um passeio aleatório com catástrofes [recurso eletrônico] /
Carlos Henrique Santos – 2024.
1 recurso online (70 f. il., color.): pdf.

Orientador: Roger William Câmara Silva.

Dissertação (Mestrado) - Universidade Federal de
Minas Gerais, Instituto de Ciências Exatas, Departamento de
Estatística.

Referências: f. 59-60.

1. Estatística – Teses. 2. Markov, Campos aleatórios de –
Teses. 3. Catástrofes (Matemática) – Teses. 3. Modelo de
população estável – Teses. I. Silva, Roger William Câmara.
II. Universidade Federal de Minas Gerais, Instituto de Ciências
Exatas, Departamento de Estatística. III. Título.

CDU 519.2(043)



UNIVERSIDADE FEDERAL DE MINAS GERAIS

PROGRAMA DE PÓS-GRADUAÇÃO EM ESTATÍSTICA



FOLHA DE APROVAÇÃO

" Um passeio aleatório com catástrofes"

CARLOS HENRIQUE SANTOS

Dissertação submetida à Banca Examinadora, designada pelo Colegiado do Programa de Pós-Graduação em ESTATÍSTICA, como requisito para obtenção do grau de Mestre em ESTATÍSTICA, área de concentração ESTATÍSTICA E PROBABILIDADE.

Aprovada em 04 de junho de 2024, pela banca constituída pelos membros:


Prof. Roger William Câmara Silva - Orientador
DEST/UFMG



Prof. Sokol Ndreca
DEST/UFMG


Prof. Pablo Almeida Gomes
IME/USP

Belo Horizonte, 4 de junho de 2024.

Agradecimentos

Agradeço à Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo mecenato que possibilitou a minha dedicação exclusiva para a formação de mestre. Também agradeço à todas outras agências de fomento à ciência, que indiretamente forneceram estrutura para a manutenção de um ambiente acadêmico. Agradeço à Universidade Federal de Minas Gerais, ao Instituto de Ciências Exatas e ao Departamento de Estatística.

Agradeço à todos docentes do departamento de estatística da UFMG que participaram, mesmo que indiretamente, da minha formação. Agradeço, em especial, ao meu professor orientador Dr. Roger William Camara Silva.

Agradeço ao senhor Carlos Magno dos Santos e à senhora Alessandra de Oliveira, meus pais queridos. Agradeço à toda minha família.

E agradeço àquele que me guardou de maus maiores, Jesus Cristo.

“... *Pois só por teu serviço navegamos.*”
(Luís de Camões, *Os Lusíadas.*)

Resumo

Dissertamos sobre os resultados e demonstrações do artigo [3], no qual os autores estudaram uma classe de modelos para crescimento populacional com catástrofes. Do ponto de vista matemático, esse tipo de modelo pode ser entendido como um passeio aleatório nos inteiros não negativos, cujas probabilidades de transições dependem do estado. O formidável dessa classe de passeios aleatórios é a sua estabilidade. De fato, uma classe ainda maior de passeios aleatórios são estáveis, veja [19].

Em acordo com o roteiro de [3], mostramos a recorrência positiva desse passeio e explicitamos a sua distribuição estacionária. Comentamos sobre a forte estabilidade desse processo. Logo após definimos um profícuo acoplamento para a cadeia de Markov, o qual utilizamos à exaustão para provar muitos resultados. Estabelecemos cotas inferiores e superiores para a distância em variação total, analisando o comportamento da convergência para a estacionariedade desse processo. Após essa análise, conseguimos discutir o principal resultado, segundo os próprios autores, de [3]: A classe de modelos apresenta o fenômeno de *cutoff*. O principal desafio na prova desse resultado é que o acoplamento definido fornece cotas para a distância em variação total apenas considerando os estados iniciais fixados, sendo o espaço de estados dessa cadeia infinito enumerável. Por fim, apresentamos alguns outros resultados interessantes, como a dependência logarítmica entre o estado inicial e o tempo de primeira extinção do processo.

Palavras-chave: modelos populacionais; catástrofes; acoplamento; cutoff; metaestabilidade.

Abstract

We discussed upon the results and proofs of the article [3], in which the authors studied a class of models for populations growth with catastrophes. From a mathematical point of view, this kind of model can be understood as a random walk in non-negative integers, whose transition probabilities are state depend. The formidable thing about this class of random walks is their stability. In fact, an even larger class of random walks are stable, see [19].

In accordance with [3]'s script, we proof the positive recurrence of this random walk and explain its stationary distribution. We commented on the strong stability of this process. Soon after, we defined a useful coupling for this Markov chain, which we used extensively to prove many results. We established lower and upper bounds for the distance in total variation, analyzing the behavior of convergence towards stationarity of this process. After this analysis we were able to discuss the main result, according to the authors themselves, of [3]: The class of models presents the phenomenon of *cutoff*. The main challenge in proving this result is that the defined coupling provides quotas for the distance in total variation only considering the fixed initial states, and the state space of this chain is countable. Finally, we present some other interesting results, such as the logarithmic dependence between the initial state and the time of first extinction of the process.

Keywords: population models; catastrophes; coupling; cutoff; metastability.

Lista de Figuras

2.1	Realização da Cadeia para $x_0 = 10$, $p = 0,4$ e $c = 0,1$	21
2.2	Realização da cadeia para $x_0 = 100$, $p = 0,4$ e $c = 0,1$	21
2.3	Realização da cadeia para $x_0 = 100$, $p = 0,4$ e $c = 0,01$	22
2.4	Realização da cadeia para $x_0 = 10$, $p = 0,4$ e $c = 0,01$	22
4.1	Ilustração cutoff.	41
4.2	Exemplo $n = 100$, intervalo $(1, 919)$	48
4.3	Exemplo $n = 500$, intervalo $(367, 5848)$	48
4.4	Exemplo $n = 1000$, intervalo $(1109, 12707)$	49

Sumário

1	Introdução	10
2	Conceitos elementares	12
2.1	Cadeias de Markov	12
2.1.1	Tipos de cadeias de Markov	13
2.2	Distribuição invariante	14
2.2.1	Distribuição invariante do passeio	15
2.3	Um comentário sobre a metaestabilidade do processo	19
3	Cotas para a convergência	23
3.1	Acoplamento em Cadeias de Markov	23
3.2	Acoplamento do modelo	24
3.2.1	Tempo de acoplamento	26
3.2.2	Cota superior	29
3.2.3	Cota inferior	31
4	O fenômeno de cutoff	37
4.1	Limite Poisson	37
4.2	Cutoff	39
4.2.1	Definição de Cutoff	39
4.2.2	Cutoff do modelo	42
4.2.3	Simulação do cutoff	47
5	Resultados adicionais	50
5.1	Representação por processos de ramificação	50
5.2	Tempo de primeira extinção	53
	Referências	59
	Apêndice A Contra exemplos Lema 2.2.1	61
A.1	Contra exemplo 1	61
A.2	Contra-exemplo 2	62
	Apêndice B Sobre a relação de X_t, H_t e N_t	64

Capítulo 1

Introdução

É sabido que, sob algumas condições, uma *Cadeia de Markov* possui única distribuição *estacionária* e a distribuição do processo converge para a distribuição estacionária [4]. Apoiado nesse resultado, foi desenvolvido a metodologia *Monte Carlo Markov Chain* (MCMC) [10, 18]. A ideia do MCMC é simular observações duma distribuição de probabilidade por meio dum algoritmo computacional que emula a dinâmica probabilística duma cadeia de Markov que tem como a sua distribuição estacionária a distribuição que se deseja ter observações. Essa metodologia foi aplicada em vários campos; problemas de otimização, inferência Bayesiana, Mecânica estatística, etc. Uma pergunta comum feita nas aplicações de MCMC é: a partir de quantas iterações um valor simulado pelo algoritmo pode ser considerado uma observação “legítima” da distribuição estacionária? Em termos teóricos, o quão rápido a distribuição da cadeia de Markov do algoritmo MCMC converge para a sua estacionária? Muitos resultados foram desenvolvidos sobre essa questão [13, 17, 12].

Nessa linha foi descoberto o fenômeno de *cutoff*, cuja a ideia foi introduzida por D. Aldous e P. Diaconis, [1, 2], para capturar o fato de que algumas cadeias de Markov convergem abruptamente, numa certa escala de tempo, para as suas distribuições estacionárias. A vantagem de cadeias de Markov que apresentam cutoff é que podemos ter uma descrição detalhada do comportamento de convergência da distribuições do processo para a distribuição estacionária. Contudo, essa análise pode ser onerosa a depender do modelo e por isso a maior parte dos resultados sobre cutoff dependem do modelo considerado.

O artigo objeto dessa dissertação será [3]. Nesse trabalho os autores estudaram um modelo para dinâmica populacional com *catástrofes*. Uma população pode ser afetada por vários fatores demográficos [16], um desses fatores são as catástrofes (ou desastres), significando que em certo momento grande parte ou toda população morre. Esses modelos com desastres são estudados deste 1970, [14], e dessa classe de modelos os que mais receberam atenção foram os *semi-estocásticos*, onde o crescimento da população é determinístico, mas as catástrofes são aleatórias. O modelo de [3] é dessa classe. Mais especificamente, o modelo que estudaremos será uma *cadeia de crescimento populacional com catástrofes binomial e imigração a tempo discreto*. Uma referência interessantíssima que apresentou outros tipos de desastres é [19]. Grande parte das pesquisas em dinâmicas

de populações consideram tempo contínuo, um análogo do modelo dessa dissertação com essa característica pode ser encontrado em [5].

A principal contribuição do nosso artigo referência, [3], segundo os próprios autores, foi que eles demonstraram que a classe de modelos que eles estudaram apresenta o fenômeno de cutoff. O objetivo dessa dissertação é apresentar e explicar os resultados e demonstrações desse artigo.

Capítulo 2

Conceitos elementares

Neste capítulo estabeleceremos algumas definições e notações básicas sobre Cadeia de Markov, e usaremos o próprio modelo do artigo referência para ilustrar esses conceitos. Consideraremos $\mathbb{Z}_+ = \mathbb{N} \cup \{0\}$.

2.1 Cadeias de Markov

Definição 2.1.1. Seja $(X_n)_{n \in \mathbb{Z}_+}$ um processo em tempo discreto tomando valores no conjunto enumerável, \mathcal{X} . Se para todo $n \in \mathbb{Z}_+$ e $i_0, \dots, i_{n-1}, i, j \in \mathcal{X}$,

$$\mathbb{P}(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) = \mathbb{P}(X_{n+1} = j | X_n = i) = \mathbb{P}(i, j), \quad (2.1)$$

em que $\mathbb{P}: \mathcal{X} \times \mathcal{X} \rightarrow [0, 1]$, $\sum_{j \in \mathcal{X}} \mathbb{P}(i, j) = 1$, então esse processo é uma *Cadeia de Markov Homogênea* com *espaço de estados* \mathcal{X} e *núcleo de transição* \mathbb{P} .

A palavra “homogênea” se refere ao tempo, visto que a função \mathbb{P} independe de n . Observe que a definição apresentada é bem restrita, é possível definições mais gerais sobre cadeias de Markov. Entretanto ela será suficiente para os propósitos desse texto. A função \mathbb{P} é comumente referida como probabilidade de transição, função de transição, lei de transição e núcleo de transição.

Uma das característica úteis desse tipo de processo é que seu comportamento probabilístico finito dimensional depende da distribuição de X_0 e das probabilidades de transição. Com efeito, para $n_k \geq n_1$ temos

$$\begin{aligned} \mathbb{P}(X_{n_1} = x_{n_1}, \dots, X_{n_k} = x_{n_k}) &= \sum_{(y_0, \dots, y_{n_k}): y_{n_i} = x_{n_i}} \mathbb{P}(X_0 = y_0, \dots, X_{n_k} = x_{n_k}) \\ &= \sum_{(y_0, \dots, y_{n_k}): y_{n_i} = x_{n_i}} \mathbb{P}(X_0 = y_0) \mathbb{P}(y_0, y_1) \dots \mathbb{P}(y_{n_k-1}, y_{n_k}). \end{aligned}$$

O modelo do artigo objeto segue essa definição. Dado $p \in (0, 1)$ e $c \in (0, 1]$, o processo $(X_n)_{n \in \mathbb{Z}_+}$ é uma Cadeia de Markov com espaço de estados \mathbb{Z}_+ e núcleo de transição $\mathbf{p}^{p,c}$, tal qual,

$$\mathbf{p}^{p,c}(i, j) = \begin{cases} p & \text{se } j = i + 1, \\ (1 - p) \binom{i}{j} (1 - c)^j c^{i-j} & \text{se } j \in \{0, \dots, i\}, \\ 0 & \text{caso contrário.} \end{cases} \quad (2.2)$$

Quando não houver perigo de engano, omitiremos os parâmetros p e c . Esse processo pode ser interpretado como uma população de tamanho X_t no momento t , a qual no próximo momento, independentemente do passado, ou, com probabilidade p , chegará um imigrante, ou, com probabilidade $1 - p$, ocorrerá uma catástrofe, e então cada indivíduo da população morre com probabilidade c independentemente dos outros indivíduos e do passado. Daí o nome do modelo tal como descrito na Introdução. Esse modelo pode ser visto como um passeio aleatório cujas transições dependem do estado. Por esse motivo, muitas vezes referiremos à este modelo como *o passeio com catástrofes*.

Antes de prosseguir para novas definições, precisamos fixar algumas notações. Dada uma cadeia de Markov $(Y_t)_{t \in \mathbb{Z}_+}$, denotaremos $\mathbb{P}(Y_{n+t} = j | Y_t = i) = \mathbb{P}^n(i, j)$, $i, j \in \mathcal{X}$. Devido à homogeneidade suposta na Definição 2.1.1, essa probabilidade não depende de t . Seja ν uma distribuição de probabilidade em \mathcal{X} , se Y_0 tiver distribuição ν , denotaremos a lei do processo por \mathbb{P}_ν e sua esperança \mathbb{E}_ν . Quando $\nu = \delta_j$, $j \in \mathcal{X}$, escrevemos $\mathbb{P}_j = \mathbb{P}_\nu$.

2.1.1 Tipos de cadeias de Markov

Definição 2.1.2. Uma cadeia de Markov $(Y_t)_{t \in \mathbb{Z}_+}$ é dita *irredutível* se, para todo $x, y \in \mathcal{X}$, existe $n \in \mathbb{N}$ tal que $\mathbb{P}^n(x, y) > 0$.

Em termos informais, o que a Definição 2.1.2 está dizendo é que uma cadeia de Markov é irredutível quando podemos alcançar qualquer estado partindo de qualquer estado, com probabilidade positiva, e numa quantidade finita de passos.

Definição 2.1.3. Para $x \in \mathcal{X}$, definimos o *período* de x para ser o maior divisor comum do conjunto de naturais $\{t \geq 1 : \mathbb{P}^t(x, x) > 0\}$.

Ou seja, o período de um estado é o máximo divisor comum dos conjuntos dos tempos onde é possível retornar ao estado com probabilidade positiva. Quando um estado têm período igual 1, nós o chamamos *aperiódico*.

Proposição 2.1.1 (Lemma 1.6 de [17]). *Numa cadeia de Markov irredutível, todos estados têm o mesmo período.*

Portanto, podemos classificar cadeias de Markov irredutíveis como aperiódicas e não aperiódicas. As do primeiro tipo serão de nosso interesse.

Definição 2.1.4. Seja $(Y_t)_{t \in \mathbb{Z}_+}$ uma cadeia de Markov. Definimos o *tempo de visita* ao $x \in \mathcal{X}$, por $\tau_x = \inf\{t \geq 0: Y_t = x\}$, e o *retorno*, por $\tau_x^+ = \inf\{t \geq 1: Y_t = x\}$.

Definição 2.1.5. O estado $x \in \mathcal{X}$ é *recorrente*, quando $\mathbb{P}_x(\tau_x^+ < \infty) = 1$. Caso contrário é *transiente*. É *recorrente positivo* quando $\mathbb{E}_x(\tau_x^+) < \infty$. Se x é recorrente mas não recorrente positivo dizemos que ele é *recorrente nulo*.

O que essa definição estabelece é a “estabilidade” dum estado. Poderíamos ter uma cadeia de Markov que visita um estado apenas uma vez, por exemplo, uma com espaço de estados \mathbb{Z}_+ e núcleo $\mathbb{P}(i, i+1) = 1$. Ou seja, para “ t grande”, a probabilidade do processo estar em algum estado é zero, isto é, a distribuição da cadeia “tende” à medida nula. A noção de recorrência positiva vêm para evitar essa situação.

Proposição 2.1.2 (Propositions 21.3 e 21.11 de [17]). *Numa cadeia de Markov irredutível, um estado é recorrente se, e somente se, todos estados o são. Um estado é recorrente positivo se, e somente se, todos estados o são.*

Qual seja, cadeias irredutíveis podem ser classificadas como recorrente positiva ou não recorrente positiva.

O processo da Equação 2.2 é irredutível aperiódico. Se $i, j \in \mathbb{Z}_+$, então

$$\mathbf{p}^{j+1}(i, j) \geq \mathbf{p}(i, 0)\mathbf{p}(0, 1) \dots \mathbf{p}(j-1, j) = (1-p)c^i p^j > 0.$$

Note que esse argumento é válido para todo c e p . Utilizando a Proposição 2.1.1, basta verificarmos o período do estado 0. Para todo $t \in \mathbb{N}$,

$$\mathbf{p}^t(0, 0) \geq (1-p)^t > 0.$$

Logo o estado 0 é aperiódico, e, portanto, cadeia é aperiódica. A recorrência positiva será verificada na próxima seção.

2.2 Distribuição invariante

Definição 2.2.1. Seja π uma medida de probabilidade definida em \mathcal{X} , e uma cadeia de Markov com núcleo de transição \mathbb{P} . Dizemos que π é uma distribuição *invariante*, ou *estacionária*, para \mathbb{P} , sempre que,

$$\sum_{x \in \mathcal{X}} \pi(x)\mathbb{P}(x, y) = \pi(y) \quad \forall y \in \mathcal{X}. \quad (2.3)$$

Nos termos da Definição 2.2.1, e observando a propriedade de Markov, podemos argumentar que se $Y_0 \sim \pi$, então $Y_1 \sim \pi$, e assim $Y_t \sim \pi, \forall t$. Daí o nome de distribuição invariante, ou estacionária.

Definição 2.2.2. Sejam μ e ν medidas de probabilidade definidas num conjunto enumerável \mathcal{X} , denotamos a *distância em variação total entre μ e ν* por

$$\|\mu - \nu\|_{TV} = \sup_{A \subset \mathcal{X}} |\mu(A) - \nu(A)| \in [0, 1].$$

Se Y_n têm distribuição μ_n e Y distribuição μ , com $\lim_{n \rightarrow \infty} \|\mu_n - \mu\|_{TV} = 0$, dizemos que Y_n converge em variação total para Y .

Genericamente, convergência em variação total é uma convergência entre medidas, e claro, esse tipo de convergência implica convergência em distribuição entre variáveis aleatórias.

A seguir, apresentaremos dois teoremas que fundamentarão os resultados do resto dessa dissertação.

Teorema 2.2.1 (Theorem 21.13 de [17]). *Uma cadeia de Markov irredutível com probabilidade de transição P e espaço de estados \mathcal{X} , é recorrente positiva se, e apenas se, existe uma distribuição de probabilidade π em \mathcal{X} , para a qual P é estacionária.*

Teorema 2.2.2 (Theorem 21.16 de [17]). *Seja $(Y_t)_{t \in \mathbb{Z}_+}$ uma cadeia de Markov, com lei de transição P e espaço de estados \mathcal{X} , irredutível e aperiódica. Se ela é recorrente positiva, então existe uma única medida de probabilidade π em \mathcal{X} , tal qual P é invariante para π e para qualquer $x \in \mathcal{X}$,*

$$\lim_{n \rightarrow \infty} \|P^n(x, \cdot) - \pi\|_{TV} = 0.$$

Perceba que os Teoremas 2.2.1 e 2.2.2 se complementam. Além disso, é possível mostrar que para uma cadeia como descrita no Teorema 2.2.2,

$$\pi(x) = \frac{1}{\mathbf{E}_x(\tau_x^+)} \quad \forall x \in \mathcal{X}. \quad (2.4)$$

Veja [17, Lemma 21.12]. Ou seja, $\pi(x) > 0$ para todo $x \in \mathcal{X}$. O processo do modelo do artigo referência detém essas boas características.

2.2.1 Distribuição invariante do passeio

Nesta seção, mostraremos que o passeio com catástrofes têm uma distribuição estacionária, e, por ser irredutível e aperiódica, a cadeia do passeio terá todas as propriedades

do Teorema 2.2.2. Para uma variável aleatória $R \in \mathbb{Z}_+$ e número real $\epsilon \in [0, 1]$, denotamos $\text{Bin}(R, \epsilon)$ uma variável aleatória, que condicionada em R , têm distribuição binomial com parâmetros R e ϵ .

Antes de prosseguirmos na demonstração do resultado, apresentamos um lema que será muito útil no decorrer dessa dissertação.

Lema 2.2.1 (Lemma 2.1 de [3]). *Suponha que R_0, R_1, \dots são variáveis aleatórias independentes tomando valores em \mathbb{Z}_+ e sejam $\epsilon_0, \epsilon_1, \dots$ uma sequência tomando valores em $[0, 1]$. Assuma que $\sum_j \epsilon_j E[R_j] < \infty$. Para $j = 0, 1, \dots$, seja $\text{Bin}_j(R_j, \epsilon_j)$ uma variável com distribuição $\text{Bin}(R_j, \epsilon_j)$, com $(\text{Bin}_j(R_j, \epsilon_j) : j \geq 1)$ independentes, condicionado em $(R_j : j \in \mathbb{Z}_+)$. Seja $R = \sum_{j=0}^{\infty} \text{Bin}_j(R_j, \epsilon_j)$ e $\epsilon > 0$. Então $\sum_{j=0}^{\infty} \text{Bin}_j(R_j, \epsilon \epsilon_j)$ e $\text{Bin}(R, \epsilon)$ coincidem em distribuição.*

O Lema 2.2.1 está apresentado tal como foi expresso no artigo referência, apesar de traduzido para o português. Numa primeira leitura ele pode ser um pouco enigmático, mas considere a seguinte explicação intuitiva do resultado.

Suponha que tenhamos n urnas com infinitas bolas pretas e brancas em cada urna. Fazemos dois procedimentos de amostragem das bolas dessas n urnas, e temos o interesse em deduzir a distribuição de probabilidade da quantidade de bolas pretas obtidas no final da amostragem.

Copiando a notação do Lema 2.2.1, ϵ_i é a probabilidade de retirarmos, aleatoriamente, uma bola preta da i -ésima urna.

Procedimento I: Decidimos, previamente, a quantidade de bolas a serem amostradas de cada urna, denotamos essa quantidade por R_i . A quantidade de bolas retiradas da urna i é independente da quantidade de bolas retiradas da urna j , sempre que $i \neq j$. Da urna 1 retiramos uma bola, se a bola amostrada for branca, guardo ela no cesto. Se ela for preta, lanço uma moeda que, têm probabilidade ϵ de dar cara, para decidir se a bola será pintada de branca. Se o resultado do lançamento for cara, guardamos a bola preta amostrada no cesto, se for coroa, pintamos a bola preta de branca e então guardamos ela no cesto. Fazemos esse procedimento, independentemente, para todas as R_1 bolas amostradas da urna 1.

Continuamos o procedimento acima, descrito para a urna 1, em todas as n urnas de maneira independente em cada urna. É claro que a quantidade de bolas amostradas da i -ésima urna é uma $\text{Bin}(R_i, \epsilon \epsilon_i)$, e $\text{Bin}_i(R_i, \epsilon \epsilon_i)$ (a quantidade de bolas retiradas da i -ésima urna) é independente de $\text{Bin}_j(R_j, \epsilon \epsilon_j)$, se $i \neq j$, dado os valores de R_i e R_j . Portanto, a quantidade de bolas pretas no cesto ao final desse procedimento tem distribuição $\sum_{j=1}^n \text{Bin}_j(R_j, \epsilon \epsilon_j)$.

Procedimento II: Feito em duas etapas. Decidimos quantas bolas serão amostradas de cada urna como anteriormente. Da urna i retiramos as R_i bolas aleatoriamente e

as guardamos no cesto, sem o passo intermediário do procedimento I. Ao final dessa primeira etapa, a distribuição da quantidade de bolas pretas coincidirá com a da variável $\sum_{j=1}^n \text{Bin}_j(R_j, \epsilon_j)$. Na próxima etapa, tomamos cada bola preta no cesto e lançamos uma moeda que têm probabilidade ϵ de dar cara. Se o resultado for cara, não pintamos a bola de branca, se for coroa pintamos ela de branca. Obviamente, a distribuição de bolas pretas no cesto ao final dessa segunda etapa é $\text{Bin}(\sum_{j=1}^n \text{Bin}_j(R_j, \epsilon_j), \epsilon)$.

Esse dois procedimento são probabilisticamente equivalentes, nos sugerindo, então, que $\sum_{j=1}^n \text{Bin}_j(R_j, \epsilon_j)$ e $\text{Bin}(\sum_{j=1}^n \text{Bin}_j(R_j, \epsilon_j), \epsilon)$ são iguais em distribuição.

Os autores do artigo referência omitiram a prova desse resultado, abaixo apresento uma. O leitor mais atento perceberá na prova que estamos usando hipóteses que não estão expressas no enunciado do Lema 2.2.1, as quais são $\text{Bin}_j(R_j, \epsilon_j)$ é independente dos $(R_i)_{i \neq j}$ dado R_j , e $\text{Bin}_0(R_0, \epsilon_0)$ é independente dos $(\text{Bin}_j(R_j, \epsilon_j) : j \geq 1)$ dado $(R_j)_{j \geq 0}$. A primeira creio que seja uma hipótese tácita entre os pesquisadores dessa área, devida a forma como a variável $\text{Bin}(R, \epsilon)$ foi definida. A segundo creio que tenha sido apenas erro de digitação. De todo modo, se uma dessas duas hipóteses não forem válidas então o Lema 2.2.1 é inválido. De fato, veja o Apêndice A. Segue a demonstração onde usaremos as propriedades da *função geradora de probabilidades*.

Demonstração. Começemos pelo caso finito. Temos que

$$\begin{aligned}
\mathbf{E}(s^{\text{Bin}(R, \epsilon)}) &= \mathbf{E}(\mathbf{E}(s^{\text{Bin}(R, \epsilon)} | R)) = \mathbf{E}((1 - \epsilon + \epsilon s)^{\sum_{j=0}^n \text{Bin}_j(R_j, \epsilon_j)}) \\
&= \mathbf{E}(\mathbf{E}((1 - \epsilon + \epsilon s)^{\sum_{j=0}^n \text{Bin}_j(R_j, \epsilon_j)} | (R_k)_{k \geq 0})) \\
&= \mathbf{E}\left(\prod_{j=0}^n \mathbf{E}((1 - \epsilon + \epsilon s)^{\text{Bin}_j(R_j, \epsilon_j)} | (R_k)_{k \geq 0})\right) \\
&= \mathbf{E}\left(\prod_{j=0}^n \mathbf{E}((1 - \epsilon + \epsilon s)^{\text{Bin}_j(R_j, \epsilon_j)} | R_j)\right) \\
&= \mathbf{E}\left(\prod_{j=0}^n (1 - \epsilon_j + \epsilon_j(1 - \epsilon + \epsilon s))^{R_j}\right) \\
&= \mathbf{E}\left(\prod_{j=0}^n (1 - \epsilon_j \epsilon + \epsilon_j \epsilon s)^{R_j}\right) \\
&= \mathbf{E}\left(\prod_{j=0}^n \mathbf{E}(s^{\text{Bin}_j(R_j, \epsilon_j \epsilon)} | R_j)\right) \\
&= \mathbf{E}\left(\prod_{j=0}^n \mathbf{E}(s^{\text{Bin}_j(R_j, \epsilon_j \epsilon)} | (R_k)_{k \geq 0})\right) \\
&= \mathbf{E}(\mathbf{E}(s^{\sum_{j=0}^n \text{Bin}_j(R_j, \epsilon_j \epsilon)} | (R_k)_{k \geq 0})) = \mathbf{E}(s^{\sum_{j=0}^n \text{Bin}_j(R_j, \epsilon_j \epsilon)}).
\end{aligned}$$

Agora, quando $n \rightarrow \infty$, lembre-se que para uma seqüência de variáveis aleatórias não-negativas, $\mathbf{P}(\sum_{k=1}^{\infty} X_k > t) = \lim_{n \rightarrow \infty} \mathbf{P}(\sum_{k=1}^n X_k > t)$. Então

$$\begin{aligned}
\mathbb{P}\left(\text{Bin}\left(\sum_{j=0}^{\infty} \text{Bin}_j(R_j, \epsilon_j), \epsilon\right) = m\right) &= \sum_{k=0}^{\infty} \mathbb{P}(\text{Bin}(k, \epsilon) = m) \mathbb{P}\left(\sum_{j=0}^{\infty} \text{Bin}_j(R_j, \epsilon_j) = k\right) \\
&= \sum_{k=0}^{\infty} \mathbb{P}(\text{Bin}(k, \epsilon) = m) \lim_{n \rightarrow \infty} \mathbb{P}\left(\sum_{j=0}^n \text{Bin}_j(R_j, \epsilon_j) = k\right) \\
&= \lim_{n \rightarrow \infty} \mathbb{P}\left(\text{Bin}\left(\sum_{j=0}^n \text{Bin}_j(R_j, \epsilon_j), \epsilon\right) = m\right) \\
&= \lim_{n \rightarrow \infty} \mathbb{P}\left(\sum_{j=0}^n \text{Bin}_j(R_j, \epsilon \epsilon_j) = m\right) \\
&= \mathbb{P}\left(\sum_{j=0}^{\infty} \text{Bin}_j(R_j, \epsilon \epsilon_j) = m\right).
\end{aligned}$$

De fato podemos aplicar o teorema da convergência dominada nessa situação, porque

$$\lim_{k \rightarrow \infty} \frac{\mathbb{P}(\text{Bin}(k+1, \epsilon) = m)}{\mathbb{P}(\text{Bin}(k, \epsilon) = m)} = \lim_{k \rightarrow \infty} \frac{\binom{k+1}{m} \epsilon^m (1-\epsilon)^{k+1-m}}{\binom{k}{m} \epsilon^m (1-\epsilon)^{k-m}} = \lim_{k \rightarrow \infty} \frac{\binom{k+1}{m}}{\binom{k}{m}} (1-\epsilon) = (1-\epsilon) < 1.$$

QED

Estamos em condição de apresentar a distribuição estacionário do modelo da Equação 2.2, usaremos $\stackrel{d}{=}$ para indicar igualdade em distribuição de variáveis aleatórias.

Proposição 2.2.1 (Proposition 2.2 de [3]). *Seja R_0, R_1, \dots uma sequência IID (independentes e identicamente distribuídas) de $\text{Geom}^-(1-p)$ (geométricas que contam o número de falhas até o primeiro sucesso), $\epsilon_j = (1-c)^j$ para $j \in \mathbb{Z}_+$. Seja, R a variável aleatória como no Lema 2.2.1 com $\epsilon = 1$, e π a distribuição dessa variável aleatória. Então π é a distribuição estacionária de \mathbf{p} .*

Demonstração. Supondo $X_0 \stackrel{d}{=} R$, mostraremos que $X_1 \stackrel{d}{=} R$ usando função geradora de probabilidade. Temos

$$\begin{aligned}
\mathbb{E}[s^{X_1}] &= \mathbb{E}[s^{X_1} | \text{catástrofe em } t=1] \mathbb{P}(\text{catástrofe em } t=1) + \\
&\quad \mathbb{E}[s^{X_1} | \text{não catástrofe em } t=1] \mathbb{P}(\text{não catástrofe em } t=1) \\
&= \mathbb{E}[s^{\text{Bin}(X_0, 1-c)}] (1-p) + \mathbb{E}[s^{X_0+1}] p \\
&= \mathbb{E}[s^{\text{Bin}(X_0, 1-c)}] (1-p) + \mathbb{E}[s^{X_0}] s p.
\end{aligned}$$

Pelo Lema 2.2.1 e como $(R_i)_{i \in \mathbb{Z}_+}$ são *IID*,

$$\begin{aligned} \mathbf{E}(s^{X_0}) &= \mathbf{E}\left(\prod_{j=0}^{\infty} s^{\text{Bin}_j(R_j, (1-c)^j)}\right) \\ &= \prod_{j=0}^{\infty} \mathbf{E}\left((1 - (1-c)^j + (1-c)^j s)^{R_j}\right) \\ &= \mathbf{E}(s^{\text{Bin}(X_0, 1-c)})\mathbf{E}(s^{R_0}). \end{aligned}$$

Desse a forma,

$$\mathbf{E}(s^{X_1}) = \left(\frac{(1-p)}{\mathbf{E}(s^{R_0})} + sp \right) \mathbf{E}(s^{X_0})$$

Observando que $\mathbf{E}(s^{R_0}) = \frac{1-p}{1-ps}$, o resultado segue.

QED

Ou seja, pelo Teorema 2.2.2, $(X_n)_{n \in \mathbb{Z}_+}$ converge em variação total para

$$\sum_{j=0}^{\infty} \text{Bin}_j(\text{Geom}_j^-(1-p), (1-c)^j).$$

Doravante, reservarei π para denotar a distribuição estacionária desse processo.

2.3 Um comentário sobre a metaestabilidade do processo

Um comportamento interessante do modelo da Equação 2.2 é a sua forte *persistência* para alguns valores de p e c . Na biologia, tempo de persistência é o tempo esperado até a extinção da população, e é uma quantidade de interesse em aplicações. Nas Figuras 2.1, 2.2, 2.3 e 2.4, apresentamos realizações do passeio para certos valores de parâmetros, x_0 seria o estado inicial da cadeia.

A reta vermelha designa um valor teórico, ela é a média estacionária do processo,

$$\mathbf{E}_{\pi}[X_t] = \sum_{j=0}^{\infty} \mathbf{E}[R_j](1-c)^j = \frac{p}{c(1-p)}.$$

Observe como o processo tende abruptamente para essa média, e se mantém passeando por ela com uma variabilidade que parece constante. Nas Figuras 2.1 e 2.2, o valor indicado pela reta vermelha é 6,6667, e nesse caso o processo toca o estado zero várias vezes. Perceba na Figura 2.1 como o processo parece estar constrangido pelo estado zero, que parece estar dentro de uma faixa de variabilidade crítica. Já nas Figuras 2.3 e 2.4, a média

estacionária é 66,6667, e nesse caso o processo não toca o zero nas vinte mil iterações observadas. Além disso, o processo parece estar variando livremente, sem aquele aspecto de truncamento observado na Figura 2.1.

Essas observações são justificadas notando que o *drift* local é

$$\mathbf{E}(X_{t+1}|X_t) - X_t = (X_t + 1)p + (1 - c)X_t(1 - p) = p\left(1 - \frac{X_t}{\mathbf{E}_\pi(X_t)}\right).$$

O processo de fato tende a passear em volta da média estacionária. Ademais é simples de deduzir que,

$$\pi(0) = \mathbf{P}\left(\sum_{j=0}^{\infty} \text{Bin}_j(\text{Geom}_j^-(1-p)) = 0\right) = \prod_{j=0}^{\infty} \frac{1-p}{1-p(1-(1-c)^j)}.$$

Pela Equação 2.4, temos que

$$\mathbf{E}_0(\tau_0^+) = \prod_{j=0}^{\infty} \left(1 + \frac{p}{1-p}(1-c)^j\right).$$

Usando que $x - \frac{x^2}{2} \leq \ln(1+x) \leq x$, $\forall x \in [0, 1)$, e fazendo $x = \frac{p}{1-p}(1-c)^j$ para todo j e somando, prova-se, para $p < 0,5$ que

$$\frac{p}{(1-p)c} - \frac{1}{2} \frac{p^2}{(1-p)^2(1-(1-c)^2)} \leq \ln \mathbf{E}_0(\tau_0^+) \leq \frac{p}{(1-p)c}.$$

Desta feita, para os parâmetros das Figuras 2.1 e 2.2, teremos $\mathbf{E}_0(\tau_0^+) \geq 244$. Já para os das Figuras 2.3 e 2.4, $\mathbf{E}_0(\tau_0^+) \geq 10^{24}$, justificando as nossas observações. Um estudo mais aprofundado sobre a persistência nesse modelo pode ser encontrado no artigo [9].

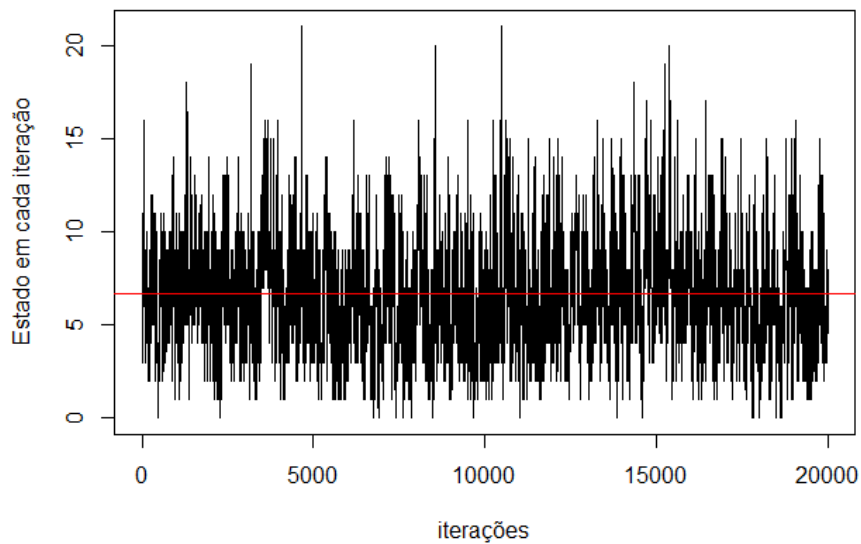


Figura 2.1: Realização da Cadeia para $x_0 = 10$, $p = 0,4$ e $c = 0,1$.

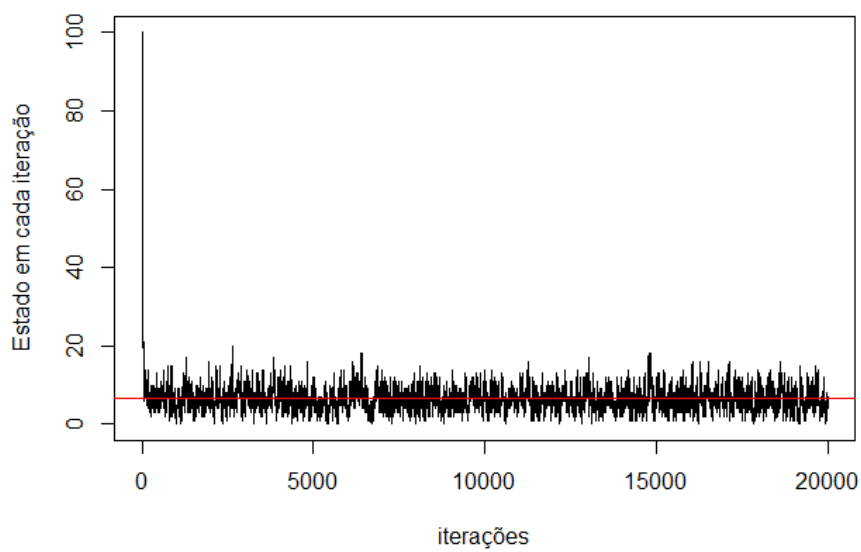


Figura 2.2: Realização da cadeia para $x_0 = 100$, $p = 0,4$ e $c = 0,1$.

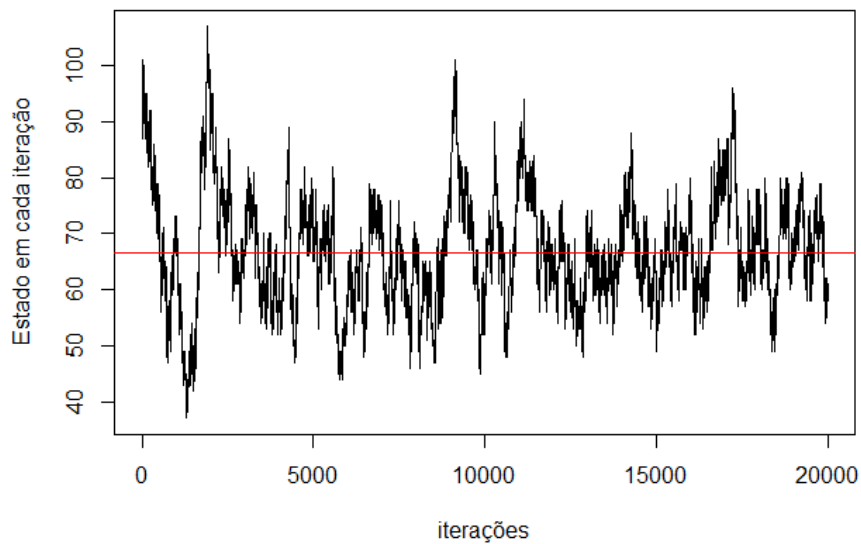


Figura 2.3: Realização da cadeia para $x_0 = 100$, $p = 0,4$ e $c = 0,01$.

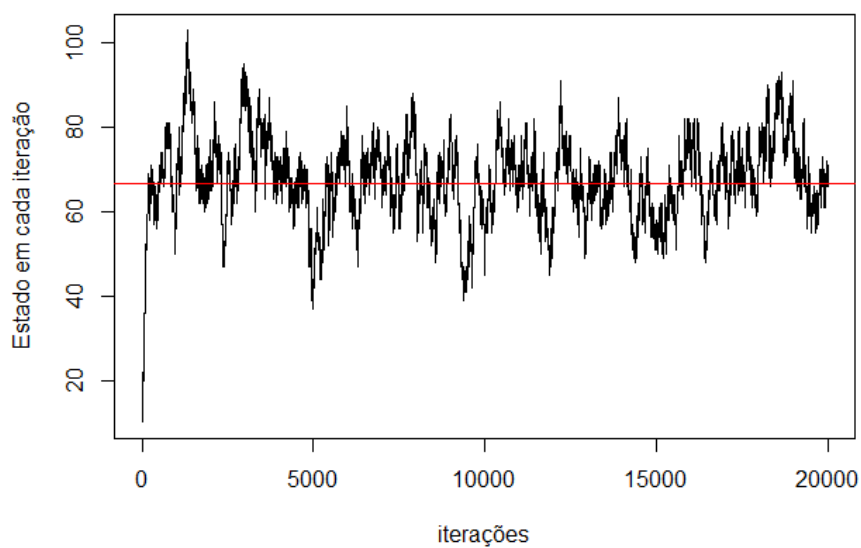


Figura 2.4: Realização da cadeia para $x_0 = 10$, $p = 0,4$ e $c = 0,01$.

Capítulo 3

Cotas para a convergência

3.1 Acoplamento em Cadeias de Markov

O método de acoplamento é bem geral e pode ser aplicado em muitos problemas de probabilidade. A ideia desse método é transformar uma questão inicialmente formulada em termos de medidas de probabilidade, numa outra equivalente, formulada em termos de variáveis aleatórias. Ao leitor interessado em aplicações gerais desse método, em [6] se encontram aplicações muito interessantes. Nosso foco será em acoplamentos para cadeias de Markov, cuja a definição segue.

Definição 3.1.1. Seja $(Z_t)_{t \in \mathbb{Z}_+}$ uma cadeia de Markov tal como na Definição 2.1.1. Um *acoplamento* para $(Z_t)_{t \in \mathbb{Z}_+}$, é um processo $(X_t, Y_t)_{t \in \mathbb{Z}_+}$, tal que, os processos marginais $(X_t)_{t \in \mathbb{Z}_+}$ e $(Y_t)_{t \in \mathbb{Z}_+}$ são cópias de $(Z_t)_{t \in \mathbb{Z}_+}$, isto é, são cadeias de Markov com mesmo núcleo de transição e espaço de estados da cadeia acoplada, apesar de terem distribuições iniciais possivelmente diferentes entre si.

Veja, o processo $(X_t, Y_t)_{t \in \mathbb{Z}_+}$ da Definição 3.1.1 está definido num certo espaço de probabilidade. Se $X_0 = x$ e $Y_0 = y$ denotaremos a medida desse espaço por $P_{x,y}$.

Teorema 3.1.1 (Theorem 5.4 de [17]). *Seja $(X_t, Y_t)_{t \in \mathbb{Z}_+}$ um acoplamento para uma cadeia de Markov como na Definição 2.1.1, $X_0 = x$, $Y_0 = y$ e com a propriedade de que $X_s = Y_s$ implica $X_t = Y_t$ $t \geq s$. Faça*

$$\tau_{couple} = \min\{t: X_s = Y_s \forall s \geq t\}. \quad (3.1)$$

Então,

$$\|P^t(x, \cdot) - P^t(y, \cdot)\|_{VT} \leq P_{x,y}(\tau_{couple} > t). \quad (3.2)$$

O resultado desse teorema ilustra o poder do acoplamento. Mas não só por isso, o acoplamento é muito útil para provar outros tipos de resultados, pois ele nos fornece um método probabilisticamente intuitivo de pensar questões sobre medidas de probabilidade

em termos de variáveis aleatórias. De fato, o artigo [3] gira em torno de um acoplamento que definiremos na próxima seção.

3.2 Acoplamento do modelo

Agora apresentaremos a regra de acoplamento para a cadeia que estamos estudando. Essa regra está no artigo referência e, como será visto no decorrer dessa dissertação, será constantemente usada para provar vários resultados. Por isso, sempre que me referir à um acoplamento será à este acoplamento.

Definição 3.2.1 (Regra do acoplamento). Sejam $x, y \in \mathbb{Z}_+$ com $x < y$. Ponha $X_0 = x, X'_0 = y$, e $H_0 = y - x$. Continue indutivamente, assumindo $((X_s, X'_s, H_s), s \leq t)$ definidos e $X'_s = X_s + H_s$ para todo $s \leq t$. Condicionado em $((X_s, X'_s, H_s), s \leq t)$,

- (c1) : Com probabilidade p , independentemente do passado, $H_{t+1} = H_t, X_{t+1} = X_t + 1$ e $X'_{t+1} = X'_t + 1$.
- (c2) : Com probabilidade $1 - p$, faça

$$X_{t+1} = \text{Bin}(X_t, 1 - c) \quad \text{e} \quad H_{t+1} = \text{Bin}(H_t, 1 - c),$$

independente um do outro e do passado. Ademais, $X'_{t+1} = X_{t+1} + H_{t+1}$.

Segue imediatamente da regra que $X'_t = H_t + X_t \quad \forall t \in \mathbb{Z}_+$, e $(H_t)_{t \in \mathbb{Z}_+}$ é uma sequência não crescente. Definamos uma sequência no mesmo espaço de probabilidade do acoplamento, que será de muita utilidade.

$$\omega_t = \begin{cases} 0 & \text{se ocorreu c1 no tempo } t; \\ 1 & \text{se ocorreu c2 no tempo } t. \end{cases} \quad (3.3)$$

Obviamente, $(\omega_t)_{t \in \mathbb{Z}_+}$ é uma sequência *IID* de $\text{Ber}(1 - p)$. Ademais, ω_{t+1} é independente de $((X_s, X'_s, H_s), s \leq t)$, pela Definição 3.2.1.

Proposição 3.2.1. *A regra da Definição 3.2.1 produz um acoplamento, nos termos da Definição 3.1.1, para o passeio definido na Equação 2.2.*

Demonstração. Devemos concluir que

$$\mathbb{P}_{x,y}(X_t = z) = \mathbf{p}^t(x, z) \quad \text{e} \quad \mathbb{P}_{x,y}(X'_t = z) = \mathbf{p}^t(y, z), \quad \forall z \in \mathbb{Z}_+ \quad \forall t \in \mathbb{N}. \quad (3.4)$$

Faremos um argumento por indução em t . Observe que

$$\begin{aligned}
P_{x,y}(X_1 = z) &= P_{x,y}(X_1 = z | X_0 = x, X'_0 = y, H_0 = y - x) \\
&= P_{x,y}(X_1 = z | X_0 = x, X'_0 = y, H_0 = y - x, \omega_1 = 0)P_{x,y}(\omega_1 = 0) + \\
&P_{x,y}(X_1 = z | X_0 = x, X'_0 = y, H_0 = y - x, \omega_1 = 1)P_{x,y}(\omega_1 = 1) \\
&= \mathbb{I}_{\{x+1\}}(z)p + (1-p) \binom{x}{z} (1-c)^z c^{x-z} \mathbb{I}_{\{0, \dots, x\}}(z) = \mathbf{p}^{p,c}(x, z).
\end{aligned}$$

Além disso, temos

$$\begin{aligned}
P_{x,y}(X'_1 = z) &= P_{x,y}(X'_1 = z | X_0 = x, X'_0 = y, H_0 = y - x) \\
&= P_{x,y}(X'_1 = z | X_0 = x, X'_0 = y, H_0 = y - x, \omega_1 = 0)P_{x,y}(\omega_1 = 0) + \\
&P_{x,y}(X'_1 = z | X_0 = x, X'_0 = y, H_0 = y - x, \omega_1 = 1)P_{x,y}(\omega_1 = 1) \\
&= \mathbb{I}_{\{y+1\}}(z)p + P_{x,y}(\text{Bin}(x, 1-c) + \text{Bin}(y-x, 1-c) = z)(1-p) \\
&= \mathbb{I}_{\{y+1\}}(z)p + (1-p) \binom{y}{z} (1-c)^z c^{y-z} \mathbb{I}_{\{0, \dots, y\}}(z) = \mathbf{p}^{p,c}(y, z).
\end{aligned}$$

Suponha, por indução, que $P_{x,y}(X_t = z) = \mathbf{p}^t(x, z)$ e $P_{x,y}(X'_t = z) = \mathbf{p}^t(y, z)$. Então,

$$\begin{aligned}
P_{x,y}(X_{t+1} = z) &= \sum_{w \in \mathbb{Z}_+} P_{x,y}(X_{t+1} = z | X_t = w) P_{x,y}(X_t = w), \\
P_{x,y}(X'_{t+1} = z) &= \sum_{w \in \mathbb{Z}_+} P_{x,y}(X'_{t+1} = z | X'_t = w) P_{x,y}(X'_t = w).
\end{aligned}$$

Além disso, observe que

$$\begin{aligned}
P_{x,y}(X_{t+1} = z | X_t = w) &= \sum_{(\nu_s; s \leq t)} \left[P_{x,y}(X_{t+1} = z | (X_s, X'_s, H_s; s \leq t) = (\nu_s; s \leq t), X_t = w) \right. \\
&P_{x,y}((X_s, X'_s, H_s; s \leq t) = (\nu_s; s \leq t) | X_t = w) \left. \right], \\
P_{x,y}(X'_{t+1} = z | X'_t = w) &= \sum_{(\nu_s; s \leq t)} \left[P_{x,y}(X'_{t+1} = z | (X_s, X'_s, H_s; s \leq t) = (\nu_s; s \leq t), X'_t = w) \right. \\
&P_{x,y}((X_s, X'_s, H_s; s \leq t) = (\nu_s; s \leq t) | X'_t = w) \left. \right].
\end{aligned}$$

Note também que

$$\begin{aligned}
& \mathbb{P}_{x,y}(X_{t+1} = z | (X_s, X'_s, H_s; s \leq t) = (\nu_s; s \leq t), X_t = w) \\
&= \mathbb{P}_{x,y}(X_{t+1} = z | (X_s, X'_s, H_s; s \leq t) = (\nu_s; s \leq t), X_t = w, \omega_{t+1} = 0) \mathbb{P}_{x,y}(\omega_{t+1} = 0) + \\
& \mathbb{P}_{x,y}(X_{t+1} = z | (X_s, X'_s, H_s; s \leq t) = (\nu_s; s \leq t), X_t = w, \omega_{t+1} = 1) \mathbb{P}_{x,y}(\omega_{t+1} = 1) \\
&= \mathbb{I}_{\{w+1\}}(z) p + \binom{w}{z} (1-c)^z c^{w-z} \mathbb{I}_{\{0, \dots, w\}}(z) (1-p) = \mathbf{p}(w, z), \\
& \mathbb{P}_{x,y}(X'_{t+1} = z | (X_s, X'_s, H_s; s \leq t) = (\nu_s; s \leq t), X'_t = w) \\
&= \mathbb{P}_{x,y}(X'_{t+1} = z | (X_s, X'_s, H_s; s \leq t) = (\nu_s; s \leq t), X'_t = w, \omega_{t+1} = 0) \mathbb{P}_{x,y}(\omega_{t+1} = 0) + \\
& \mathbb{P}_{x,y}(X'_{t+1} = z | (X_s, X'_s, H_s; s \leq t) = (\nu_s; s \leq t), X'_t = w, \omega_{t+1} = 1) \mathbb{P}_{x,y}(\omega_{t+1} = 1) \\
&= \mathbb{I}_{\{w+1\}}(z) p + \mathbb{P}_{x,y}(\text{Bin}(\nu_s(1), (1-c)) + \text{Bin}(\nu_s(3), (1-c)) = z) (1-p) \\
&= \mathbb{I}_{\{w+1\}}(z) p + \mathbb{P}_{x,y}(\text{Bin}(\nu_s(1) + \nu_s(3), (1-c)) = z) (1-p) \\
&= \mathbb{I}_{\{w+1\}}(z) p + \mathbb{P}_{x,y}(\text{Bin}(w, (1-c)) = z) (1-p) = \mathbf{p}(w, z).
\end{aligned}$$

Dessa forma, segue que

$$\begin{aligned}
& \mathbb{P}_{x,y}(X_{t+1} = z | X_t = w) = \\
& \mathbf{p}(w, z) \sum_{(\nu_s; s \leq t)} \mathbb{P}_{x,y}((X_s, X'_s, H_s; s \leq t) = (\nu_s; s \leq t) | X_t = w) = \mathbf{p}(w, z) \quad e \\
& \mathbb{P}_{x,y}(X'_{t+1} = z | X'_t = w) = \\
& \mathbf{p}(w, z) \sum_{(\nu_s; s \leq t)} \mathbb{P}_{x,y}((X_s, X'_s, H_s; s \leq t) = (\nu_s; s \leq t) | X'_t = w) = \mathbf{p}(w, z).
\end{aligned}$$

Finalmente, concluimos que

$$\begin{aligned}
\mathbb{P}_{x,y}(X_{t+1} = z) &= \sum_{w \in \mathbb{Z}_+} \mathbb{P}_{x,y}(X_{t+1} = z | X_t = w) \mathbb{P}_{x,y}(X_t = w) = \sum_{w \in \mathbb{Z}_+} \mathbf{p}(w, z) \mathbf{p}^t(x, w) = \mathbf{p}^{t+1}(x, z), \\
\mathbb{P}_{x,y}(X'_{t+1} = z) &= \sum_{w \in \mathbb{Z}_+} \mathbb{P}_{x,y}(X'_{t+1} = z | X'_t = w) \mathbb{P}_{x,y}(X'_t = w) = \sum_{w \in \mathbb{Z}_+} \mathbf{p}(w, z) \mathbf{p}^t(y, w) = \mathbf{p}^{t+1}(y, z).
\end{aligned}$$

QED

3.2.1 Tempo de acoplamento

Pela Definição 3.2.1, o acoplamento têm a propriedade de que $X_s = Y_s$ implica $X_t = Y_t \quad \forall t \geq s$, e portanto podemos usar o tempo de acoplamento para cotar superiormente a distância em variação total do processo em todo tempo. Nesta seção, analisaremos a variável

$$\xi = \inf\{t \geq 0 : X_t = X'_t\} = \inf\{t \geq 0 : H_t = 0\}. \quad (3.5)$$

Como $(H_t)_{t \in \mathbb{Z}_+}$ é um processo monótono não decrescente, $\{\xi \leq t\} = \{H_t = 0\}$, $\forall t$. Usaremos dessa relação para provar muitos resultados para ξ . Seja N_t a quantidade de catástrofes até o momento t , logicamente $N_t = \sum_{k=1}^t \omega_k \stackrel{d}{=} \text{Bin}(t, 1-p)$.

Lema 3.2.1. *Vale que*

$$\mathbb{P}_{x,y}(H_t \in \cdot | N_t) \stackrel{d}{=} \text{Bin}(y-x, (1-c)^{N_t}). \quad (3.6)$$

A prova do Lema 3.2.1 não é feita no artigo, mas será feita aqui. Mas antes, uma consideração sobre a intuição dessa relação. A variável H_t , no contexto da regra de acoplamento, é a discrepância entre X'_t e X_t . Sabemos que a população de H_t nunca aumenta, ou ela se mantém constante, quando não acontece uma catástrofe, ou ela diminui como uma binomial no momento de uma catástrofe, isto é, cada indivíduo morre independentemente com probabilidade c . Então, dado que até o tempo t ocorreram m catástrofes ($N_t = m$), um indivíduo que inicialmente estava em H_0 , estará em H_t , se sobreviver às m catástrofes ocorridas até o tempo t . Como a ocorrência de catástrofes e a sobrevivência dos indivíduos são eventos independentes, perceberemos que $H_t | \{N_t = m\} \stackrel{d}{=} \text{Bin}(y-x, (1-c)^m)$.

Demonstração. Faremos uma indução sobre t . Para $t = 1$, $N_1 \stackrel{d}{=} \text{Bin}(1, 1-p)$. Então, pela regra do acoplamento temos,

$$\begin{aligned} & \mathbb{P}_{x,y}(H_1 = j | N_1 = 0) \\ &= \sum_{\nu_0} \mathbb{P}_{x,y}(H_1 = j | (X_0, X'_0, H_0) = \nu_0, N_1 = 0) \mathbb{P}_{x,y}((X_0, X'_0, H_0) = \nu_0 | N_1 = 0) \\ &= \mathbb{P}_{x,y}(H_1 = j | X_0 = x, X'_0 = y, H_0 = y-x, N_1 = 0) = \mathbb{I}_{\{y-x\}}(j). \end{aligned}$$

Similarmente,

$$\begin{aligned} \mathbb{P}_{x,y}(H_1 = j | N_1 = 1) &= \mathbb{P}_{x,y}(H_1 = j | X_0 = x, X'_0 = y, H_0 = y-x, N_1 = 1) \\ &= \binom{y-x}{j} (1-c)^j c^{y-x-j} \mathbb{I}_{\{0, \dots, y-x\}}(j). \end{aligned}$$

Portanto,

$$\mathbb{P}_{x,y}(H_1 \in \cdot | N_1) \stackrel{d}{=} \text{Bin}(y-x, (1-c)^{N_1}).$$

Suponha a validade da sentença para $t = k$, isto é,

$$\mathbb{P}_{x,y}(H_k \in \cdot | N_k) \stackrel{d}{=} \text{Bin}(y-x, (1-c)^{N_k}).$$

Fazendo $t = k+1$ e $m > 0$, obtemos

$$\begin{aligned} \mathbb{P}_{x,y}(H_{k+1} = j | N_{k+1} = m) &= \sum_{s \in \mathbb{Z}_+} \mathbb{P}_{x,y}(H_{k+1} = j | N_{k+1} = m, N_k = s) \mathbb{P}_{x,y}(N_k = s | N_{k+1} = m) \\ &= \mathbb{P}_{x,y}(H_{k+1} = j | N_{k+1} = m, N_k = m-1) \mathbb{P}_{x,y}(N_k = m-1 | N_{k+1} = m) + \\ & \mathbb{P}_{x,y}(H_{k+1} = j | N_{k+1} = m, N_k = m) \mathbb{P}_{x,y}(N_k = m | N_{k+1} = m). \end{aligned}$$

A segunda igualdade do argumento acima é válido pois

$$\{N_{k+1} = m, N_k = m + k\} = \emptyset \quad \forall k \in \mathbb{Z} \setminus \{0, -1\}.$$

Observe que, $\{N_{k+1} = m, N_k = m\} = \{\omega_{k+1} = 0, N_k = m\}$, denote $\mathbf{V}_s = (X_s, X'_s, H_s)$. Logo,

$$\begin{aligned} \mathbb{P}_{x,y}(H_{k+1} = j | N_{k+1} = m, N_k = m) &= \mathbb{P}_{x,y}(H_{k+1} = j | \omega_{k+1} = 0, N_k = m) \\ &= \frac{\mathbb{P}_{x,y}(H_{k+1} = j, \omega_{k+1} = 0, N_k = m)}{\mathbb{P}_{x,y}(\omega_{k+1} = 0, N_k = m)} \\ &= \frac{\mathbb{E}_{x,y}(\mathbb{P}_{x,y}(H_{k+1} = j, \omega_{k+1} = 0, N_k = m | (\mathbf{V}_s; s \leq k)))}{\mathbb{E}_{x,y}(\mathbb{P}_{x,y}(\omega_{k+1} = 0, N_k = m | (\mathbf{V}_s; s \leq k)))} \\ &= \frac{\mathbb{E}_{x,y}(\mathbb{P}_{x,y}(H_k = j, \omega_{k+1} = 0, N_k = m | (\mathbf{V}_s; s \leq k)))}{\mathbb{E}_{x,y}(\mathbb{P}_{x,y}(\omega_{k+1} = 0 | (\mathbf{V}_s; s \leq k)) \mathbb{P}_{x,y}(N_k = m | (\mathbf{V}_s; s \leq k)))} \\ &= \frac{\mathbb{E}_{x,y}(\mathbb{P}_{x,y}(\omega_{k+1} = 0) \mathbb{P}_{x,y}(H_k = j, N_k = m | (\mathbf{V}_s; s \leq k)))}{\mathbb{P}_{x,y}(\omega_{k+1} = 0) \mathbb{E}_{x,y}(\mathbb{P}_{x,y}(N_k = m | (\mathbf{V}_s; s \leq k)))} \\ &= \mathbb{P}_{x,y}(H_k = j | N_k = m) \\ &= \mathbb{P}_{x,y}(\text{Bin}(y - x, (1 - c)^m) = j). \end{aligned}$$

A quarta igualdade segue porque, dado a processo até o momento k ,

$$\{H_{k+1} = j, \omega_{k+1} = 0\} = \{H_k = j, \omega_{k+1} = 0\},$$

e ω_{k+1} é independente de N_k . Para a quinta igualdade, lembre ω_t é independente do passado. Na sétima igualdade usamos a hipótese de indução. Agora analisemos a outra parcela. Observe que

$$\begin{aligned} \mathbb{P}_{x,y}(H_{k+1} = j | N_{k+1} = m, N_k = m - 1) &= \mathbb{P}_{x,y}(H_{k+1} = j | \omega_{k+1} = 1, N_k = m - 1) \\ &= \sum_{i \in \mathbb{Z}_+} \mathbb{P}_{x,y}(H_{k+1} = j | N_k = m - 1, H_k = i, \omega_{k+1} = 1) \mathbb{P}_{x,y}(H_k = i | N_k = m - 1, \omega_{k+1} = 1). \end{aligned}$$

Mas,

$$\begin{aligned} \mathbb{P}_{x,y}(H_k = i | N_k = m - 1, \omega_{k+1} = 1) &= \frac{\mathbb{P}_{x,y}(\omega_{k+1} = 1 | H_k = i, N_k = m - 1)}{\mathbb{P}_{x,y}(\omega_{k+1} = 1 | H_k = i)} \mathbb{P}_{x,y}(H_k = i | N_k = m - 1) \\ &= \mathbb{P}_{x,y}(H_k = i | N_k = m - 1) \\ &= \mathbb{P}_{x,y}(\text{Bin}(y - x, (1 - c)^{m-1}) = i). \end{aligned}$$

Na segunda igualdade usamos que ω_t é independente do passado. Já o outro fator,

$$\begin{aligned} \mathbb{P}_{x,y}(H_{k+1} = j | H_k = i, N_k = m - 1, \omega_{k+1} = 1) &= \mathbb{E}_{x,y}[\mathbb{P}_{x,y}(H_{k+1} = j | H_k = i, N_k = m - 1, \omega_{k+1} = 1, (\mathbf{V}_s; s \leq k)) | H_k = i, N_k = m - 1, \omega_{k+1} = 1] \\ &= \mathbb{E}_{x,y}[\mathbb{P}_{x,y}(\text{Bin}(i, (1 - c)) = j) | H_k = i, N_k = m - 1, C_{k+1} = 1] \\ &= \mathbb{P}_{x,y}(\text{Bin}(i, (1 - c)) = j). \end{aligned}$$

Na segunda igualdade, dado ω_{k+1} e $H_k = i$, $H_{k+1} \stackrel{d}{=} \text{Bin}(i, 1 - c)$ independentemente do passado. Qual seja,

$$\begin{aligned} & \mathbb{P}_{x,y}(H_{k+1} = j | N_{k+1} = m, N_k = m - 1) \\ &= \sum_{i \in \mathbb{Z}_+} \mathbb{P}_{x,y}(\text{Bin}(i, (1 - c)) = j) \mathbb{P}_{x,y}(\text{Bin}(y - x, (1 - c)^{m-1}) = i) \\ &= \mathbb{E}_{x,y}[\mathbb{P}_{x,y}(\text{Bin}(\text{Bin}(y - x, (1 - c)^{m-1}), (1 - c)) = j)] \\ &= \mathbb{P}_{x,y}(\text{Bin}(y - x, (1 - c)^m) = j). \end{aligned}$$

A terceira igualdade vale devido o Lema 2.2.1. Portanto,

$$\begin{aligned} & \mathbb{P}_{x,y}(H_{k+1} = j | N_{k+1} = m) \\ &= [\mathbb{P}_{x,y}(N_k = m - 1 | N_{k+1} = m) + \mathbb{P}_{x,y}(N_k = m | N_{k+1} = m)] \mathbb{P}_{x,y}(\text{Bin}(y - x, (1 - c)^m) = j) \\ &= \mathbb{P}_{x,y}(\text{Bin}(y - x, (1 - c)^m) = j). \end{aligned}$$

Isso tudo considerando $m > 0$. Se $m = 0$ significa que não houve catástrofes até $t + 1$, assim

$$\mathbb{P}_{x,y}(H_{k+1} = j | N_{k+1} = 0) = \mathbb{I}_{y-x}(j) = \mathbb{P}_{x,y}(\text{Bin}(y - x, (1 - c)^0) = j).$$

QED

3.2.2 Cota superior

Denotemos $d_t(x, y) = \|\mathbf{p}^t(x, \cdot) - \mathbf{p}^t(y, \cdot)\|_{VT}$. Utilizando o Lema 3.2.1 podemos provar o seguinte resultado.

Teorema 3.2.1 (Theorem 2.2 de [3]). *Seja $\alpha = 1 - c(1 - p)$. Então, para $x, y, t \in \mathbb{Z}_+$*

$$d_t(x, y) \leq |y - x| \alpha^t. \quad (3.7)$$

Demonstração. Sem perda de generalidade considere $x < y$. Para provar esse resultado cotaremos $\mathbb{P}_{x,y}(\xi > t)$, e pelo Teorema 3.1.1 o resultado estará provado. Temos que

$$\begin{aligned} \mathbb{P}_{x,y}(\xi > t) &= \mathbb{P}_{x,y}(H_t > 0) \\ &= 1 - \mathbb{E}_{x,y}(\mathbb{P}_{x,y}(H_t = 0 | N_t)) \\ &= 1 - \mathbb{E}_{x,y}((1 - (1 - c)^{N_t})^{y-x}) \\ &\leq (y - x) \mathbb{E}_{x,y}[(1 - c)^{N_t}] \\ &= (y - x)(1 - c(1 - p))^t. \end{aligned}$$

Na única desigualdade da argumentação usamos a Desigualdade de Bernoulli, isto é,

$$(1+x)^n \geq 1+nx, \quad x \geq -1 \text{ e } n \in \mathbb{N},$$

aplicado em $-(1-(1-c)^{N_t})^{y-x}$, pois $-(1-c)^{N_t} > -1$.

QED

Considerando $x < y$, um fato que vale ser mencionado é que $\mathbb{P}_{x,y}(\xi > t)$ e $(y-x)\alpha^t$ são muito similares quando t é grande. Isto é,

$$\begin{aligned} \lim_{t \rightarrow \infty} \frac{\mathbb{P}_{x,y}(\xi > t)}{(y-x)\mathbb{E}_{x,y}[(1-c)^{N_t}]} &= \lim_{t \rightarrow \infty} \frac{1 - \mathbb{E}_{x,y}[(1-(1-c)^{N_t})^{y-x}]}{(y-x)\mathbb{E}_{x,y}[(1-c)^{N_t}]} \\ &= \lim_{t \rightarrow \infty} \frac{1 - \mathbb{E}_{x,y}[\sum_{k=0}^{y-x} (-1)^k \binom{y-x}{k} (1-c)^{kN_t}]}{(y-x)\mathbb{E}_{x,y}[(1-c)^{N_t}]} \\ &= \lim_{t \rightarrow \infty} \frac{(y-x)\mathbb{E}_{x,y}[(1-c)^{N_t}]}{(y-x)\mathbb{E}_{x,y}[(1-c)^{N_t}]} + \lim_{t \rightarrow \infty} \frac{\sum_{k=2}^{y-x} (-1)^{k+1} \binom{y-x}{k} \mathbb{E}_{x,y}[(1-c)^{kN_t}]}{(y-x)\mathbb{E}_{x,y}[(1-c)^{N_t}]} \\ &= 1 + \lim_{t \rightarrow \infty} \sum_{k=2}^{y-x} \frac{(-1)^{k+1}}{y-x} \binom{y-x}{k} \frac{(1-(1-p) + (1-p)(1-c)^k)^t}{(1-(1-p) + (1-p)(1-c))^t} \\ &= 1 + \sum_{k=2}^{y-x} \frac{(-1)^{k+1}}{y-x} \binom{y-x}{k} \lim_{t \rightarrow \infty} \left[\frac{(p + (1-p)(1-c)^k)}{(p + (1-p)(1-c))} \right]^t = 1. \end{aligned}$$

Lembremos que π é a distribuição estacionária do processo do modelo que estamos estudando, e denotamos $\mu = \frac{p}{c(1-p)} = \mathbb{E}_\pi(X_t)$.

Corolário 3.2.1 (Corollary 3.4 de [3]). *Para todo $x, t \in \mathbb{Z}_+$,*

$$d_t(x, \pi) = \|\mathbb{P}^t(x, \cdot) - \pi\|_{VT} \leq \left(x - \mu + 2 \sum_{y>x} (y-x)\pi(y) \right) \alpha^t.$$

Em particular,

$$d_t(0, \pi) \leq \mu \alpha^t.$$

Demonstração. Para todo $A \subseteq \mathbb{Z}_+$,

$$|\mathbb{P}_x(X_t \in A) - \pi(A)| \leq \sum_{y \in \mathbb{Z}_+} |\mathbb{P}_x(X_t \in A) - \mathbb{P}_y(X_t \in A)| \pi(y) \leq \left[\sum_{y \in \mathbb{Z}_+} |y-x| \pi(y) \right] \alpha^t$$

Mas,

$$\sum_y |y-x| \pi(y) = \sum_{y \leq x} (x-y)\pi(y) + \sum_{y > x} (y-x)\pi(y) = x - \sum_y y\pi(y) + 2 \sum_{y > x} (y-x)\pi(y).$$

QED

3.2.3 Cota inferior

Denotaremos

$$\tilde{p} = \frac{p}{\alpha} = \frac{p}{p + (1-c)(1-p)}.$$

Usaremos $\mathbf{P}_x^{(\tilde{p})}$, $\mathbf{E}_x^{(\tilde{p})}$ e $\pi^{(\tilde{p})}$, para indicar, respectivamente, a Lei, a Esperança e a distribuição estacionária do passeio com estado inicial x e probabilidade de transição $\mathbf{p}^{\tilde{p},c}$. Denotaremos $\mathbf{P}_{x,y}^{(\tilde{p})}$ para a medida de probabilidade do acoplamento do passeio com probabilidade catástrofe $1 - \tilde{p}$.

Teorema 3.2.2 (Theorem 3.3 de [3]). *Seja $x, y, t \in \mathbb{Z}_+$, $x < y$. Então*

$$d_t(x, y) \geq \alpha^t \max_{j \in \mathbb{Z}_+} \sum_{k=x}^{y-1} \mathbf{P}_k^{(\tilde{p})}(X_t = j).$$

Para x e y constantes, como $\mathbf{P}_k^{(\tilde{p})}(X_t = j) \rightarrow \pi^{(\tilde{p})}$ quando t cresce, a razão entre essa cota inferior e a cota do Teorema 3.2.1 converge, quando t cresce, para uma constante, ou seja, são da mesma ordem de grandeza no tempo. É notável a constante α , ela aparece em muitos contextos. Vale uma interpretação probabilística sobre ela. É possível argumentar que

$$\mathbf{P}(X_{t+1} > 0 | X_t = 1) = \alpha.$$

Então α seria a probabilidade da população não se extinguir na próxima época, dado que a população possui um indivíduo. Uma outra interpretação, agora no contexto do acoplamento, seria notar

$$\mathbf{P}_{x,y}(H_{t+1} = 1 | H_t = 1) = \alpha,$$

qual seja, a probabilidade de não haver acoplamento no momento $t + 1$ dado que X'_t e X_t se diferem apenas por uma unidade. Para provar o Teorema 3.2.2, usamos quatro lemas.

Lema 3.2.2. *A distribuição de X_t , dado N_t , não depende de p .*

Demonstração. Faremos uma indução em t . Para $t = 1$, perceba que $N_1 = \omega_1$. Então, usando a regra do acoplamento, obtemos

$$\mathbf{P}_{x,y}(X_1 = k | N_1 = 0) = \mathbf{P}_{x,y}(X_1 = k | \omega_1 = 0, X_0 = x, X'_0 = y, H_0 = y - x) = \mathbb{I}_{x+1}(k)$$

e

$$\begin{aligned} \mathbf{P}_{x,y}(X_1 = k | N_1 = 1) &= \mathbf{P}_{x,y}(X_1 = k | \omega_1 = 1, X_0 = x, X'_0 = y, H_0 = y - x) \\ &= \binom{x}{k} (1-c)^k c^{x-k} \mathbb{I}_{\{0, \dots, x\}}(k). \end{aligned}$$

Nos garantindo que X_1 dado N_1 não depende de p . Suponha que $\mathbb{P}_{x,y}(X_t = k | N_t = n)$ não depende de p para todo $n, k \in \mathbb{Z}_+$. Então

$$\begin{aligned} \mathbb{P}_{x,y}(X_{t+1} = k | N_{t+1} = n) &= \mathbb{P}_{x,y}(X_{t+1} = k | N_{t+1} = n, \omega_{t+1} = 1) \mathbb{P}_{x,y}(\omega_{t+1} = 1 | N_{t+1} = n) + \\ &\quad \mathbb{P}_{x,y}(X_{t+1} = k | N_{t+1} = n, \omega_{t+1} = 0) \mathbb{P}_{x,y}(\omega_{t+1} = 0 | N_{t+1} = n). \end{aligned}$$

Cada parcela dessa soma não depende de p . De fato,

$$\{\omega_{t+1} = 1, N_{t+1} = n\} = \{\omega_{t+1} = 1, N_t = n - 1\},$$

$$\begin{aligned} \mathbb{P}_{x,y}(\omega_{t+1} = 1 | N_{t+1} = n) &= \frac{\mathbb{P}_{x,y}(\omega_{t+1} = 1, N_t = n - 1)}{\mathbb{P}_{x,y}(N_{t+1} = n)} \\ &= \frac{(1-p) \binom{t}{n-1} (1-p)^{n-1} p^{t-n+1}}{\binom{t+1}{n} (1-p)^n p^{t+1-n}} = \frac{\binom{t}{n-1}}{\binom{t+1}{n}}. \end{aligned}$$

Além disso, pela regra do acoplamento, segue que

$$\begin{aligned} \{X_{t+1} = k, \omega_{t+1} = 0\} &= \{X_t = k - 1, \omega_{t+1} = 0\}, \\ \{N_{t+1} = n, \omega_{t+1} = 0\} &= \{N_t = n - 1, \omega_{t+1} = 0\}, \end{aligned}$$

e ω_t é independente do passado. Então

$$\begin{aligned} \mathbb{P}_{x,y}(X_{t+1} = k | N_{t+1} = n, \omega_{t+1} = 0) &= \mathbb{P}_{x,y}(X_t = k - 1 | N_t = n, \omega_{t+1} = 0) \\ &= \mathbb{P}_{x,y}(X_t = k - 1 | N_t = n). \end{aligned}$$

Pela hipótese de indução $\mathbb{P}_{x,y}(X_t = k - 1 | N_t = n)$ não depende de p . Por fim,

$$\begin{aligned} \mathbb{P}_{x,y}(X_{t+1} = k | N_{t+1} = n, \omega_{t+1} = 1) &= \mathbb{P}_{x,y}(X_{t+1} = k | N_t = n - 1, \omega_{t+1} = 1) \\ &= \sum_i \binom{i}{k} (1-c)^k c^{i-k} \mathbb{I}_{\{0, \dots, i\}}(k) \mathbb{P}_{x,y}(X_t = i | N_t = n - 1). \end{aligned}$$

E pela hipótese de indução, essa parcela também não depende de p . QED

Lema 3.2.3 (Lemma 3.5 de [3]). *Para todo $x, t \in \mathbb{Z}_+$,*

1. $\mathbb{P}_{x,x+1}(\xi > t) = \alpha^t$;
2. $\mathbb{P}_{x,x+1}(X_t \in \cdot | \xi > t) = \mathbb{P}_x^{(\tilde{p})}(X_t \in \cdot)$.

Demonstração. O primeiro item do lema decorre imediatamente da argumentação do Teorema 3.2.1, pois se $y = x + 1$ não necessitamos utilizar a desigualdade de Bernoulli. Para a segundo item, os autores do artigo [3] afirmaram que condicionado em N_t , ξ e X_t são independentes. Temos

$$\begin{aligned} \mathbb{P}_{x,x+1}(X_t = j, \xi > t | N_t = k) &= \mathbb{P}_{x,x+1}(X_t = j | N_t = k) \mathbb{P}_{x,x+1}(\xi > t | N_t = k) \\ &= \mathbb{P}_{x,x+1}(X_t = j | N_t = k) (1-c)^k. \end{aligned}$$

Além disso,

$$\begin{aligned}
(1-c)^k \mathbb{P}_{x,x+1}(N_t = k) &= \binom{t}{k} ((1-p)(1-c))^k p^{t-k} \\
&= \binom{t}{k} (\alpha - p)^k p^{t-k} \\
&= \alpha^t \binom{t}{k} (1 - \tilde{p})^k \tilde{p}^{t-k} \\
&= \alpha^t \mathbb{P}_{x,x+1}^{(\tilde{p})}(N_t = k).
\end{aligned}$$

Então,

$$\mathbb{P}_{x,x+1}(X_t = j, \xi > t) = \alpha^t \sum_k \mathbb{P}_{x,x+1}(X_t = j | N_t = k) \mathbb{P}_{x,x+1}^{(\tilde{p})}(N_t = k).$$

Usando o Lema 3.2.3 e a validade do acoplamento, obtemos

$$\mathbb{P}_{x,x+1}(X_t = j, \xi > t) = \alpha^t \mathbb{P}_{x,x+1}^{(\tilde{p})}(X_t = j) = \alpha^t \mathbb{P}_x^{(\tilde{p})}(X_t = j).$$

QED

A afirmação de que X_t e ξ são independentes dado N_t é falsa. Veja o seguinte contra-exemplo.

Contra-Exemplo. Uma condição necessária para X_t e ξ serem independentes, dado N_t , seria

$$\mathbb{P}_{x,y}(X_2 = s | \xi \leq 1, N_2 = 1) = \mathbb{P}_{x,y}(X_2 = s | N_2 = 1).$$

Todavia,

$$\mathbb{P}_{x,y}(X_2 = s | \xi \leq 1, N_2 = 1) = \mathbb{P}(\text{Bin}(x, (1-c)) = s - 1).$$

E,

$$\begin{aligned}
\mathbb{P}_{x,y}(X_2 = s | N_2 = 1) &= \mathbb{P}_{x,y}(X_2 = s | N_2 = 1, \omega_1 = 1) \mathbb{P}_{x,y}(\omega_1 = 1 | N_2 = 1) + \\
&\mathbb{P}_{x,y}(X_2 = s | N_2 = 1, \omega_1 = 0, \omega_2 = 1) \mathbb{P}_{x,y}(\omega_2 = 1 | N_2 = 1, \omega_1 = 0) \mathbb{P}_{x,y}(\omega_1 = 0 | N_2 = 1) = \\
&\mathbb{P}(\text{Bin}(x, (1-c)) = s - 1) \frac{1}{2} + \mathbb{P}(\text{Bin}(x+1, (1-c)) = s) \frac{1}{2}.
\end{aligned}$$

Como as probabilidades não coincidem em geral, isso é suficiente para a falsidade da sentença.

Contudo, é possível provar que os eventos $\{X_t = k\}$ e $\{H_i = h\}$ são independentes dado N_t , para todo $k, t, h \in \mathbb{Z}_+$ e $i \geq t$. A demonstração pode ser encontrada no apêndice B. Note que esse resultado garante a argumentação do Lema 3.2.3, pois no argumento foi apenas usado que os eventos $\{X_t = j\}$ e $\{\xi > t\}$ são independentes dado N_t , e isso é válido pois $\{\xi > t\} = \{H_t > 0\}$.

Lema 3.2.4 (Lemma 3.6 de [3]). *Para $j \in \mathbb{Z}_+$, seja $A_j = \{0, \dots, j\}$. Então*

$$\mathbb{P}_x(X_t \in A_j) - \mathbb{P}_{x+1}(X_t \in A_j) = \alpha^t \mathbb{P}_x^{(\tilde{p})}(X_t = j).$$

Demonstração. Para todo t ,

$$\begin{aligned} \mathbb{P}_x(X_t \in A_j) - \mathbb{P}_{x+1}(X_t \in A_j) &= \sum_{i \in A_j} \mathbf{p}^t(x, i) - \sum_{i \in A_j} \mathbf{p}^t(x+1, i) \\ &= \sum_{i \in A_j} \mathbb{P}_{x, x+1}(X_t = j) - \sum_{i \in A_j} \mathbb{P}_{x, x+1}(X'_t = j) \\ &= \mathbb{P}_{x, x+1}(X_t \in A_j) - \mathbb{P}_{x, x+1}(X'_t \in A_j) \\ &= \mathbb{P}_{x, x+1}(X_t \in A_j, \xi > t) + \mathbb{P}_{x, x+1}(X_t \in A_j, \xi \leq t) - \\ &\quad \mathbb{P}_{x, x+1}(X'_t \in A_j, \xi > t) - \mathbb{P}_{x, x+1}(X'_t \in A_j, \xi \leq t) \\ &= \mathbb{P}_{x, x+1}(X_t \in A_j, \xi > t) - \mathbb{P}_{x, x+1}(X'_t \in A_j, \xi > t) \\ &= \mathbb{E}_{x, x+1}(\mathbb{I}_{A_j}(X_t) - \mathbb{I}_{A_j}(X'_t), \xi > t) \\ &= \mathbb{E}_{x, x+1} \left(\sum_{k=0}^j \mathbb{I}_{\{k\}}(X_t) - \sum_{k=0}^j \mathbb{I}_{\{k\}}(X'_t), \xi > t \right) \\ &= \mathbb{E}_{x, x+1} \left(\sum_{k=0}^{j-1} \mathbb{I}_{\{k\}}(X_t) - \sum_{k=1}^j \mathbb{I}_{\{k\}}(X'_t), \xi > t \right) + \\ &\quad \mathbb{E}_{x, x+1}(\mathbb{I}_{\{j\}}(X_t), \xi > t) - \mathbb{E}_{x, x+1}(\mathbb{I}_{\{0\}}(X'_t), \xi > t) \\ &= \mathbb{E}_{x, x+1} \left(\sum_{k=0}^{j-1} (\mathbb{I}_{\{k\}}(X_t) - \mathbb{I}_{\{k+1\}}(X'_t)), \xi > t \right) + \\ &\quad \mathbb{P}_{x, x+1}(X_t = j, \xi > t) - \mathbb{P}_{x, x+1}(X'_t = 0, \xi > t). \end{aligned}$$

Sob $\xi > t$ temos $X'_t = X_t + 1$ e então, $\mathbb{I}_{\{k+1\}}(X'_t) = \mathbb{I}_{\{k\}}(X_t)$, e $X'_t = 0 \implies \xi \leq t$, logo $\{X'_t = 0, \xi > t\} = \emptyset$. Portanto, usando o resultado provado na prova do Lema 3.2.3 segue

$$\mathbb{P}_x(X_t \in A_j) - \mathbb{P}_{x+1}(X_t \in A_j) = \mathbb{P}_{x, x+1}(X_t = j, \xi > t) = \alpha^t \mathbb{P}_x^{(\tilde{p})}(X_t = j).$$

QED

Prova do Teorema 3.2.2. Para todo j , somando zero e usando o Lema 3.2.4, temos

$$\begin{aligned} d_t(x, y) &\geq \mathbb{P}_x(X_t \in A_j) - \mathbb{P}_y(X_t \in A_j) \\ &= \sum_{k=x}^{y-1} (\mathbb{P}_k(X_t \in A_j) - \mathbb{P}_{k+1}(X_t \in A_j)) \\ &= \alpha^t \sum_{k=x}^{y-1} \mathbb{P}_k^{(\tilde{p})}(X_t = j). \end{aligned}$$

Como vale para todo j , o resultado é válido.

QED

Para fechar o capítulo, apresentamos um resultado que generaliza o item 2 do Lema 3.2.3.

Teorema 3.2.3 (Theorem 3.7 de [3]). $\mathbb{P}_{x,y}(X_t \in \cdot | \xi > t)$ converge em distribuição para $\pi^{(\tilde{p})}$ quando $t \rightarrow \infty$.

Demonstração. Argumentaremos que

$$\mathbb{P}_{x,y}(X_t = j | H_t = 1) \rightarrow \pi^{(\tilde{p})}(j) \quad e \quad \frac{\mathbb{P}_{x,y}(H_t = k)}{\mathbb{P}_{x,y}(H_t = 1)} \rightarrow 0,$$

para $k \geq 2$. Isso será suficiente, pois

$$\begin{aligned} \mathbb{P}_{x,y}(X_t = j | \xi > t) &= \frac{\mathbb{P}_{x,y}(X_t = j, \xi > t)}{\mathbb{P}_{x,y}(\xi > t)} \\ &= \frac{\sum_{k=1}^{y-x} \mathbb{P}_{x,y}(H_t = k, X_t = j)}{\sum_{k=1}^{y-x} \mathbb{P}_{x,y}(H_t = k)} \\ &= \frac{\mathbb{P}_{x,y}(X_t = j | H_t = 1)}{1 + \sum_{k=2}^{y-x} \left[\frac{\mathbb{P}_{x,y}(H_t = k)}{\mathbb{P}_{x,y}(H_t = 1)} \right]} + \frac{\sum_{k=2}^{y-x} \frac{\mathbb{P}_{x,y}(X_t = j, H_t = k)}{\mathbb{P}_{x,y}(H_t = 1)}}{1 + \sum_{k=2}^{y-x} \left[\frac{\mathbb{P}_{x,y}(H_t = k)}{\mathbb{P}_{x,y}(H_t = 1)} \right]}. \end{aligned}$$

Denotando $M = y - x$ e $\theta = 1 - c$. Temos,

$$\begin{aligned} \mathbb{P}_{x,y}(H_t = k) &= \mathbf{E}_{x,y}(\mathbf{P}_{x,y}(H_t = k | N_t)) \\ &= \mathbf{E}_{x,y}\left(\binom{y-x}{k} (1-c)^{N_t k} (1 - (1-c)^{N_t})^{y-x-k}\right) \\ &= \binom{M}{k} \mathbf{E}_{x,y}(\theta^{N_t k} (1 - \theta^{N_t})^{M-k}). \end{aligned}$$

Também,

$$\begin{aligned} \mathbf{E}_{x,y}(\theta^{N_t k} (1 - \theta^{N_t})^{M-k}) &= \mathbf{E}_{x,y}\left[\theta^{N_t k} \sum_{i=0}^{M-k} (-1)^i \binom{M-k}{i} \theta^{N_t i}\right] \\ &= \sum_{i=0}^{M-k} (-1)^i \binom{M-k}{i} \mathbf{E}_{x,y}[\theta^{(k+i)N_t}] \\ &= (p + (1-p)\theta^k)^t + \sum_{i=1}^{M-k} \binom{M-k}{i} (-1)^i (p + (1-p)\theta^{i+k})^t \\ &= (p + (1-p)\theta^k)^t \left[1 + \sum_{i=1}^{M-k} \binom{M-k}{i} (-1)^i \left(\frac{p + (1-p)\theta^{i+k}}{p + (1-p)\theta^k}\right)^t\right] \\ &= (p + (1-p)\theta^k)^t (1 + o(1)). \end{aligned}$$

Além disso, para $k \geq 2$

$$\frac{\mathbb{P}_{x,y}(H_t = k)}{\mathbb{P}_{x,y}(H_t = 1)} = \frac{(p + (1-p)\theta^k)^t}{(p + (1-p)\theta)^t} \left(\frac{1}{1 + o(1)} + \frac{o(1)}{1 + o(1)} \right) \xrightarrow{t \rightarrow \infty} 0.$$

Para a outra afirmação, usando o Lema 3.2.2, o argumento da demonstração do Lema 3.2.3 e a relação anteriormente expressa entre N_t , H_t e X_t , temos

$$\begin{aligned}
\mathbb{P}_{x,y}(X_t = j, H_t = 1) &= \sum_{k=0}^{\infty} \mathbb{P}_{x,y}(X_t = j | N_t = k) \mathbb{P}_{x,y}(H_t = 1 | N_t = k) \mathbb{P}_{x,y}(N_t = k) \\
&= \sum_{k=0}^{\infty} \mathbb{P}_x^{(\tilde{p})}(X_t = j | N_t = k) (y-x) (1 - (1-c)^k)^{y-x-1} (1-c)^k \mathbb{P}_{x,y}(N_t = k) \\
&= (y-x) \alpha^t \mathbb{E}_x^{(\tilde{p})} [\mathbb{I}_{\{j\}}(X_t) (1 - (1-c)^{N_t})^{x-y-1}] \\
&= (y-x) \alpha^t \mathbb{P}_x^{(\tilde{p})}(X_t = j) (1 + o(1)).
\end{aligned}$$

A quarta igualdade é válida porque

$$\begin{aligned}
&\mathbb{E}_x^{(\tilde{p})} (\mathbb{I}_{\{j\}}(X_t) (1 - (1-c)^{N_t})^{y-x-1}) \\
&= \mathbb{E}_x^{(\tilde{p})} \left(\mathbb{I}_{\{j\}}(X_t) \left[\sum_{k=0}^{y-x-1} \binom{y-x-1}{k} (-1)^k (1-c)^{N_t k} \right] \right) \\
&= \mathbb{P}_x^{(\tilde{p})}(X_t = j) + \sum_{k=1}^{y-x-1} \mathbb{E}_x^{(\tilde{p})} ((1-c)^{N_t k} \mathbb{I}_{\{j\}}(X_t)) (-1)^k \binom{y-x-1}{k} \\
&= \mathbb{P}_x^{(\tilde{p})}(X_t = j) (1 + o(1)).
\end{aligned}$$

A terceira igualdade é verdadeira visto que para todo k

$$\frac{\mathbb{E}_x^{(\tilde{p})} ((1-c)^{N_t k} \mathbb{I}_{\{j\}}(X_t))}{\mathbb{P}_x^{(\tilde{p})}(X_t = j)} \leq \frac{\mathbb{E}_x^{(\tilde{p})} ((1-c)^{N_t k})}{\mathbb{P}_x^{(\tilde{p})}(X_t = j)} = \frac{(\tilde{p} + (1-\tilde{p})(1-c)^k)^t}{\mathbb{P}_x^{(\tilde{p})}(X_t = j)} \xrightarrow{t \rightarrow \infty} 0.$$

Já que $(X_t)_{t \geq 0}$ é recorrente positiva, e $\mathbb{P}_x^{(\tilde{p})}(X_t = j) \rightarrow \pi^{(\tilde{p})}(j)$, onde $0 < \pi^{(\tilde{p})}(j) < 1$. Tudo isso nos possibilita concluir que $\mathbb{P}_{x,y}(X_t = j | H_t = 1) = \mathbb{P}_x^{(\tilde{p})}(X_t = j) + o(1)$. QED

Capítulo 4

O fenômeno de cutoff

Neste capítulo apresentamos o principal resultado do artigo referência, isto é, que a classe de modelos da Equação 2.2 apresenta cutoff num sentido que será explicado. Mas antes mostraremos um comportamento interessante da distribuição estacionária quando os parâmetros p e c estão próximos de zero, esse resultado será importante para provar o resultado sobre cutoff.

Para uma sequência de cadeias de Markov $((X_t^{(n)})_{t \in \mathbb{Z}_+})_{n \in \mathbb{N}}$, onde $(X_t^{(n)})_{t \in \mathbb{Z}_+}$ é a cadeia da Equação 2.2 com núcleo de transição \mathbf{p}^{p_n, c_n} , denotamos, $\pi^{(n)}$, $d_t^{(n)}$, $\mathbf{P}_x^{(n)}$, para designar, respectivamente, a distribuição estacionária, a distância em variação total no momento t e a lei do processo partindo de $x \in \mathbb{Z}_+$, da n -ésima cadeia dessa sequência.

4.1 Limite Poisson

Teorema 4.1.1 (Theorem 4.2 de [3]). *Seja $p_n, c_n \in (0, 1)$, $p_n \rightarrow 0$ e $\frac{p_n}{c_n} = \beta \in (0, \infty)$. Então $\pi^{(n)}$ converge em distribuição para $\text{Pois}(\beta)$ quando $n \rightarrow \infty$.*

Demonstração. A prova desse teorema, no artigo, foi feita parcialmente, para $t > 0$. Eles usaram a função geradora de momentos. O argumento apresentado aqui é válido $t \in \mathbb{R}$, e para isso lançaremos mão da seguinte desigualdade.

$$0 \leq x - \ln(x + 1) \leq 2x^2, \quad |x| < 0,93. \quad (4.1)$$

Para t fixado, temos

$$\begin{aligned}
\Lambda^{(n)}(t) &= \ln \left[\mathbf{E} \left(\exp \left\{ t \sum_{j=0}^{\infty} \text{Bin}_j(R_j^{(n)}, (1-c_n)^j) \right\} \right) \right] \\
&= \ln \left[\mathbf{E} \left(\prod_{j=0}^{\infty} \exp \left\{ t \text{Bin}_j(R_j^{(n)}, (1-c_n)^j) \right\} \right) \right] \\
&= \ln \left[\mathbf{E} \left(\prod_{j=0}^{\infty} (1 - (1-c_n)^j + (1-c_n)^j \exp\{t\})^{R_j^{(n)}} \right) \right] \\
&= \sum_{j=0}^{\infty} \ln \left[\mathbf{E} \left((1 - (1-c_n)^j + (1-c_n)^j \exp\{t\})^{R_j^{(n)}} \right) \right].
\end{aligned}$$

Mas veja que para todo j , $R_j^{(n)} \stackrel{d}{=} \text{Geom}^-(1-p_n)$ e

$$p_n(1 - (1-c_n)^j + (1-c_n)^j \exp\{t\}) < 1 \Leftrightarrow \exp\{t\} < 1 + \frac{1-p_n}{p_n(1-c_n)^j}.$$

Vendo que $j_1 \geq j_2 \implies \frac{1}{(1-c_n)^{j_2}} \leq \frac{1}{(1-c_n)^{j_1}}$, notamos que para todo n suficientemente grande, $\exp\{t\} < 1 + \frac{1-p_n}{p_n(1-c_n)^j}, \forall j$. Assim, para todo $j \in \mathbb{N}$ e n grande o suficiente, teremos

$$\begin{aligned}
\mathbf{E} \left((1 - (1-c_n)^j + (1-c_n)^j \exp\{t\})^{R_j^{(n)}} \right) &= \sum_{k=0}^{\infty} (1-p_n)(p_n(1 - (1-c_n)^j + (1-c_n)^j \exp\{t\}))^k \\
&= \frac{1-p_n}{1-p_n(1 - (1-c_n)^j + (1-c_n)^j \exp\{t\})}.
\end{aligned}$$

Assim, a seguinte soma está bem definida para todo n suficientemente grande,

$$\begin{aligned}
\Lambda^{(n)}(t) &= \sum_{j=0}^{\infty} \ln \left\{ \frac{1-p_n}{1-p_n(1 - (1-c_n)^j(1 - \exp\{t\}))} \right\} \\
&= - \sum_{j=0}^{\infty} \ln \left\{ \frac{1-p_n + p_n(1-c_n)^j(1 - \exp\{t\})}{1-p_n} \right\} \\
&= - \sum_{j=0}^{\infty} \ln \left\{ 1 + \frac{p_n(1-c_n)^j}{1-p_n}(1 - \exp\{t\}) \right\}
\end{aligned}$$

Agora usaremos a desigualdade da Equação 4.1, com $x = \frac{p_n(1-c_n)^j}{1-p_n}(1 - \exp\{t\})$. Observando que para todo t e j ,

$$\left| \frac{p_n(1-c_n)^j}{1-p_n}(1 - \exp\{t\}) \right| < \left| \frac{p_n}{1-p_n}(1 - \exp\{t\}) \right|.$$

Então, como t é fixo, para todo n suficientemente grande

$$\left| \frac{p_n(1-c_n)^j}{1-p_n}(1 - \exp\{t\}) \right| < 0.93, \quad \forall j.$$

Ou seja, para todo j ,

$$\begin{aligned}
0 &\leq \frac{p_n(1-c_n)^j}{1-p_n}(1-\exp\{t\}) - \ln\left\{1 + \frac{p_n(1-c_n)^j}{1-p_n}(1-\exp\{t\})\right\} \leq 2\left(\frac{p_n(1-c_n)^j}{1-p_n}\right)^2 (1-\exp\{t\})^2 \\
\implies 0 &\leq \sum_{j=0}^{\infty} \frac{p_n(1-c_n)^j}{1-p_n}(1-\exp\{t\}) + \Lambda^{(n)}(t) \leq 2\left(\frac{p_n(1-\exp\{t\})^2}{1-p_n}\right)^2 \sum_{j=0}^{\infty} (1-c_n)^{2j}.
\end{aligned}$$

Vendo que,

$$\frac{p_n}{1-p_n}(1-\exp\{t\}) \sum_{j=0}^{\infty} (1-c_n)^j = \frac{p_n}{(1-p_n)c_n}(1-\exp\{t\}) \xrightarrow{n \rightarrow \infty} \beta(1-\exp\{t\})$$

e

$$\begin{aligned}
2\left(\frac{p_n(1-\exp\{t\})^2}{1-p_n}\right)^2 \sum_{j=0}^{\infty} (1-c_n)^{2j} &\leq 2\left(\frac{p_n(1-\exp\{t\})^2}{1-p_n}\right)^2 \sum_{j=0}^{\infty} (1-c_n)^j \\
&= \frac{2p_n}{(1-p_n)^2} \frac{p_n}{c_n} (1-\exp\{t\})^2 \xrightarrow{n \rightarrow \infty} 0,
\end{aligned}$$

concluimos que

$$\lim_{n \rightarrow \infty} \ln \left[\mathbb{E} \left(\exp \left\{ t \sum_{j=0}^{\infty} \text{Bin}_j(R_j^{(n)}, (1-c_n)^j) \right\} \right) \right] = -\beta(1-\exp\{t\}), \quad \forall t \in \mathbb{R}.$$

QED

4.2 Cutoff

4.2.1 Definição de Cutoff

Para apresentar a definição desse fenômeno e algumas consequências, precisamos definir outros objetos. Gostaríamos antes de pontuar que a distância em variação total, tal como na Definição 2.2.2, é uma *métrica*. no sentido usual de análise, para o espaço das medidas de probabilidade. Então é válido, por exemplo, a desigualdade triangular.

Definição 4.2.1. Seja $(Y_t)_{t \in \mathbb{Z}_+}$, uma cadeia de Markov tal como na Definição 2.1.1. Seja ela irredutível, aperiódica e recorrente positiva, com distribuição estacionária η . Denotamos a *função variação total* por $d(t) = \sup_{x \in \mathcal{X}} \|\mathbb{P}^t(x, \cdot) - \eta\|_{VT}$.

A função $d(t)$ é não crescente em t , pois

$$\|\mathbb{P}^{t+1}(x, \cdot) - \eta\|_{VT} \leq \|\mathbb{P}^t(x, \cdot) - \eta\|_{VT} \quad \forall x \in \mathcal{X},$$

veja [17, Capítulo 4].

Definição 4.2.2. O ϵ -tempo de mistura é

$$t_{mix}(\epsilon) = \inf\{t \in \mathbb{N} : d(t) \leq \epsilon\}.$$

Denotaremos t_{mix} para $\frac{1}{4}$ -tempo de mistura.

É imediato que $\delta \geq \epsilon \implies t_{mix}(\delta) \leq t_{mix}(\epsilon)$. O fenômeno de cutoff é definido numa sequência de cadeias de Markov, e usaremos n para denotar a grandeza da n -ésima cadeia de Markov da sequência. Por exemplo, $t_{mix}^{(n)}(\epsilon)$ será o ϵ -tempo de mistura da n -ésima cadeia de Markov, d_n a função variação total da n -ésima, etc. A definição a seguir foi retirada de [17, Chapter 18].

Definição 4.2.3. Uma sequência de cadeias de Markov, $((Y_t^{(n)})_{t \in \mathbb{Z}_+})_{n \in \mathbb{N}}$, têm *cutoff com janela* de tamanho $O(w_n)$ se $w_n = o(t_{mix}^{(n)})$,

$$\begin{aligned} \lim_{\alpha \rightarrow -\infty} \liminf_{n \rightarrow \infty} d_n(t_{mix}^{(n)} + \alpha w_n) &= 1, \\ \lim_{\alpha \rightarrow \infty} \limsup_{n \rightarrow \infty} d_n(t_{mix}^{(n)} + \alpha w_n) &= 0. \end{aligned}$$

Pode ser um pouco difícil para entender o que essa definição está dizendo. Mas usando o primeiro limite da Definição 4.2.3, é possível deduzir, pela monotonicidade de $d(t)$, que

$$\forall \epsilon > 0, \exists M(\epsilon) < 0 \text{ e } k(\epsilon) \in \mathbb{N} \text{ tal que } \alpha \leq M(\epsilon) \text{ e } k \geq k(\epsilon) \implies d_k(t_{mix}^{(k)} + \alpha w_k) \geq 1 - \epsilon.$$

Com isso, para todo $0 < C < 1$,

$$d_n(C t_{mix}^{(n)}) = d_n(t_{mix}^{(n)} - \frac{(1-C)t_{mix}^{(n)}}{w_n} w_n) \rightarrow 1,$$

quando $n \rightarrow \infty$, pois $w_n \in o(t_{mix}^{(n)})$. De um argumento análogo, mas usando o segundo limite da Definição 4.2.3, podemos deduzir que

$$\lim_{n \rightarrow \infty} d_n(C t_{mix}^{(n)}) = \begin{cases} 1, & \text{se } C < 1, \\ 0, & \text{se } C > 1. \end{cases} \quad (4.2)$$

Ou seja, para a cadeia $(Y_t^{(n)})_{t \in \mathbb{Z}_+}$ com n grande, a sua função variação total vista na escala de tempo criada por $t_{mix}^{(n)}$ se comporta como ilustrado na Figura 4.2.1. Essa queda repentina exemplificada dá o nome ao fenômeno: *cutoff*. O $\frac{1}{4}$ -tempo de mistura não possui nada de especial, poderíamos usar qualquer ϵ -tempo de mistura na Definição 4.2.3. Também podemos provar que

$$\forall \epsilon \in (0, 1), \quad \lim_{n \rightarrow \infty} \frac{t_{mix}^{(n)}(1 - \epsilon)}{t_{mix}^{(n)}(\epsilon)} = 1. \quad (4.3)$$

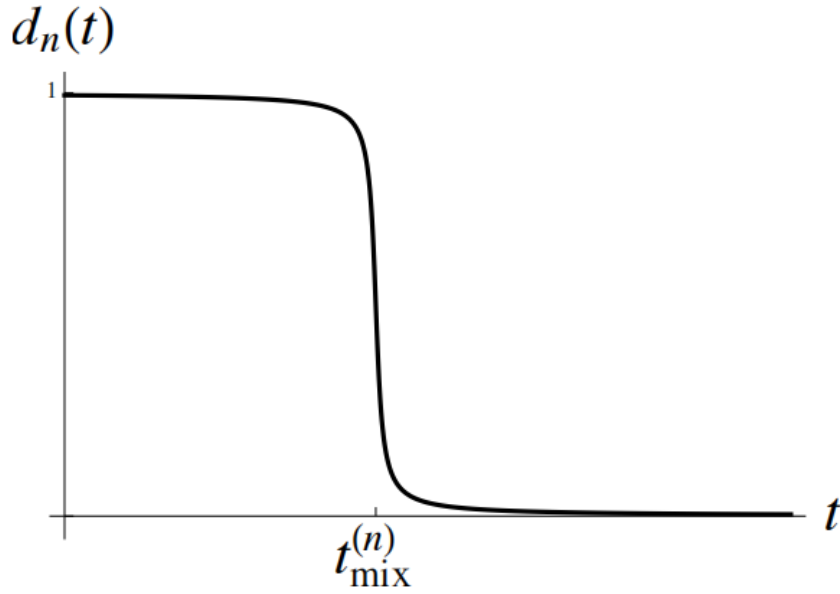


Figura 4.1: Ilustração cutoff.

Não provaremos esses resultados pois as demonstrações são essencialmente iguais aos argumentos que apresentaremos na Proposição 4.2.1, para provar um resultado um pouco mais geral.

Portanto a Definição 4.2.3 está dizendo que a lei $\mathbb{P}_x(Y_t^{(n)} \in \cdot)$ se aproxima em variação total da sua distribuição estacionária $\eta^{(n)}$, para n grande, numa janela de tempo cuja a magnitude, se comparada com a ordem dos tempos de mistura, é irrisória. Para fechar essa seção, apresentamos um resultado muito útil.

Proposição 4.2.1. *Seja $((Y_t^{(n)})_{t \in \mathbb{Z}_+})_{n \in \mathbb{N}}$ uma sequência de cadeias de Markov. Se existem sequências $(f(n))_{n \in \mathbb{N}}$ e $(\nu_n)_{n \in \mathbb{N}}$ tais quais $\nu_n = o(f(n))$ e*

$$\begin{aligned} \lim_{\alpha \rightarrow -\infty} \liminf_{n \rightarrow \infty} d_n(f(n) + \alpha \nu_n) &= 1 \\ \lim_{\alpha \rightarrow \infty} \limsup_{n \rightarrow \infty} d_n(f(n) + \alpha \nu_n) &= 0. \end{aligned} \quad (4.4)$$

Então existe uma sequência $(w_n)_{n \in \mathbb{N}}$ tal que a Definição 4.2.3 é satisfeita, e

$$\lim_{n \rightarrow \infty} \frac{f(n)}{t_{mix}^{(n)}} = 1.$$

Demonstração. Primeiro mostremos que $\frac{f(n)}{t_{mix}^{(n)}} \rightarrow 1$. Usando a definição de limite e a monotonicidade de $d_n(t)$, podemos deduzir, da Equação 4.4, as seguintes sentenças.

$$\forall \epsilon > 0, \exists M(\epsilon) < 0 \text{ e } k(\epsilon) \in \mathbb{N} \text{ tal que } \alpha \leq M(\epsilon) \text{ e } k \geq k(\epsilon) \implies d_k(f(k) + \alpha \nu_k) \geq 1 - \epsilon.$$

$$\forall \epsilon > 0, \exists N(\epsilon) > 0 \text{ e } m(\epsilon) \in \mathbb{N} \text{ tal que } \alpha \geq N(\epsilon) \text{ e } k \geq m(\epsilon) \implies d_k(f(k) + \alpha \nu_k) \leq \epsilon.$$

Fixe $\beta \geq N(\frac{1}{4})$ e $\alpha \leq M(\frac{1}{4})$ Então para todo $k \geq \max\{m(\frac{1}{4}), k(\frac{1}{4})\}$

$$f(k) + \beta\nu_k \leq t_{mix}^{(k)} \leq f(k) + \alpha\nu_k,$$

garantindo o resultado.

Mostraremos apenas o primeiro limite da Definição 4.2.3, o argumento para o segundo é similar. Seja $w_n = \nu_n + |f(n) - t_{mix}^{(n)}|$. Dado $\epsilon > 0$ faça $k \geq k(\epsilon)$ e $\alpha \leq M(\epsilon) < 0$ tal que $d_k(f(k) + \alpha\nu_k) \geq 1 - \epsilon$.

Seja $t_{mix}^{(k)} < f(k)$. Pela monotonicidade e por α ser negativo, obtemos

$$d_k(t_{mix}^{(k)} + \alpha w_k) \geq d_k(f(k) + \alpha w_k) \geq d_k(f(k) + \alpha\nu_k) \geq 1 - \epsilon.$$

Se $t_{mix}^{(k)} > f(k)$, então $|f(k) - t_{mix}^{(k)}| = t_{mix}^{(k)} - f(k)$ e, como α pode ser arbitrariamente pequeno,

$$d_k(t_{mix}^{(k)} + \alpha w_k) \geq d_k(f(k) + \alpha\nu_k + (\alpha + 1)(t_{mix}^{(k)} - f(k))) \geq d_k(f(k) + \alpha\nu_k) \geq 1 - \epsilon.$$

QED

Fizemos questão de apresentar esse resultado pois na prática usa-se essa proposição, visto que em muitos contextos é complicado obter exatamente os tempos de misturas. Isso é o que ocorre no modelo que estamos estudando, note que temos cotas apenas para $d_t(x, \pi)$, não para $d(t) = \sup_{x \in \mathbb{Z}_+} d_t(x, \pi)$. Desse modo, não conseguimos ter ideia da grandeza dos ϵ -tempos de mistura do nosso modelo. Quando se prova que uma sequência de cadeias de Markov satisfaz os limites da Proposição 4.2.1, dizemos que a sequência têm cutoff em $f(n)$ com tamanho de janela $O(\nu_n)$.

4.2.2 Cutoff do modelo

Nesta seção apresentamos o principal resultado do artigo referência.

Teorema 4.2.1 (Theorem 4.3 de [3]). *Para $n \in \mathbb{Z}_+$, $p_n, c_n \in (0, 1)$ com $p_n \rightarrow 0$ e $\lim_{n \rightarrow \infty} \frac{p_n}{c_n} = \beta \in (0, \infty)$. Seja $(y_n)_{n \in \mathbb{Z}_+}$ uma sequência de números reais satisfazendo $\lim_{n \rightarrow \infty} y_n = \infty$. Faça*

$$t_n = \frac{\ln y_n}{c_n}.$$

Então, para todo $\epsilon > 0$:

1.

$$\lim_{\epsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \sup \left\{ d_t^{(n)}(y_n, \pi^{(n)}); t > t_n + \frac{1}{\epsilon c_n} \right\} = 0.$$

2.

$$\lim_{n \rightarrow \infty} \inf \left\{ d_t^{(n)}(y_n, \pi^{(n)}); t < t_n - b_n \right\} = 1,$$

onde

$$b_n = (1 + \epsilon) \left(\frac{1}{2} \ln y_n + \frac{\ln \ln y_n}{c_n} \right).$$

Segundo os autores do artigo [3], a sequência de cadeias de Markov, com a n -ésima cadeia sendo aquela com probabilidade de transição \mathbf{p}^{p_n, c_n} , c_n e p_n tal como no enunciado do Teorema 4.2.1, apresenta cutoff em t_n com janela de tamanho

$$O(\max(\ln(y_n), \frac{\ln(\ln(y_n))}{c_n})).$$

Perceba que a sentença do Teorema 4.2.1 satisfaz a Proposição 4.2.1, com os parâmetros enunciados. É notório a forma intrincada do Teorema 4.2.1, isso é devido ao fato de termos cotas para a distância em variação total apenas para estados iniciais fixos, assim dificultando a demonstração do fenômeno de cutoff, que é definido para a função variação total. A demonstração do Teorema 4.2.1 é feito através de lemas e corolários. Segue o lema.

Lema 4.2.1 (Lemma 4.4 de [3]). *Assuma as condições do Teorema 4.2.1 válidas. Para $\theta > 0$, seja*

$$\lambda_n(\theta) := \frac{\ln y_n + \theta}{c_n} \text{ e } \nu_n(\theta) := \frac{\ln y_n - \ln \ln y_n - \ln \frac{p_n}{c_n} - \frac{\theta}{(\ln y_n)^{1/4}}}{-\ln(1 - c_n)}, \quad n \in \mathbb{N}.$$

Então,

1. $\lim_{\theta \rightarrow \infty} \left\{ \limsup_{n \rightarrow \infty} \sup_{t > \lambda_n(\theta)} d_t^{(n)}(0, y_n) \right\} = 0.$
2. $\lim_{\theta \rightarrow \infty} \left\{ \liminf_{n \rightarrow \infty} \inf_{t < \nu_n(\theta)} d_t^{(n)}(0, y_n) \right\} = 1.$

Demonstração. Denote $\alpha_n = 1 - c_n(1 - p_n)$. Para o primeiro item, usando o Teorema 3.2.1, para todo $t > \lambda_n(\theta)$

$$\begin{aligned} \ln d_t^{(n)}(0, y_n) &\leq \ln y_n + t \ln \alpha_n \\ &\leq \ln y_n + \lambda_n(\theta) \ln \alpha_n \\ &= \ln y_n + (\ln y_n + \theta) \frac{\ln(1 - c_n(1 - p_n))}{c_n(1 - p_n)} (1 - p_n) \\ &\leq \ln y_n - (\ln y_n + \theta) = -\theta. \end{aligned}$$

Tomando o $\theta \rightarrow \infty$ e usando a continuidade do logaritmo, concluímos o primeiro ítem.

A demonstração do segundo item é maior. Usaremos as conhecidas desigualdades de Chernoff-Hoeffding para a distribuição binomial. Se $X \stackrel{d}{=} \text{Bin}(m, p)$ para algum $m \in \mathbb{N}$ e $p \in (0, 1)$, então para todo $\delta \in (0, 1)$,

$$\mathbb{P}(X > (1 + \delta)mp) \leq \exp\left\{-\frac{\delta^2 pm}{2}\right\} \text{ e } \mathbb{P}(X < (1 - \delta)mp) \leq \exp\left\{-\frac{\delta^2 pm}{3}\right\}. \quad (4.5)$$

Para um estudo detalhado sobre esse tipo de desigualdade veja [11].

Continuando, observe que sob $\mathbb{P}_0^{(n)}$, X_t é *estocasticamente dominado* pela variável $\text{Bin}(t, p_n)$, isto é, $\mathbb{P}_0^{(n)}(X_t \geq a) \leq \mathbb{P}(\text{Bin}(t, p_n) \geq a)$. Para entender essa relação basta considerar um acoplamento similar ao da Definição 3.2.1, onde X_t e X'_t começam em 0, mas na catástrofe X'_t se mantém constante. Definamos,

$$\gamma_n(\theta) = \left(1 + \frac{\theta}{2(\ln y_n)^{1/4}}\right) p_n \nu_n(\theta).$$

Por conveniência de notação, denotaremos ν_n e γ_n ao invés de $\nu_n(\theta)$ e $\gamma_n(\theta)$. Pela Equação 4.5, para qualquer $t \leq \nu_n$,

$$\mathbb{P}_0^{(n)}(X_t \geq \gamma_n) \leq \mathbb{P}(\text{Bin}(t, p_n) \geq \gamma_n) \leq \mathbb{P}(\text{Bin}(\nu_n, p_n) \geq \gamma_n) \leq \exp\left[-\frac{\theta^2 p_n \nu_n}{8\sqrt{\ln y_n}}\right].$$

Por outro lado, sob $\mathbb{P}_{y_n}^{(n)}$, X_t domina estocasticamente $\text{Bin}(y_n, (1 - c)^{N_t})$. Isso é verdade pois na regra do acoplamento $H_t + X_t = X'_t$ com $x = 0$ e $y = y_n$, em toda trajetória. Em particular, domina $\text{Bin}(y_n, (1 - c)^t)$. Note que

$$\begin{aligned} y_n(1 - c_n)^{\nu_n} &= \exp\{\ln y_n + \nu_n \ln(1 - c_n)\} \\ &= \exp\left\{\ln y_n + \frac{(\ln y_n - \ln \ln y_n - \ln \frac{p_n}{c_n} - \frac{\theta}{(\ln y_n)^{1/4}})}{-\ln(1 - c_n)} \ln(1 - c_n)\right\} \\ &= \exp\left\{\ln \ln y_n + \ln \frac{p_n}{c_n} + \frac{\theta}{(\ln y_n)^{1/4}}\right\} \\ &= \ln(y_n) \left(\frac{p_n}{c_n}\right) \exp\left[\frac{\theta}{(\ln y_n)^{1/4}}\right]. \end{aligned} \quad (4.6)$$

Como $t \leq \nu_n$, para todo n suficientemente grande,

$$\begin{aligned}
\frac{\gamma_n}{y_n(1-c)^t} &\leq \frac{\left(1 + \frac{\theta}{2(\ln y_n)^{1/4}}\right) p_n \nu_n}{y_n(1-c)^{\nu_n}} = \frac{\left(1 + \frac{\theta}{2(\ln y_n)^{1/4}}\right) c_n \nu_n}{\ln(y_n) \exp\left[\frac{\theta}{(\ln y_n)^{1/4}}\right]} \\
&\leq \frac{\left(1 + \frac{\theta}{2(\ln y_n)^{1/4}}\right) c_n \left(1 - \frac{\ln \ln y_n}{\ln(y_n)} - \ln \frac{p_n}{c_n} / \ln(y_n) - \frac{\theta}{(\ln y_n)^{5/4}}\right)}{\exp\left[\frac{\theta}{(\ln y_n)^{1/4}}\right] - \ln(1-c_n)} \\
&\leq \frac{\left(1 + \frac{\theta}{2(\ln y_n)^{1/4}}\right)}{\exp\left[\frac{\theta}{(\ln y_n)^{1/4}}\right]} \leq \left(1 + \frac{\theta}{2(\ln y_n)^{1/4}}\right) \left(1 - \frac{\theta}{2(\ln y_n)^{1/4}}\right) \\
&= 1 - \frac{\theta^2}{4(\ln y_n)^{1/2}}. \tag{4.7}
\end{aligned}$$

Na quarta desigualdade usamos $\exp\{-x\} \leq 1 - \frac{x}{2}$, $1,5 > x > 0$. Como $t \leq \nu_n$, e pelas Equações 4.5 e 4.7, temos

$$\mathbf{P}_{y_n}^{(n)}(X_t \leq \gamma_n) \leq \mathbf{P}(\text{Bin}(y_n, (1-c)^{\nu_n}) \leq \gamma_n) \leq \exp\left[-\frac{y_n(1-c_n)^{\nu_n} \theta^4}{48 \ln y_n}\right].$$

Assim temos que pela Equação 4.6,

$$\limsup_{n \rightarrow \infty} \sup_{t \leq \nu_n} \mathbf{P}_{y_n}^{(n)}(X_t \leq \gamma_n) \leq \exp\left\{-\frac{\beta \theta^4}{48}\right\}.$$

Ou seja,

$$\begin{aligned}
\liminf_{n \rightarrow \infty} \left(\inf_{t \leq \nu_n(\theta)} d_t^{(n)}(0, y_n) \right) &\geq \liminf_{n \rightarrow \infty} \left(1 - \exp\left[-\frac{y_n(1-c_n)^{\nu_n} \theta^4}{48 \ln y_n}\right] - \exp\left[-\frac{\theta^2 p_n \nu_n}{8\sqrt{\ln y_n}}\right] \right) \\
&\geq 1 - \exp\left\{-\frac{\beta \theta^4}{48}\right\}.
\end{aligned}$$

QED

Corolário 4.2.1 (Corollary 4.5 [3]). *Sob as hipóteses do Teorema 4.2.1,*

1. $\lim_{\epsilon \rightarrow \infty} \limsup_{n \rightarrow \infty} \sup_{t \geq t_n \frac{1}{\epsilon c_n}} d_t^{(n)}(0, y_n) = 0$.
2. *Para todo* $\epsilon > 0$, $\lim_{n \rightarrow \infty} \inf_{t \leq t_n - b_n} d_t^{(n)}(0, y_n) = 1$.

Demonstração. A prova desse corolário se resume em provar que o item dois do Lema 4.2.1 implica no item dois desse corolário. Para isso note,

$$\nu_n(\theta) \geq \frac{\ln y_n - \ln \ln y_n - \theta}{-\ln(1-c_n)} \iff -\theta \leq \frac{-\ln \frac{p_n}{c_n}}{1 - \frac{1}{(\ln y_n)^{1/4}}}.$$

Como $-\ln \frac{p_n}{c_n} \rightarrow -\ln \beta$ e $1 - \frac{1}{(\ln y_n)^{1/4}} \rightarrow 1$, $\exists n_0 \in \mathbb{N}$ tal que, fazendo θ bem grande, as desigualdades enunciadas valem para todo $n \geq n_0$. Usando o teorema da convergência monótona e resultados de cálculo diferenciável e integrável, pode-se mostrar

$$\sum_{i=0}^{\infty} x^i = \frac{1}{1-x} \implies -\ln(1-x) = \sum_{i=0}^{\infty} \frac{x^{i+1}}{i+1}, \quad \forall x \in (0, 1).$$

Além disso, para todo n suficientemente grande,

$$\begin{aligned} \frac{-1}{\ln(1-c_n)} &= \frac{1}{\sum_{i=1}^{\infty} c_n^i/i} \\ &= \frac{1}{c_n} \frac{1}{1 - (-c_n/2 - c_n^2/3 - \dots)} \\ &= \frac{1}{\sum_{j=0}^{\infty} (-c_n/2 - c_n^2/3 - \dots)^j} \\ &= \frac{c_n}{1 + (-c_n/2 + O(c_n^2)) + \sum_{j=2}^{\infty} (-c_n/2 + O(c_n^2))^j} \\ &= \frac{1 - c_n/2 + O(c_n^2)}{c_n}. \end{aligned}$$

A terceira equação segue, já que $|-c_n/2 - c_n^2/3 - \dots| \leq \exp(c_n) - 1 < 1$ para todo n grande. Ou seja,

$$\begin{aligned} \nu_n(\theta) &\geq \frac{\ln y_n - \ln \ln y_n - \theta}{c_n} (1 - c_n/2 + O(c_n^2)) \\ &= \frac{\ln y_n}{c_n} - \frac{\ln y_n}{2} + \ln(y_n) O(c_n) - \frac{\ln \ln y_n}{c_n} + \frac{\ln \ln y_n}{2} - \ln \ln(y_n) O(c_n) - \frac{\theta}{c_n} (1 - c_n/2 + O(c_n^2)) \\ &= t_n - \left(\frac{\ln y_n}{2} + \frac{\ln \ln y_n}{c_n} \right) + O(c_n) \ln y_n + \frac{\ln \ln y_n}{2} - \frac{\theta}{c_n} (1 + o(1)) \\ &= t_n - \left[\underbrace{1 - \frac{O(c_n) \ln y_n}{\frac{\ln y_n}{2} + \frac{\ln \ln y_n}{c_n}}}_{(1)} - \underbrace{\frac{\frac{1}{2} \ln \ln y_n}{\frac{1}{2} \ln y_n + \frac{\ln \ln y_n}{c_n}}}_{(2)} + \underbrace{\frac{\frac{\theta}{c_n} (1 - o(1))}{\frac{1}{2} \ln y_n + \frac{\ln \ln y_n}{c_n}}}_{(3)} \right] \left(\frac{1}{2} \ln y_n + \frac{\ln \ln y_n}{c_n} \right). \end{aligned}$$

(1) - $\frac{O(c_n) \ln y_n}{\frac{\ln y_n}{2} + \frac{\ln \ln y_n}{c_n}} = \frac{O(c_n)}{\frac{1}{2} + \frac{\ln \ln y_n}{\ln(y_n)c_n}}$. Se $\frac{\ln \ln y_n}{\ln(y_n)c_n} \rightarrow c \leq \infty$, como $O(c_n) \rightarrow 0$ teremos (1) $\rightarrow 0$, $\frac{\ln \ln y_n}{\ln(y_n)c_n} \rightarrow \infty$ ainda teremos (1) $\rightarrow 0$.

(2) - Pois $\frac{\ln \ln y_n}{\ln y_n} \rightarrow 0$.

(3) = $\frac{\theta(1+o(1))}{\left(\frac{c_n}{2} \ln y_n + \ln \ln y_n \right)} \rightarrow 0$ visto que θ é constante. Note que essa convergência

independe do que ocorre com $c_n \ln y_n$.

Então dado $\theta > 0$ e $\epsilon > 0$, para todo n grande,

$$\nu_n(\theta) \geq t_n - (1+(1)+(2)+(3)) \left(\frac{1}{2} \ln y_n + \frac{\ln \ln y_n}{c_n} \right) > t_n - (1+\epsilon) \left(\frac{1}{2} \ln y_n + \frac{\ln \ln y_n}{c_n} \right) = t_n - b_n.$$

Portanto, para todo n grande o suficiente

$$\inf_{t \leq t_n - b_n} d_t^{(n)}(0, y_n) \geq \inf_{t \leq \nu_n(\theta)} d_t^{(n)}(0, y_n).$$

Isso, com o Lema 4.2.1, prova o segundo item do corolário. QED

Demonstração Teorema 4.2.1. Começemos pelo primeiro limite.

$$d_t^{(n)}(y_n, \pi^{(n)}) \leq d_t^{(n)}(y_n, 0) + d_t^{(n)}(0, \pi^{(n)}) \leq d_t^{(n)}(y_n, 0) + \mu_n \alpha_n^t.$$

Note que $\mu_n \rightarrow \beta$ pelo Teorema 4.1.1 e $\ln \alpha_n^t = t \ln(1 - c_n(1 - p_n)) \leq -\frac{c_n}{2}t$ quando $p_n \leq 1/2$. Portanto

$$\lim_{n \rightarrow \infty} \sup_{t > t_n + \frac{1}{\epsilon c_n}} \mu_n \alpha_n^t = 0.$$

Isso, com o Corolário 4.2.1, garante o primeiro item. Para o segundo item, observe

$$d_t^{(n)}(y_n, \pi^{(n)}) \geq \mathbf{P}_{y_n}^{(n)}(X_t > \gamma_n) - \pi^{(n)}(\{\gamma_n, \gamma_n + 1, \dots\}).$$

Como já argumentamos na demonstração do Corolário 4.2.1, $\nu_n > t_n - b_n$ para todo n grande. Então pelo argumento do Lema 4.2.1

$$\liminf_n \inf_{t \leq t_n - b_n} \mathbf{P}_{y_n}^{(n)}(X_t > \gamma_n) \geq 1 - \exp\left\{-\frac{\beta \theta^4}{48}\right\},$$

para todo $\theta > 0$. Vendo que $\mathbf{P}_{\pi^{(n)}}^{(n)}(X_t > \gamma_n) \leq \frac{\mu_n}{\gamma_n} \rightarrow 0$, provamos o teorema. QED

4.2.3 Simulação do cutoff

Fizemos algumas simulações para ilustrar a validade do Teorema 4.2.1. Tomamos:

$$y_n = n, \quad p_n = \frac{70}{70 + n}, \quad c_n = \frac{1}{n}.$$

Simulamos 60000 trajetórias de $(X_t^{(n)})_{t \in \mathbb{Z}_+}$, partindo de $X_0^{(n)} = 10$, e com elas estimamos empiricamente $\mathbf{p}^t(10, \cdot)$ para

$$t \in (t_n - 3 \max(\log(y_n), \frac{\log(\log(y_n))}{c_n}), t_n + 3 \max(\log(y_n), \frac{\log(\log(y_n))}{c_n})).$$

E com essas estimativas, calculamos $\nu_n(t) = d_t(10, \pi^{(n)})$.

Se o resultado que provamos nas seções anteriores estivesse correto deveríamos ver a $\nu_n(t)$ caindo de 1 para 0 na janela do cutoff. Isso é o que acontece, nos indicando que o resultado é coerente.

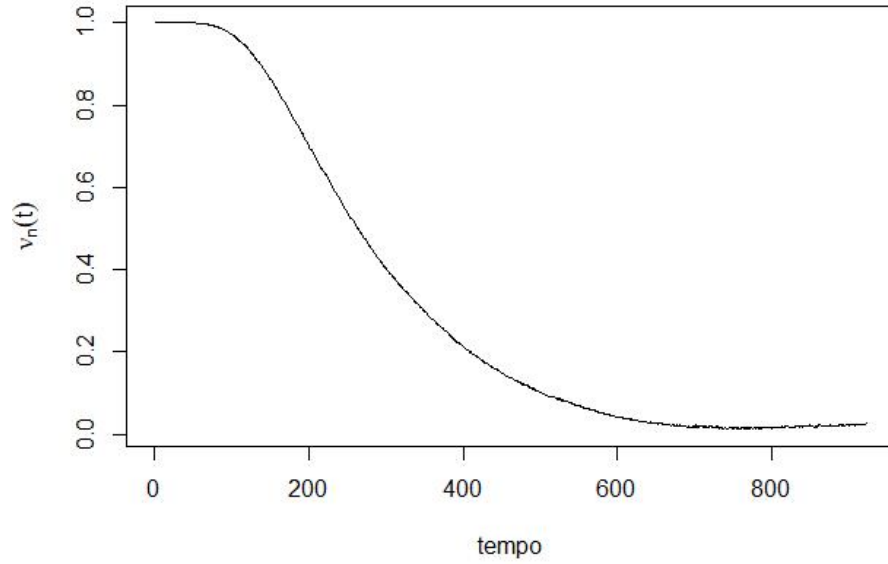


Figura 4.2: Exemplo $n = 100$, intervalo (1, 919).

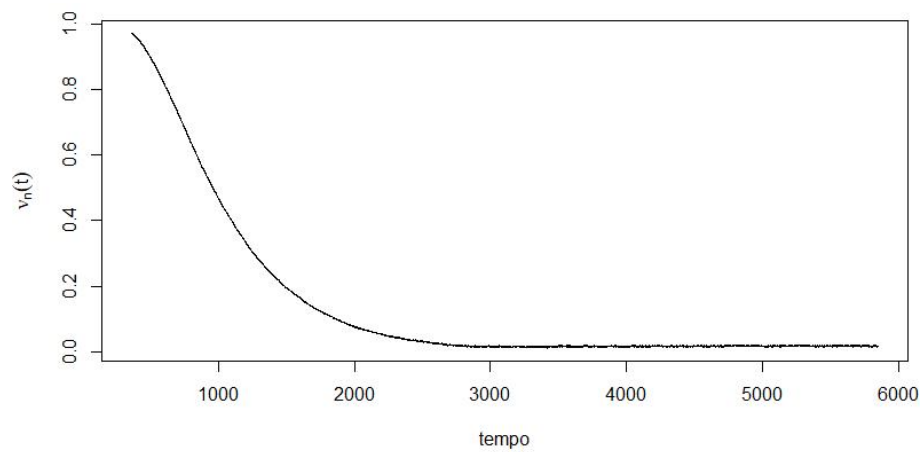


Figura 4.3: Exemplo $n = 500$, intervalo (367, 5848).

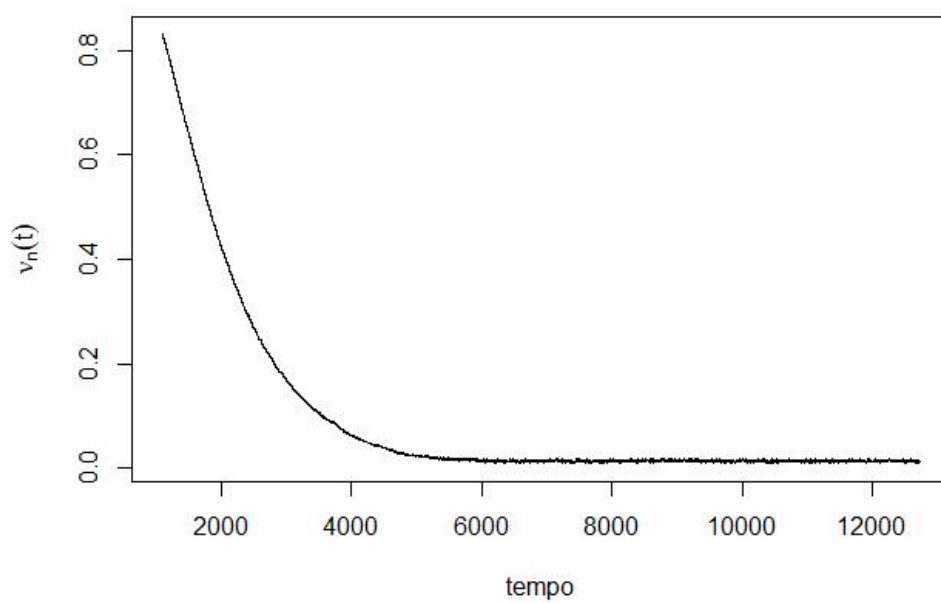


Figura 4.4: Exemplo $n = 1000$, intervalo (1109, 12707).

Capítulo 5

Resultados adicionais

A intenção desse capítulo é apresentar alguns resultados interessantes, mas que destoam muito do objetivo dos capítulos anteriores, o qual era o estudo do fenômeno de cutoff para a Cadeia da Equação 2.2.

5.1 Representação por processos de ramificação

O modelo definido na Equação 2.2 pode ser pensado como um *processo de ramificação com imigração num ambiente aleatório*, nos termos de [15]. Recordando que ω_t é a variável indicadora de catástrofe no momento t , defina $v_t = 1 - \omega_t$, o passeio pode ser escrito como

$$X_{t+1} = \sum_{k=1}^{X_t} U_{k,t} + v_t. \quad (5.1)$$

Em que v_t é interpretado como o número de imigrantes se juntando à população na t -ésima geração e $U_{k,t}$ é a quantidade de descendentes do k -ésimo sujeito vivendo na t -ésima geração. Dado v_t , $\{U_{k,t}\}_k$ são IID com distribuição comum $\text{Ber}(v_t + (1-c)(1-v_t))$, e independentes de X_t .

Esse tipo de representação é muito útil pois escrevemos a variável do processo como uma soma de variáveis aleatórias não negativas. Ademais, a teoria de processos de ramificação é bem desenvolvida. Por exemplo, usando essa representação e [15, Theorem 4.2], pode-se provar que o *tempo até a extinção*, isto é, τ_0 tal como na Definição 2.1.4, tem cauda exponencial, qual seja, $P_0(\tau_0^+ \geq t) \leq a \exp(-bt)$ para algum $a, b > 0$. Essa observação seria suficiente para argumentarmos a recorrência positiva no começo da dissertação.

Construamos o seguinte processo. $T_0 = 0$ e $T_n = \inf\{k > T_{n-1} : \omega_k = 1\}$, note que T_n é uma variável aleatória que designa a época da n -ésima catástrofe. Essa é uma sequência de *tempos de parada*, e o processo $Z_n = X_{T_n}$ forma uma cadeia de Markov pela

*propriedade forte de Markov*¹. Informalmente falando, $(Z_t)_{t \in \mathbb{Z}_+}$ é a cadeia da Equação 2.2 no momento da t -ésima catástrofe. Diremos que $(Z_t)_{t \in \mathbb{Z}_+}$ é a *cadeia(processo) das catástrofes* de $(X_t)_{t \in \mathbb{Z}_+}$. Esse processo é aperiódico, $P(Z_{n+1} = z | Z_n = z) > 0$ pois, se após uma catástrofe a população têm z indivíduos, a probabilidade de na próxima catástrofe a população ficar com z indivíduos é estritamente positiva. É irreduzível visto que a geométrica toma valores em \mathbb{Z}_+ . Além disso, possui uma distribuição estacionária.

Proposição 5.1.1 (Proposition 5.2 de [3]). *A cadeia de Markov $(Z_t)_{t \in \mathbb{Z}_+}$ tem uma distribuição estacionária Z_∞ . Sua função geradora de probabilidades é dada por*

$$\mathbf{E}[s^{Z_\infty}] = \prod_{k=1}^{\infty} \frac{1-p}{1-p(1-(1-c)^k + (1-c)^k s)}, \quad s \in [0, 1].$$

Nos termos da Proposição 2.2.1, $Z_\infty = R - R_0 = \sum_{j=1}^{\infty} \text{Bin}_j(R_j, (1-c)^j)$.

Demonstração. Seguiremos o roteiro da prova da Proposição 2.2.1. Seja $Z_0 \stackrel{d}{=} \sum_{j=1}^{\infty} \text{Bin}_j(R_j, (1-c)^j)$. Perceba que $Z_1 \stackrel{d}{=} \text{Bin}(Z_0 + \text{Geom}^-(1-p), (1-c))$, Z_0 e $\text{Geom}^-(1-p)$ são independentes, pois a ocorrência de catástrofe num certo momento não depende do passado. Então,

$$\begin{aligned} \mathbf{E}(s^{Z_1}) &= \mathbf{E}(\mathbf{E}(s^{Z_1} | Z_0 + \text{Geom}^-(1-p))) \\ &= \mathbf{E}((1 - (1-c) + (1-c)s)^{Z_0 + \text{Geom}^-(1-p)}) \\ &= \prod_{k=1}^{\infty} \frac{1-p}{1-p(1-(1-c)^k + (1-c)^k(1-(1-c) + (1-c)s))} \frac{1-p}{1-p(1-(1-c) + (1-c)s)} \\ &= \prod_{k=1}^{\infty} \frac{1-p}{1-p(1-(1-c)^{k+1} + (1-c)^{k+1})} \frac{1-p}{1-p(1-(1-c) + (1-c)s)} \\ &= \prod_{k=1}^{\infty} \frac{1-p}{1-p(1-(1-c)^k + (1-c)^k s)}. \end{aligned}$$

QED

O resultado da Proposição 5.1.1 está de acordo com a intuição, note que a distribuição de X_t é igual à $Z_{\vartheta(t)} + t - \vartheta(t)$, onde $\vartheta(t)$ é o momento da última catástrofe antes da época t . A variável $t - \vartheta(t)$ é uma geométrica $1-p$ truncada em t , na medida que t cresce essa geométrica tende a ser, com efeito, uma geométrica. A abordagem para demonstrar esse resultado no artigo referência foi outra. Eles representaram Z_t também como um processo de ramificação e usaram, novamente, o artigo [15]. A utilidade do processo das catástrofes é muito grande, porque a estrutura probabilística desse processo é, em geral, mais simples que a do processo original. Por exemplo, [14] e [7], deduziram distribuições estacionárias utilizando a ideia da cadeia das catástrofes para outros passeios aleatórios. Ademais, o processo das catástrofes da Equação 2.2 é mais estável que a cadeia original quando observamos o caso de catástrofes raras, porém severas.

¹Não entraremos nos detalhes desses conceitos. Aos leitores interessados veja [4, Seção 2.5]

Proposição 5.1.2. *Seja $p_n \rightarrow 1$ e $\frac{1-c_n}{1-p_n} \rightarrow \beta \in (0, \infty)$. Denote $X_\infty^{(n)}$ a variável com lei igual à distribuição estacionária de \mathbf{p}^{p_n, c_n} . Seja $Z_\infty^{(n)}$ a variável com lei igual à distribuição estacionária do processo das catástrofes de \mathbf{p}^{p_n, c_n} . Então $X_\infty^{(n)}$ converge para ∞ em distribuição, mas $Z_\infty^{(n)}$ converge para $\text{Geom}^-\left(\frac{1}{1+\beta}\right)$ em distribuição quando n cresce.*

Demonstração. Observe que $\forall M > 0$,

$$\mathbb{P}(X_\infty^{(n)} \leq M) \leq \mathbb{P}(\text{Geom}^-(1-p_n) \leq M) = 1 - (p_n)^M \xrightarrow[n \rightarrow \infty]{} 0.$$

Garantindo que $X_\infty^{(n)}$ converge para infinito em distribuição. Para mostrar que $Z_\infty^{(n)}$ converge para $\text{Geom}^-\left(\frac{1}{1+\beta}\right)$ em distribuição, mostraremos que as funções geradoras de probabilidade dos $(Z_\infty^{(n)})_{n \in \mathbb{N}}$ convergem para a de $\text{Geom}^-\left(\frac{1}{1+\beta}\right)$, veja [8, Continuity Theorem p. 280]. Note que

$$\mathbb{E}(s^{\text{Geom}^-(p)}) = \frac{p}{1 - (1-p)s}, \quad s, p \in [0, 1].$$

Por outro lado, pela Proposição 5.1.1,

$$\mathbb{E}(s^{Z_\infty^{(n)}}) = \prod_{k=1}^{\infty} \frac{1-p_n}{1-p_n(1-(1-c_n)^k + (1-c_n)^k s)}.$$

Avaliemos os casos. O fator associado ao $k=1$ no limite em n converge,

$$\begin{aligned} \frac{1-p_n}{1-p_n(1-(1-c_n) + (1-c_n)s)} &= \frac{1}{1 + \frac{p_n(1-c_n)}{1-p_n}(1-s)} \\ &\xrightarrow[n \rightarrow \infty]{} \frac{1}{1 + \beta(1-s)} \\ &= \frac{\frac{1}{1+\beta}}{1 - (1 - \frac{1}{1+\beta})s}. \end{aligned}$$

Para o outro fator,

$$\sum_{k=2}^{\infty} \ln \left\{ \frac{1-p_n}{1-p_n(1-(1-c_n)^k + (1-c_n)^k s)} \right\} = - \sum_{k=2}^{\infty} \ln \left\{ 1 + \frac{p_n(1-c_n)^k}{1-p_n}(1-s) \right\}.$$

Observe que

$$0 \leq x - \ln(1+x) \leq \frac{x^2}{2} \quad x \in (0, 1).$$

Fazendo $x = \frac{p_n(1-c_n)^k}{1-p_n}(1-s)$ e notando que para todo $k \geq 2$,

$$\frac{p_n(1-c_n)^k}{1-p_n}(1-s) \leq \frac{p_n(1-c_n)^2}{1-p_n}(1-s),$$

podemos dizer que para todo n grande o suficiente, dado s fixo,

$$0 \leq \frac{p_n(1-c_n)^k}{1-p_n}(1-s) - \ln \left[1 + \frac{p_n(1-c_n)^k}{1-p_n}(1-s) \right] \leq \frac{1}{2} \left[\frac{p_n(1-c_n)^k}{1-p_n}(1-s) \right]^2 \quad \forall k \geq 2.$$

Isso implica

$$0 \leq \sum_{k=2}^{\infty} \frac{p_n(1-c_n)^k}{1-p_n} (1-s) - \sum_{k=2}^{\infty} \ln \left[1 + \frac{p_n(1-c_n)^k}{1-p_n} (1-s) \right] \leq \frac{1}{2} \sum_{k=2}^{\infty} \left[\frac{p_n(1-c_n)^k}{1-p_n} (1-s) \right]^2.$$

Observando que,

$$\sum_{k=2}^{\infty} \frac{p_n(1-c_n)^k}{1-p_n} (1-s) = (1-s) \frac{p_n(1-c_n)^2}{c_n(1-p_n)} \xrightarrow{n \rightarrow \infty} 0,$$

teremos

$$\prod_{k=2}^{\infty} \frac{1-p_n}{1-p_n(1-(1-c_n)^k + (1-c_n)^k s)} \xrightarrow{n \rightarrow \infty} 1.$$

Portanto,

$$\lim_{n \rightarrow \infty} \mathbf{E}(s^{Z_{\infty}^{(n)}}) = \frac{\frac{1}{1+\beta}}{1 - (1 - \frac{1}{1+\beta})s}.$$

QED

5.2 Tempo de primeira extinção

Vamos analisar o que ocorre com tempo de primeira extinção quando a população inicial é grande. Denotaremos o tempo até a primeira extinção por τ_0 , essa variável é como na Definição 2.1.4 para o processo da Equação 2.2. Em modelos para crescimento populacional, a relação entre o tamanho da população inicial e o tempo até extinção é, em geral, aproximadamente logarítmica, veja [16, seção 2]. Verificaremos que isso ocorre nesse modelo também. Denote $(X_t^{(0)})_{t \in \mathbb{Z}_+}$ e $(X_t^{(n)})_{t \in \mathbb{Z}_+}$ a cadeia da Equação 2.2 com população inicial 0 e n , respectivamente. Pela regra do acoplamento na Definição 3.2.1, para todo $t \geq 0$

$$X_t^{(n)} = X_t^{(0)} + H_t^{(n)}. \quad (5.2)$$

Além disso, defina

$$\tau^{(n)} = \inf\{t \in \mathbb{Z}_+; X_t^{(n)} = 0\} \quad e \quad \xi^{(n)} = \inf\{t \in \mathbb{Z}_+; H_t^{(n)} = 0\}.$$

Primeiro apresentemos alguns resultados que serão úteis para provar a dependência logarítmica entre a população inicial e τ_0 , mas que são interessantes por eles mesmos.

Proposição 5.2.1. $X_{\xi^{(n)}}^{(0)}$ converge em distribuição para Z_{∞} .

Demonstração. Observe que na regra do acoplamento, o tempo em que $H_t^{(n)} = 0$ pela primeira vez é um tempo no qual ocorreu catástrofe. Denote $(Z_t^{(0)})_{t \in \mathbb{Z}_+}$ a cadeia das catástrofes de $(X_t^{(0)})_{t \in \mathbb{Z}_+}$, $(Z_t^{(n)})_{t \in \mathbb{Z}_+}$ a cadeia das catástrofes de $(X_t^{(n)})_{t \in \mathbb{Z}_+}$ e $(V_t^{(n)})_{t \in \mathbb{Z}_+}$ para o processo das catástrofes de $(H_t^{(n)})_{t \in \mathbb{Z}_+}$. Perceba que pelo Lema 2.2.1, $V_t^{(n)} \stackrel{d}{=} \text{Bin}(n, (1-c)^t)$ e $X_{\xi^{(n)}}^{(0)} \stackrel{d}{=} Z_{\eta^{(n)}}^{(0)}$, onde $\eta^{(n)} = \inf\{t: V_t^{(n)} = 0\}$. Observe que $\eta^{(n)}$ indica em qual catástrofe houve o acoplamento para o modelo original, isto é, se houve acoplamento da primeira catástrofe, $\eta^{(n)} = 1$, se na segunda, $\eta^{(n)} = 2$, e assim por diante. Veja que $\forall M \in \mathbb{N}$,

$$\mathbb{P}(\eta^{(n)} \leq M) = \mathbb{P}(V_M^{(n)} = 0) = (1 - (1-c)^M)^n \rightarrow 0.$$

Ou seja, por Borel-Cantelli, $\eta^{(n)}$ converge para o infinito quase sempre, quando n cresce. Agora provemos que os eventos $\{Z_k^{(0)} = z\}$ e $\{\eta^{(n)} = k\}$ são independentes para todo $z, k \in \mathbb{Z}_+$. Para isso, argumentaremos em paralelo que $\{Z_k^{(0)} = z\}$ e $\{V_k^{(n)} = v\}$ para todo $k, z, v \in \mathbb{Z}_+$, pois isso nos ajudará a provar o primeiro.

Antes de prosseguir façamos uma observação fundamental. Pela Definição 3.2.1, quando ocorre catástrofe, isto é, quando estamos em (c2), X_t toma valor como uma binomial e H_t como outra binomial, as quais são independentes uma da outra e são independentes do passado. Denotaremos $\text{B}(m, (1-c), Z_t)$ para ser a binomial de $(X_t)_{t \in \mathbb{Z}_+}$ na t -ésima catástrofe, dado que logo no tempo anterior à t -ésima catástrofe a população tinha m integrantes. Similarmente para o processo $(H_t)_{t \in \mathbb{Z}_+}$ denotaremos $\text{B}(m, (1-c), V_t)$. Por fim, uma outra notação, R_t será a quantidade de imigrantes que chegaram entre a t -ésima menos uma e a t -ésima catástrofe. Então,

$$\begin{aligned} \mathbb{P}(Z_1^{(0)} = z, V_1^{(n)} = v) &= \mathbb{P}(\text{B}(R_1, (1-c), Z_1) = z, \text{B}(n, (1-c), V_1) = v) \\ &= \sum_r \mathbb{P}(\text{B}(r, (1-c), Z_1) = z) \mathbb{P}(\text{B}(n, (1-c), V_1) = v) \mathbb{P}(R_1 = r) \\ &= \mathbb{P}(Z_1^{(0)} = z) \mathbb{P}(V_1^{(n)} = v). \end{aligned}$$

A segunda igualdade é verdadeira porque as binomiais são independentes do passado, em particular, independentes de R_1 . Em particular,

$$\mathbb{P}(Z_1^{(0)} = z, \eta^{(n)} = 1) = \mathbb{P}(Z_1^{(0)} = z, V_1^{(n)} = 0) = \mathbb{P}(Z_1^{(0)} = z) \mathbb{P}(\eta^{(n)} = 1).$$

Suponha $\mathbb{P}(Z_k^{(0)} = z, V_k^{(n)} = v) = \mathbb{P}(Z_k^{(0)} = z) \mathbb{P}(V_k^{(n)} = v)$. Então,

$$\begin{aligned} \mathbb{P}(Z_{k+1}^{(0)} = z, V_{k+1}^{(n)} = v) &= \mathbb{P}(\text{B}(R_{k+1} + Z_k^{(0)}, (1-c), Z_{k+1}) = z, \text{B}(V_k^{(n)}, (1-c), V_{k+1}) = v) \\ &= \sum_{x,y,r} \mathbb{P}(\text{B}(r+x, (1-c), Z_{k+1}) = z) \mathbb{P}(\text{B}(y, (1-c), V_{k+1}) = v) \mathbb{P}(R_{k+1} = r) \mathbb{P}(Z_k^{(0)} = x) \mathbb{P}(V_k^{(n)} = y) \\ &= \mathbb{P}(Z_{k+1}^{(0)} = z) \mathbb{P}(V_{k+1}^{(n)} = v). \end{aligned}$$

Na segunda igualdade as binomiais são independentes do passado e uma da outra, o R_t é independente do que ocorreu na $t-1$ -ésima catástrofe e usamos a hipótese de

indução. Também, com o mesmo argumento, temos

$$\begin{aligned}
& \mathbb{P}(Z_{k+1}^{(0)} = z, \eta^{(n)} = k + 1) \\
&= \mathbb{P}(B(R_{k+1} + Z(0)_k, (1 - c), Z_{k+1}) = z, B(V_k^{(n)}, (1 - c), V_{k+1}) = 0, V_k^{(n)} > 0) \\
&= \sum_{x,r,y>0} \mathbb{P}(B(r + x, (1 - c), Z_{k+1}) = z) \mathbb{P}(B(y, (1 - c), V_{k+1}) = 0) \mathbb{P}(R_{k+1} = r) \mathbb{P}(Z_k^{(0)} = x) \mathbb{P}(V_k^{(n)} = y) \\
&= \mathbb{P}(Z_{k+1}^{(0)} = z) \mathbb{P}(\eta^{(n)} = k + 1).
\end{aligned}$$

Portanto, pelo Teorema da Convergência Dominada,

$$\begin{aligned}
\mathbb{P}(X_{\xi^{(n)}}^{(0)} = x) &= \mathbb{P}(Z_{\eta^{(n)}}^{(0)} = x) \\
&= \sum_{m=1}^{\infty} \mathbb{P}(Z_m^{(0)} = x) \mathbb{P}(\eta^{(n)} = m) \\
&= \mathbb{E}\left(\sum_{m=1}^{\infty} \mathbb{P}(Z_m^{(0)} = x) \mathbb{I}_{\{\eta^{(n)}=m\}}\right) \\
&= \mathbb{E}(\mathbb{P}(Z_{\eta^{(n)}}^{(0)} = x)) \xrightarrow{n \rightarrow \infty} \mathbb{P}(Z_{\infty}^{(0)} = x).
\end{aligned}$$

QED

Proposição 5.2.2 (Proposition 5.4 de [3]). $\tau^{(n)} - \xi^{(n)}$ converge em distribuição para $\mathbb{P}_{Z_{\infty}}(\tau_0 \in \cdot)$ com $n \rightarrow \infty$.

Demonstração. Antes de iniciar a prova do resultado propriamente, façamos algumas observações preliminares. Pela Equação 5.2 é claro que $\tau^{(n)} \geq \xi^{(n)}$. A variável $\rho^{(n)} = \tau^{(n)} - \xi^{(n)}$ depende do passado do sistema acoplado, isto é, do sistema até o tempo de acoplamento, apenas através da variável $X_{\xi^{(n)}}^{(n)} = X_{\xi^{(n)}}^{(0)}$. De fato, seja

$$\mathcal{F}_t = \sigma(X_0^{(0)}, H_0^{(n)}, \dots, X_t^{(0)}, H_t^{(n)}), \quad t \in \mathbb{Z}_+.$$

Note que $\{\xi^n \leq k\} = \{H_k^n = 0\} \in \mathcal{F}_k$, ou seja, $\xi^{(n)}$ é um tempo de parada para essa filtração. Também,

$$\mathbb{P}(\xi^{(n)} \leq +\infty) \geq \mathbb{P}(\xi^{(n)} \leq 1) = \mathbb{P}(H_1^{(n)} = 0) = \mathbb{E}((1 - (1 - c)^{N_1})^n) = (1 - p)c^n > 0.$$

Com isso, podemos usar a propriedade forte de Markov na cadeia $(X_t^{(0)}, H_t^{(n)})_{t \in \mathbb{Z}_+}$.

$$\begin{aligned}
\mathbb{P}(\rho^{(n)} = k | \mathcal{F}_{\xi^{(n)}}) &= \mathbb{P}(\tau^{(n)} = k + \xi^{(n)} | \mathcal{F}_{\xi^{(n)}}) \\
&= \mathbb{P}(X_{\xi^{(n)}+k}^{(0)} = 0, H_{\xi^{(n)}+k}^{(n)} = 0, \cap_{j=0}^{k-1} \{X_{\xi^{(n)}+j}^{(0)} > 0, H_{\xi^{(n)}+j}^{(n)} = 0\} | \mathcal{F}_{\xi^{(n)}}) \\
&= \mathbb{P}(X_{\xi^{(n)}+k}^{(0)} = 0, H_{\xi^{(n)}+k}^{(n)} = 0, \cap_{j=0}^{k-1} \{X_{\xi^{(n)}+j}^{(0)} > 0, H_{\xi^{(n)}+j}^{(n)} = 0\} | X_{\xi^{(n)}}^{(0)}, H_{\xi^{(n)}}^{(n)}) \\
&= \mathbb{P}(\rho^{(n)} = k | X_{\xi^{(n)}}^{(0)}).
\end{aligned}$$

Perceba que, $H_{\xi^{(n)}}^{(n)} = 0$ logo $\mathbf{P}(\rho^{(n)} = k | X_{\xi^{(n)}}^{(0)}) = \mathbf{P}(\rho^{(n)} = k | X_{\xi^{(n)}}^{(0)}, H_{\xi^{(n)}}^{(n)})$. Então a distribuição de $\rho^{(n)}$ coincide com a de $\tau^{(X_{\xi^{(n)}}^{(0)})}$ (a época de extinção do passeio onde a distribuição inicial é $X_{\xi^{(n)}}^{(0)}$). De fato,

$$\begin{aligned} \mathbf{P}(\rho^{(n)} = k) &= \mathbf{P}(\tau^{(n)} = k + \xi^{(n)}) \\ &= \mathbf{P}(X_{k+\xi^{(n)}}^{(n)} = 0, \cap_{j=0}^{\xi^{(n)}+k-1} \{X_j > 0\}) \\ &= \sum_{m>0} \mathbf{P}(X_{k+\xi^{(n)}}^{(n)} = 0, \cap_{j=1}^{k-1} \{X_{j+\xi^{(n)}}^{(n)} > 0\} | X_{\xi^{(n)}}^{(n)} = m) \mathbf{P}(X_{\xi^{(n)}}^{(n)} = m) \\ &= \sum_{m>0} \mathbf{P}(X_k^{(n)} = 0, \cap_{j=1}^{k-1} \{X_j^{(n)} > 0\} | X_0^{(n)} = m) \mathbf{P}(X_{\xi^{(n)}}^{(0)} = m) = \mathbf{P}(\tau^{(X_{\xi^{(n)}}^{(0)})} = k). \end{aligned}$$

Na terceira e na quarta igualdade, usamos que $X_i^{(n)} > 0$ sempre que $i < \xi^{(n)}$ e o [4, Teorema 2.5.8]. Agora estamos em posição para provar enunciado dessa proposição. Argumentaremos uma indução. Temos,

$$\mathbf{P}(X_{\xi^{(n)}}^{(0)} = k) \xrightarrow{n \rightarrow \infty} \mathbf{P}(Z_{\infty} = k) = \mathbf{P}_{Z_{\infty}}(X_0 = k).$$

O limite vale pela Proposição 5.2.1. Suponha,

$$\begin{aligned} &\mathbf{P}(X_{\xi^{(n)}}^{(0)} > 0, X_{\xi^{(n)}+1}^{(0)} > 0, \dots, X_{\xi^{(n)}+m-1}^{(0)} > 0, X_{\xi^{(n)}+m}^{(0)} = k) \\ &\xrightarrow{n \rightarrow \infty} \mathbf{P}_{Z_{\infty}}(X_0 > 0, X_1 > 0, \dots, X_{m-1} > 0, X_m = k), \end{aligned}$$

para todo k . Então, pela propriedade forte de Markov,

$$\begin{aligned} &\mathbf{P}(X_{\xi^{(n)}}^{(0)} > 0, X_{\xi^{(n)}+1}^{(0)} > 0, \dots, X_{\xi^{(n)}+m}^{(0)} > 0, X_{\xi^{(n)}+m+1}^{(0)} = k) = \\ &\sum_{j=1}^{\infty} \mathbf{P}(X_{\xi^{(n)}}^{(0)} > 0, X_{\xi^{(n)}+1}^{(0)} > 0, \dots, X_{\xi^{(n)}+m}^{(0)} = j) \mathbf{p}(j, k) \\ &\xrightarrow{n \rightarrow \infty} \sum_{j=1}^{\infty} \mathbf{P}_{Z_{\infty}}(X_0 > 0, X_1 > 0, \dots, X_{m-1} > 0, X_m = k) \mathbf{p}(j, k) = \\ &\mathbf{P}_{Z_{\infty}}(X_0 > 0, X_1 > 0, \dots, X_m > 0, X_{m+1} = k). \end{aligned}$$

Podemos usar o Teorema da Convergência Dominada porque,

$$\begin{aligned} \sum_{j=0}^{\infty} \mathbf{p}(j, k) &= p + \sum_{r=0}^{\infty} (1-p)(1-c)^k \binom{k+r}{k} c^r < 0 \text{ já que} \\ \frac{\binom{k+r+1}{k} c^{r+1}}{\binom{k+r}{k} c^r} &= \frac{k+r+1}{r+1} c \xrightarrow{r \rightarrow \infty} c < 1. \end{aligned}$$

Em particular,

$$\begin{aligned} \mathbf{P}(\rho^{(n)} = k) &= \mathbf{P}(\tau^{(X_{\xi^{(n)}}^{(0)})} = k) = \mathbf{P}(X_{\xi^{(n)}}^{(0)} > 0, X_{\xi^{(n)}+1}^{(0)} > 0, \dots, X_{\xi^{(n)}+k-1}^{(0)} > 0, X_{\xi^{(n)}+k}^{(0)} = 0) \\ &\rightarrow \mathbf{P}_{Z_{\infty}}(X_0 > 0, X_1 > 0, \dots, X_{k-1} > 0, X_k = 0) = \mathbf{P}_{Z_{\infty}}(\tau_0 = k). \end{aligned}$$

QED

Proposição 5.2.3 (Proposition 5.5 de [3]). $\tau^{(n)}/d_n$ converge em probabilidade para 1. Onde $d_n = -\frac{\ln(n)}{(1-p)\ln(1-c)}$.

Demonstração. Argumentaremos que $\xi^{(n)}/d_n \rightarrow 1$ em probabilidade. Como $\tau^{(n)} - \xi^{(n)}$ converge em distribuição para $\mathbb{P}_{Z_\infty}(\tau_0 \in \cdot)$, uma variável integrável, então $(\tau^{(n)} - \xi^{(n)})/d_n$ convergirá para 0 em distribuição, logo em probabilidade. Por Slutsky, o resultado estará provado.

Lembre que para um acoplamento qualquer,

$$\mathbb{P}_{x,y}(\xi \leq t) = \mathbb{E}_{x,y}(\mathbb{P}_{x,y}(H_t = 0 | N_t)) = \mathbb{E}_{x,y}((1 - (1 - c)^{N_t})^{y-x}).$$

Sendo $\epsilon \in (0, \frac{1}{2})$, e $A_t = \{|N_t/t - (1 - p)| < (1 - p)\epsilon\}$ os autores afirmaram que,

$$\mathbb{E}[(1 - (1 - c)^{(1-\epsilon)(1-p)t})^n, A_t] \leq \mathbb{E}[(1 - (1 - c)^{N_t})^n] \leq \mathbb{E}[(1 - (1 - c)^{t(1+\epsilon)(1-p)})^n] + \mathbb{P}(A_t^c).$$

De fato, pois

$$\begin{aligned} \left| \frac{i}{t} - (1 - p) \right| &< (1 - p)\epsilon \\ \Leftrightarrow (1 - \epsilon)(1 - p)t &< i < (1 + \epsilon)(1 - p)t \\ \Leftrightarrow (1 - c)^{(1-\epsilon)(1-p)t} &> (1 - c)^i > (1 - c)^{(1+\epsilon)(1-p)t} \\ \Leftrightarrow 1 - (1 - c)^{(1-\epsilon)(1-p)t} &< 1 - (1 - c)^i < 1 - (1 - c)^{(1+\epsilon)(1-p)t}. \end{aligned}$$

Com isso,

$$\mathbb{E}[(1 - (1 - c)^{N_t})^n] \mathbb{E}[(1 - (1 - c)^{N_t})^n, A_t] + \mathbb{E}[(1 - (1 - c)^{N_t})^n, A_t^c] \leq (1 - (1 - c)^{t(1+\epsilon)(1-p)})^n + \mathbb{P}(A_t^c)$$

$$\text{e } \mathbb{E}[(1 - (1 - c)^{N_t})^n] \geq (1 - (1 - c)^{t(1+\epsilon)(1-p)})^n \mathbb{P}(A_t).$$

Pela Lei Forte, sabemos que $\lim_{t \rightarrow \infty} \mathbb{P}(A_t) = 1$.

Usando as desigualdades acima, para todo $t \leq (1 - \epsilon)d_n$ vale o seguinte

$$\begin{aligned} \mathbb{P}(\xi^{(n)} \leq t) &\leq \mathbb{P}(\xi^{(n)} \leq (1 - \epsilon)d_n) \\ &\leq (1 - (1 - c)^{(1-\epsilon^2)(1-p)d_n})^n \mathbb{P}(A_{(1-\epsilon)d_n}) + \mathbb{P}(A_{(1-\epsilon)d_n}^c) \\ &= (1 - (1 - c)^{(1-\epsilon^2)(1-p)(-\frac{\ln(n)}{(1-p)\ln(1-c)})})^n (1 + o(1)) + o(1) \\ &= (1 - \left[(1 - c)^{\frac{\ln(n)}{\ln(1-c)}} \right]^{-(1-\epsilon^2)})^n (1 + o(1)) + o(1) \\ &= (1 - n^{-(1-\epsilon^2)})^n (1 + o(1)) + o(1) \\ &= (1 + \frac{-n^{\epsilon^2}}{n})^n (1 + o(1)) + o(1) \\ &\approx \exp\{-n^{\epsilon^2}\} (1 + o(1)) + o(1) \xrightarrow[n \rightarrow \infty]{} 0. \end{aligned}$$

Por outro lado, para todo $t \geq (1 + 2\epsilon)d_n$

$$\begin{aligned}
\mathbb{P}(\xi^{(n)} \leq t) &\geq \mathbb{P}(\xi^{(n)} \leq (1 + 2\epsilon)d_n) \\
&\geq (1 - (1 - c)^{(1-\epsilon)(1+2\epsilon)(1-p)d_n})^n \mathbb{P}(A_{(1+2\epsilon)d_n}) \\
&= (1 - (1 - c)^{(1-\epsilon)(1+2\epsilon)(1-p)\left(-\frac{\ln(n)}{(1-p)\ln(1-c)}\right)})^n (1 + o(1)) \\
&= (1 - \left[(1 - c)^{\frac{\ln(n)}{\ln(1-c)}} \right]^{-(1-\epsilon)(1+2\epsilon)})^n (1 + o(1)) \\
&= (1 - n^{-(1-\epsilon)(1+2\epsilon)})^n (1 + o(1)) \\
&= (1 - \frac{1}{n^{(1-\epsilon)(1+2\epsilon)}})^n (1 + o(1)) \\
&\approx \exp\{-n^{\epsilon(2\epsilon-1)}\} (1 + o(1)) \xrightarrow[n \rightarrow \infty]{} 1.
\end{aligned}$$

Pois $\epsilon < \frac{1}{2}$.

QED

Referências

- [1] David Aldous. Random walks on finite groups and rapidly mixing markov chains. In *Séminaire de Probabilités XVII 1981/82: Proceedings*, pages 243–297. Springer, 1983.
- [2] David Aldous and Persi Diaconis. Strong uniform times and finite random walks. *Advances in Applied Mathematics*, 8(1):69–97, 1987.
- [3] Iddo Ben-Ari, Alexander Roitershtein, and Rinaldo B Schinazi. A random walk with catastrophes. *Electronic journal of probability*, 24(28):1–21, 2019.
- [4] Pierre Brémaud. *Markov chains: Gibbs fields, Monte Carlo simulation, and queues*. Springer Science & Business Media, 2001.
- [5] Peter J Brockwell, J Gani, and Sidney I Resnick. Birth, immigration and catastrophe processes. *Advances in Applied Probability*, 14(4):709–731, 1982.
- [6] Frank Den Hollander. Probability theory: The coupling method. *Lecture notes available online (<http://websites.math.leidenuniv.nl/probability/lecturenotes/CouplingLectures.pdf>)*, 2012.
- [7] Antonis Economou. The compound poisson immigration process subject to binomial catastrophes. *Journal of applied probability*, 41(2):508–523, 2004.
- [8] William Feller. *An introduction to probability theory and its applications*, volume 1. John Wiley & Sons, 1967.
- [9] Luiz Renato Fontes and Rinaldo B Schinazi. Metastability of a random walk with catastrophes. *Electronic Communications in Probability*, 24(70):1–8, 2019.
- [10] W Keith Hastings. Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57(1):97–109, 1970.
- [11] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *The collected works of Wassily Hoeffding*, pages 409–426, 1994.
- [12] Steve Huntsman. Fast markov chain monte carlo algorithms via lie groups. In *International Conference on Artificial Intelligence and Statistics*, pages 2841–2851. PMLR, 2020.

-
- [13] Marius Iosifescu. *Finite Markov processes and their applications*. John Wiley & Sons, 1980.
- [14] Norman Kaplan, Aidan Sudbury, and Trygve S Nilsen. A branching process with disasters. *Journal of Applied Probability*, 12(1):47–59, 1975.
- [15] Eric S Key. Limiting distributions and regeneration times for multitype branching processes with immigration in a random environment. *The Annals of Probability*, 15(1):344–353, 1987.
- [16] Russell Lande. Risks of population extinction from demographic and environmental stochasticity and random catastrophes. *The American Naturalist*, 142(6):911–927, 1993.
- [17] David A Levin and Yuval Peres. *Markov chains and mixing times*. American Mathematical Soc., 2017.
- [18] Nicholas Metropolis, Arianna W Rosenbluth, Marshall N Rosenbluth, Augusta H Teller, and Edward Teller. Equation of state calculations by fast computing machines. *The journal of chemical physics*, 21(6):1087–1092, 1953.
- [19] Marcel F Neuts. An interesting random walk on the non-negative integers. *Journal of applied probability*, 31(1):48–58, 1994.

Apêndice A

Contra exemplos Lema 2.2.1

A.1 Contra exemplo 1

Creio que na verdade foi um erro de digitação. Mas os autores enunciaram que $(\text{Bin}_j(R_j, \epsilon_j) : j \geq 1)$ é independente condicionado em $(R_j : j \in \mathbb{Z}_+)$ dando a entender que não, necessariamente, $\text{Bin}_0(R_0, \epsilon_0)$ é independente dos $\{\text{Bin}_{j_1}(R_{j_1}, \epsilon_{j_1}), \dots, \text{Bin}_{j_n}(R_{j_n}, \epsilon_{j_n})\}$ dado $(R_j)_{j \geq 0} \forall \{j_1, \dots, j_n\} \subset \mathbb{N}$. Meu ponto é que $\text{Bin}_0(R_0, \epsilon_0)$ deve ser independente dos $\{\text{Bin}_j(R_j, \epsilon_j); j \geq 1\}$ dado $(R_j)_{j \geq 0}$ pois se não for eu consigo construir um exemplo onde o lema é inválido. Com efeito, seja $R_1, R_2, R_3 \text{ IID } \sim \text{Ber}(\frac{1}{2})$.

$$\begin{aligned}
& \mathbb{P}(\text{Bin}_0(R_0, \epsilon_0) = j_0, \text{Bin}_1(R_1, \epsilon_1) = j_1, \text{Bin}_2(R_2, \epsilon_2) = j_2 | R_0 = r_0, R_1 = r_1, R_2 = r_2) \\
&= \binom{r_2}{j_2} \epsilon_2^{j_2} (1 - \epsilon_2)^{r_2 - j_2} \mathbb{I}_{\{0, \dots, r_2\}}(j_2) \mathbb{P}(\text{Bin}_0(R_0, \epsilon_0) = j_0, \text{Bin}_1(R_1, \epsilon_1) = j_1 | R_0 = r_0, R_1 = r_1, R_2 = r_2) \\
&= \binom{r_2}{j_2} \epsilon_2^{j_2} (1 - \epsilon_2)^{r_2 - j_2} \mathbb{I}_{\{0, \dots, r_2\}}(j_2) \left\{ \binom{r_0}{j_0} \epsilon_0^{j_0} (1 - \epsilon_0)^{r_0 - j_0} \binom{r_1}{j_1} \epsilon_1^{j_1} (1 - \epsilon_1)^{r_1 - j_1} + \right. \\
& \left. \mathbb{I}_{\{1\}}(r_0) (-1)^{j_0} (\min(\epsilon_0, 1 - \epsilon_0))^{2r_0} (-1)^{r_1} \mathbb{I}_{\{1\}}(r_1) (-1)^{j_1} (\min(\epsilon_1, 1 - \epsilon_1))^{2r_1} (-1)^{r_0} \right\} \times \\
& \mathbb{I}_{\{0, \dots, r_0\}}(j_0) \mathbb{I}_{\{0, 1\}}(r_0) \mathbb{I}_{\{0, \dots, r_1\}}(j_1) \mathbb{I}_{\{0, 1\}}(r_1).
\end{aligned}$$

Perceba que dessa definição, $\text{Bin}_2(R_2, \epsilon_2)$ é independente de $\{\text{Bin}_1(R_1, \epsilon_1), \text{Bin}_0(R_0, \epsilon_0)\}$ dado (R_0, R_1, R_2) . Além disso, $\text{Bin}_i(R_i, \epsilon_i)$ dado R_i é binomialmente distribuído, R_i e ϵ_i , sendo independente de $R_j, j \neq i$.

Nesses termos, temos que $\text{Bin}(\text{Bin}_0(R_0, \epsilon_0) + \text{Bin}_1(R_1, \epsilon_1) + \text{Bin}_2(R_2, \epsilon_2), \epsilon) \stackrel{d}{\neq} \text{Bin}_0(R_0, \epsilon_0 \epsilon) + \text{Bin}_1(R_1, \epsilon_1 \epsilon) + \text{Bin}_2(R_2, \epsilon_2 \epsilon)$. Com efeito, seja $R_0, R_1, R_2 \text{ IID } \sim \text{Ber}(\frac{1}{2}), \epsilon = \epsilon_0 = \dots =$

$$\epsilon_2 = \frac{1}{2}.$$

$$\begin{aligned} & \mathbb{P}(\text{Bin}(\text{Bin}_0(R_0, \frac{1}{2}) + \text{Bin}_1(R_1, \frac{1}{2}) + \text{Bin}_2(R_2, \frac{1}{2}), \frac{1}{2}) = 3) \\ &= \left(\frac{1}{2}\right)^3 \mathbb{P}(\text{Bin}(\text{Bin}_0(R_0, \frac{1}{2}) + \text{Bin}_1(R_1, \frac{1}{2}) + \text{Bin}_2(R_2, \frac{1}{2}) = 3 | R_0 = R_1 = R_2 = 1) \left(\frac{1}{2}\right)^3 \\ &= \left(\frac{1}{2}\right)^6 \sum_{m_1 \geq 0} \sum_{m_2 \geq 0} \mathbb{P}(\text{Bin}(\text{Bin}_0(R_0, \frac{1}{2}) = 3 - m_1 - m_2, \text{Bin}_1(R_1, \frac{1}{2}) = m_1, \text{Bin}_2(R_2, \frac{1}{2}) = m_2 | \\ & R_0 = R_1 = R_2 = 1) \\ &= \left(\frac{1}{2}\right)^6 \sum_{m_1 \geq 0} \sum_{m_2 \geq 0} \binom{1}{m_2} \left(\frac{1}{2}\right)^1 \mathbb{I}_{\{0,1\}}(m_2) \left\{ \binom{1}{m_1} \left(\frac{1}{2}\right)^1 \binom{1}{3 - m_1 - m_2} \left(\frac{1}{2}\right)^1 + \right. \\ & \left. (-1)^{3 - m_1 - m_2} \frac{1}{4} (-1)^1 (-1)^{m_1} \frac{1}{4} (-1)^1 \right\} \mathbb{I}_{\{0,1\}}(3 - m_1 - m_2) \mathbb{I}_{\{0,1\}}(m_1) \\ &= \left(\frac{1}{2}\right)^6 \frac{1}{2} \left\{ \frac{1}{2} \frac{1}{2} + \left(\frac{1}{4}\right)^2 \right\} = \left(\frac{1}{2}\right)^6 \frac{5}{32}. \end{aligned}$$

Por outro lado,

$$\begin{aligned} & \mathbb{P}(\text{Bin}_0(R_0, \frac{1}{4}) + \text{Bin}_1(R_1, \frac{1}{4}) + \text{Bin}_2(R_2, \frac{1}{4}) = 3) \\ &= \mathbb{P}(\text{Bin}_0(R_0, \frac{1}{4}) + \text{Bin}_1(R_1, \frac{1}{4}) + \text{Bin}_2(R_2, \frac{1}{4}) = 3 | R_0 = R_1 = R_2 = 1) \left(\frac{1}{2}\right)^3 \\ &= \left[\frac{1}{4} \left\{ \frac{1}{4} \frac{1}{4} + \left(\frac{1}{4}\right)^2 \left(\frac{1}{4}\right)^2 \right\} \right] \left(\frac{1}{2}\right)^3 = \left(\frac{1}{2}\right)^5 \left\{ \frac{17}{(16)^2} \right\} \end{aligned}$$

Como $\left(\frac{1}{2}\right)^5 \frac{17}{(16)^2} \neq \left(\frac{1}{2}\right)^6 \frac{5}{32}$, finalizamos o contra-exemplo.

A.2 Contra-exemplo 2

No enunciado do lema 2.2.1 não foi dito explicitamente que $\text{Bin}_j(R_j, \epsilon_j)$ é independente dos $(R_i)_{j \neq i}$ dado R_j . Isso pode ser um costume da área, algo que já deveria ser entendido quando eles definiram a variável $\text{Bin}_j(R_j, \epsilon_j)$. De qualquer forma, na demonstração do lema usamos essa hipótese e por isso apresentamos aqui um contra-exemplo ao enunciado do lema caso essa hipótese não seja válida. Seja $R_1, R_2 \text{ IID } \sim \text{Ber}(\frac{1}{2})$. Definamos

$$\begin{aligned} & \mathbb{P}(\text{Bin}_1(R_1, p) = k | R_1 = r_1, R_2 = r_2) \\ &= \left[\binom{r_1}{k} p^k (1-p)^{r_1 - k} + \mathbb{I}_{\{1\}}(r_1) (-1)^k (\min(1-p, p))^{2r_1} (-1)^{r_2} \right] \mathbb{I}_{\{0, \dots, r_1\}}(k) \mathbb{I}_{\{0,1\}}(r_1). \end{aligned}$$

Similarmente para $\text{Bin}_2(R_2, p)$. Dessa definição, se pode deduzir que $\text{Bin}_1(R_1, p)$ dado R_1 é binomialmente distribuído. Consideraremos $\text{Bin}_1(R_1, \epsilon_1), \text{Bin}_2(R_2, \epsilon_2)$ independentes

dados (R_1, R_2) . Mostraremos que nesse contexto, $\text{Bin}(\text{Bin}_1(R_1, \frac{1}{2}) + \text{Bin}_2(R_2, \frac{1}{2}), \frac{1}{2}) \stackrel{d}{\neq} \text{Bin}_1(R_1, \frac{1}{4}) + \text{Bin}_2(R_2, \frac{1}{4})$. Sem entrar em muitos detalhes, teremos

$$\begin{aligned} & \mathbb{P}(\text{Bin}(\text{Bin}_1(R_1, \frac{1}{2}) + \text{Bin}_2(R_2, \frac{1}{2}), \frac{1}{2}) = 2) \\ &= \left(\frac{1}{2}\right)^2 \mathbb{P}(\text{Bin}_1(R_1, \frac{1}{2}) + \text{Bin}_2(R_2, \frac{1}{2}) = 2 | R_1 = R_2 = 1) \left(\frac{1}{2}\right)^2 = \left(\frac{1}{2}\right)^4 \left(\frac{3}{4}\right)^2, \\ & \mathbb{P}(\text{Bin}_1(R_1, \frac{1}{4}) + \text{Bin}_2(R_2, \frac{1}{4}) = 2) = \mathbb{P}(\text{Bin}_1(R_1, \frac{1}{4}) + \text{Bin}_2(R_2, \frac{1}{4}) = 2 | R_1 = R_2 = 1) \left(\frac{1}{2}\right)^2 \\ &= \left[\frac{5}{16}\right]^2 \left(\frac{1}{2}\right)^2. \end{aligned}$$

Como essas probabilidades são diferentes, as distribuições das variáveis aleatórias são diferentes.

Apêndice B

Sobre a relação de X_t , H_t e N_t

Então devemos provar,

$$P_{x,y}(H_i = h, X_t = s | N_t = k) = P_{x,y}(H_i = h | N_t = k)P_{x,y}(X_t = s | N_t = k), \quad \forall i \geq t, \forall s, \forall t, \forall h, \forall k.$$

Demonstração. Usarei P ao invés de $P_{x,y}$ no que se segue, mas que fique sempre claro que estou considerando os estados iniciais do acoplamento $x < y$. Note,

$$\begin{aligned} P(H_i = h, X_t = s | N_t = k) &= P(H_i = h | N_t = k, X_t = s)P(X_t = s | N_t = k) \\ &= P(X_t = s | N_t = k, H_i = h)P(H_i = h | N_t = k). \end{aligned}$$

Também

$$P(H_i = h | N_t = k, X_t = s) = P(H_i = h | N_t = k) \Leftrightarrow P(X_t = s | N_t = k, H_i = h) = P(X_t = s | N_t = k).$$

Provarei, por indução, a seguinte identidade.

$$P(X_t = s | N_t = k, H_i = h, N_i = n) = P(X_t = s | N_t = k, N_i = n) \quad \forall i \forall t. \quad (\text{B.1})$$

Para $i = t = 1$ e $n \neq k$,

$$P(H_1 = h, |N_1 = n, X_1 = s, N_1 = k) = 0 = P(H_1 = h, |N_1 = n, N_1 = k).$$

Se $n=k=0$

$$\begin{aligned} P(H_1 = h, |N_1 = n, X_1 = s, N_1 = k) &= P(H_1 = h, |N_1 = 0, X_1 = s) \\ &= P(H_1 = h, |N_1 = 0, X_1 = s, X_0 = x, X'_0 = y, H_0 = y - x) \\ &= \mathbb{I}(h)_{\{y-x\}} \\ &= P(H_1 = h | N_1 = 0). \end{aligned}$$

Se $n=k=1$

$$\begin{aligned} P(H_1 = h, |N_1 = n, X_1 = s, N_1 = k) &= P(H_1 = h, |N_1 = 1, X_1 = s) \\ &= P(H_1 = h, |N_1 = 1, X_1 = s, X_0 = x, X'_0 = y, H_0 = y - x) \\ &= P(\text{Bin}(y - x, 1 - c) = h) \\ &= P(H_1 = h | N_1 = 1). \end{aligned}$$

Nos raciocínios acima, na terceira igualdade regra da Definição 3.2.1 e $\omega_1 = N_1$. A quarta igualdade é verdadeira pelo Lema 3.2.1. Por indução, suponha

$$\mathbb{P}(X_t = s | N_1 = n, H_1 = h, N_t = k) = \mathbb{P}(X_t = s | N_1 = n, N_t = k).$$

Considerando s, n, h, k devidos e o passo de indução,

$$\begin{aligned} \mathbb{P}(X_{t+1} = s | N_1 = n, H_1 = h, N_{t+1} = k) = & \\ & \overbrace{\mathbb{P}(X_{t+1} = s | N_1 = n, H_1 = h, N_{t+1} = k, \omega_{t+1} = 0)}^{(II)} \overbrace{\mathbb{P}(\omega_{t+1} = 0 | N_1 = n, H_1 = h, N_{t+1} = k)}^{(I)} + \\ & \overbrace{\mathbb{P}(X_{t+1} = s | N_1 = n, H_1 = h, N_{t+1} = k, \omega_{t+1} = 1)}^{(III)} \mathbb{P}(\omega_{t+1} = 1 | N_1 = n, H_1 = h, N_{t+1} = k). \end{aligned}$$

Analisemos os casos,

(I)

$$\begin{aligned} \mathbb{P}(N_1 = n, N_{t+1} = k | \omega_{t+1} = 0, H_1 = h) &= \frac{\mathbb{P}(\sum_{m=2}^{t+1} \omega_m = k - n, \omega_{t+1} = 0, H_1 = h)}{\mathbb{P}(\omega_{t+1} = 0, H_1 = h)} \\ &= \mathbb{P}(N_1 = n, N_{t+1} = k | \omega_{t+1} = 0). \end{aligned}$$

A segunda equação é pela independência de ω_m do passado.

(II)

$$\begin{aligned} \mathbb{P}(X_{t+1} = s | N_1 = n, H_1 = h, N_{t+1} = k, \omega_{t+1} = 0) &= \mathbb{P}(X_t = s - 1 | N_1 = n, H_1 = h, N_t = k, \omega_{t+1} = 0) \\ &= \mathbb{P}(X_t = s - 1 | N_1 = n, H_1 = h, N_t = k) \\ &= \mathbb{P}(X_t = s - 1 | N_1 = n, N_t = k) \\ &= \mathbb{P}(X_t = s - 1 | N_1 = n, N_t = k, \omega_{t+1} = 0) \\ &= \mathbb{P}(X_{t+1} = s | N_1 = n, N_{t+1} = k, \omega_{t+1} = 0). \end{aligned}$$

A primeira igualdade vale porque

$$\{\omega_{t+1} = 0\} \cap \{N_{t+1} = k\} = \{\omega_{t+1} = 0\} \cap \{N_t = k\}$$

e

$$\{\omega_{t+1} = 0\} \cap \{X_{t+1} = s\} = \{\omega_{t+1} = 0\} \cap \{X_t = s - 1\}.$$

A segunda vale, pois ω_{t+1} é independente do passado. Na terceira usamos a hipótese de indução.

(III)

$$\begin{aligned}
& \mathbf{P}(X_{t+1} = s | N_1 = n, H_1 = h, \omega_{t+1} = 1, N_{t+1} = k) \\
&= \mathbf{P}(X_{t+1} = s | N_1 = n, H_1 = h, \omega_{t+1} = 1, N_t = k - 1) \\
&= \sum_j \left[\mathbf{P}(X_{t+1} = s | N_1 = n, H_1 = h, \omega_{t+1} = 1, N_t = k - 1, X_t = j) \times \right. \\
& \left. \mathbf{P}(X_t = j | N_1 = n, H_1 = h, \omega_{t+1} = 1, N_t = k - 1) \right] \\
&= \sum_j \left[\mathbf{P}(X_{t+1} = s | N_1 = n, \omega_{t+1} = 1, N_t = k - 1, X_t = j) \mathbf{P}(X_t = j | N_1 = n, \omega_{t+1} = 1, N_t = k - 1) \right] \\
&= \mathbf{P}(X_{t+1} = s | N_1 = n, \omega_{t+1} = 1, N_{t+1} = k).
\end{aligned}$$

Vale a terceira igualdade, porque

$$\begin{aligned}
& \mathbf{P}(X_{t+1} = s | N_1 = n, H_1 = h, \omega_{t+1} = 1, N_t = k - 1, X_t = j) = \mathbf{P}(\text{Bin}(j, 1 - c) = s) \\
&= \mathbf{P}(X_{t+1} = s | N_1 = n, \omega_{t+1} = 1, N_t = k - 1, X_t = j)
\end{aligned}$$

pela regra do acoplamento. Também,

$$\begin{aligned}
\mathbf{P}(X_t = j | N_1 = n, H_1 = h, \omega_{t+1} = 1, N_t = k - 1) &= \mathbf{P}(X_t = j | N_1 = n, H_1 = h, N_t = k - 1) \\
&= \mathbf{P}(X_t = j | N_1 = n, N_t = k - 1) \\
&= \mathbf{P}(X_t = j | N_1 = n, \omega_{t+1} = 1, N_t = k - 1).
\end{aligned}$$

Na primeira igualdade usei a independência de ω_m do passado, na segunda a hipótese de indução e na terceira novamente independência de ω_m do passado. Ou seja, provamos

$$\mathbf{P}(X_t = s | N_1 = n, H_1 = h, N_t = k) = \mathbf{P}(X_t = s | N_1 = n, N_t = k) \quad \forall t.$$

Suponha,

$$\mathbf{P}(X_t = s | N_i = n, H_i = h, N_t = k) = \mathbf{P}(X_t = s | N_i = n, N_t = k) \quad \forall t.$$

Para o caso $i + 1$ primeiro provarei que

$$\mathbf{P}(X_t = s | N_t = n, H_t = h, N_t = k) = \mathbf{P}(X_t = s | N_t = n, N_t = k)$$

para todo t . Se $n \neq k$, o resultado vale trivialmente. Para caso $t = 1$ provei no começo da demonstração. Seja valido que

$$\mathbf{P}(X_{t-1} = s | H_{t-1} = h, N_{t-1} = k) = \mathbf{P}(X_{t-1} = s | N_{t-1} = k).$$

Perceba que temos duas hipóteses. Então

$$\begin{aligned}
\mathbf{P}(X_t = s | H_t = h, N_t = k) &= \overbrace{\mathbf{P}(X_t = s | H_t = h, N_t = k, \omega_t = 1)}^{(III)} \overbrace{\mathbf{P}(\omega_t = 1 | H_t = h, N_t = k)}^{(I)} + \\
& \overbrace{\mathbf{P}(X_t = s | H_t = h, N_t = k, \omega_t = 0)}^{(II)} \mathbf{P}(\omega_t = 0 | H_t = h, N_t = k)
\end{aligned}$$

(I) - Queremos argumentar que $P(\omega_t = 1|H_t = h, N_t = k) = P(\omega_t = 1|N_t = k)$.
Então,

$$\begin{aligned}
P(H_t = h|\omega_t = 1, N_t = k) &= P(H_t = h|\omega_t = 1, N_{t-1} = k - 1) \\
&= \sum_j P(H_t = h|\omega_t = 1, N_{t-1} = k - 1, H_{t-1} = j)P(H_{t-1} = j|\omega_t = 1, N_{t-1} = k - 1) \\
&= \sum_j P(\text{Bin}(j, (1 - c)) = h)P(\text{Bin}(y - x, (1 - c)^{k-1}) = j) \\
&= P(\text{Bin}(y - x, (1 - c)^k) = h) = P(H_t = h|N_t = k)
\end{aligned}$$

Na segunda equação usamos que ω_m é independente do passado, relação de H_t com N_t e a regra de transição do acoplamento para H_t . Na terceira igualdade o Lema 2.2.1.

(II)

Utilizarei as propriedades de $\{\omega_t = 0\}$ e a hipótese de indução.

$$\begin{aligned}
P(X_t = s|H_t = h, N_t = k, \omega_t = 0) &= P(X_{t-1} = s - 1|H_{t-1} = h, N_{t-1} = k, \omega_t = 0) \\
&= P(X_{t-1} = s - 1|H_{t-1} = h, N_{t-1} = k) \\
&= P(X_{t-1} = s - 1|N_{t-1} = k) \\
&= P(X_t = s|N_t = k, \omega_t = 0).
\end{aligned}$$

(III)

$$\begin{aligned}
P(X_t = j|H_t = h, \omega_t = 1, N_t = k) &= P(X_t = j|H_t = h, \omega_t = 1, N_{t-1} = k - 1) \\
&= \sum_m P(X_t = j|H_t = h, \omega_t = 1, N_{t-1} = k - 1, X_{t-1} = m)P(X_{t-1} = m|H_t = h, \omega_t = 1, N_{t-1} = k - 1)
\end{aligned}$$

Mas,

$$\begin{aligned}
P(H_t = h|X_{t-1} = m, \omega_t = 1, N_{t-1} = k - 1) &= \\
\sum_j P(H_t = h|X_{t-1} = m, \omega_t = 1, N_{t-1} = k - 1, H_{t-1} = j)P(H_{t-1} = j|X_{t-1} = m, \omega_t = 1, N_{t-1} = k - 1) &= \\
= \sum_j P(H_t = h|\omega_t = 1, N_{t-1} = k - 1, H_{t-1} = j)P(H_{t-1} = j|\omega_t = 1, N_{t-1} = k - 1) &= \\
= P(H_t = h|\omega_t = 1, N_{t-1} = k - 1) &
\end{aligned}$$

Na segunda identidade usamos a independência de ω_m do passado, a hipótese de indução, e a regra de transição para H_t . Similar ao argumento dos casos anteriores.

Com isso, e usando que pela regra de transição do acoplamento

$$\begin{aligned} \mathbb{P}(X_t = j | H_t = h, \omega_t = 1, N_{t-1} = k-1, X_{t-1} = m) &= \mathbb{P}(X_t = j | \omega_t = 1, X_{t-1} = m) \\ &= \mathbb{P}(X_t = j | \omega_t = 1, N_{t-1} = k-1, X_{t-1} = m) \end{aligned}$$

Concluimos,

$$\mathbb{P}(X_t = j | H_t = h, \omega_t = 1, N_t = k) = \mathbb{P}(X_t = j | \omega_t = 1, N_{t-1} = k-1) = \mathbb{P}(X_t = j | \omega_t = 1, N_t = k).$$

Disso tudo, nós temos

$$\begin{aligned} \mathbb{P}(X_t = s | H_t = h, N_t = k) &= \mathbb{P}(X_t = s | N_t = k, \omega_t = 1) \mathbb{P}(\omega_t = 1 | N_t = k) + \\ &\mathbb{P}(X_t = s | N_t = k, \omega_t = 0) \mathbb{P}(\omega_t = 0 | N_t = k) = \mathbb{P}(X_t = s | N_t = k). \end{aligned}$$

Feito isso, vamos provar que

$$\mathbb{P}(X_t = s | N_{i+1} = n, H_{i+1} = h, N_t = k) = \mathbb{P}(X_t = s | N_{i+1} = n, N_t = k) \quad \forall t.$$

Para $t = i + 1$ já provei. Suponha válido para $t \geq i + 1$, isto é,

$$\mathbb{P}(X_t = s | N_{i+1} = n, H_{i+1} = h, N_t = k) = \mathbb{P}(X_t = s | N_{i+1} = n, N_t = k),$$

então

$$\begin{aligned} \mathbb{P}(X_{t+1} = s | N_{i+1} = n, H_{i+1} = h, N_{t+1} = k) &= \\ \mathbb{P}(X_{t+1} = s | N_{i+1} = n, H_{i+1} = h, N_{t+1} = k, \omega_{t+1} = 1) &\mathbb{P}(\omega_{t+1} = 1 | N_{i+1} = n, H_{i+1} = h, N_{t+1} = k) + \\ \mathbb{P}(X_{t+1} = s | N_{i+1} = n, H_{i+1} = h, N_{t+1} = k, \omega_{t+1} = 0) &\mathbb{P}(\omega_{t+1} = 0 | N_{i+1} = n, H_{i+1} = h, N_{t+1} = k). \end{aligned}$$

Usando independência de ω_m do passado e a hipótese de indução,

$$\begin{aligned} \mathbb{P}(X_{t+1} = s | N_{i+1} = n, H_{i+1} = h, N_{t+1} = k, \omega_{t+1} = 0) &= \mathbb{P}(X_t = s-1 | N_{i+1} = n, H_{i+1} = h, N_t = k) \\ &= \mathbb{P}(X_t = s-1 | N_{i+1} = n, N_t = k) = \mathbb{P}(X_{t+1} = s | N_{i+1} = n, N_{t+1} = k, \omega_{t+1} = 0). \end{aligned}$$

O outro termo,

$$\begin{aligned} \mathbb{P}(X_{t+1} = s | N_{i+1} = n, H_{i+1} = h, N_{t+1} = k, \omega_{t+1} = 1) &= \\ &= \mathbb{P}(X_{t+1} = s | N_{i+1} = n, H_{i+1} = h, N_t = k-1, \omega_{t+1} = 1) \\ &= \sum_r \left[\mathbb{P}(X_{t+1} = s | N_{i+1} = n, H_{i+1} = h, N_t = k-1, \omega_{t+1} = 1, X_t = r) \times \right. \\ &\left. \mathbb{P}(X_t = r | N_{i+1} = n, H_{i+1} = h, N_t = k-1, \omega_{t+1} = 1) \right] \\ &= \sum_r \mathbb{P}(X_{t+1} = s | N_{i+1} = n, N_t = k-1, \omega_{t+1} = 1, X_t = r) \mathbb{P}(X_t = r | N_{i+1} = n, N_t = k-1, \omega_{t+1} = 1) \\ &= \mathbb{P}(X_{t+1} = s | N_{i+1} = n, N_t = k-1, \omega_{t+1} = 1) = \mathbb{P}(X_{t+1} = s | N_{i+1} = n, N_{t+1} = k, \omega_{t+1} = 1). \end{aligned}$$

Na terceira igualdade usamos argumentos similares aos que usei antes, independência ω_m do passado, $t + 1 > i + 1$, hipótese de indução (temos duas hipóteses, usei a última delas), definição da regra de transição do acoplamento. Que

$$\mathbb{P}(\omega_{t+1} = 1 | N_{i+1} = n, H_{i+1} = h, N_{t+1} = k) = \mathbb{P}(\omega_{t+1} = 1 | N_{i+1} = n, N_{t+1} = k)$$

é imediato. Portanto, provamos

$$\mathbb{P}(X_{t+1} = s | N_{i+1} = n, H_{i+1} = h, N_{t+1} = k) = \mathbb{P}(X_{t+1} = s | N_{i+1} = n, N_{t+1} = k).$$

Logo, $\mathbb{P}(X_t = s | N_{i+1} = n, H_{i+1} = h, N_t = k) = \mathbb{P}(X_t = s | N_{i+1} = n, N_t = k)$ para todo $t \geq i + 1$. Para $i + 1 > t$

$$\begin{aligned} & \mathbb{P}(H_{i+1} = h | N_{i+1} = n, X_t = s, N_t = k) = \\ & \mathbb{P}(H_{i+1} = h | N_{i+1} = n, X_t = s, N_t = k, \omega_{i+1} = 1) \mathbb{P}(\omega_{i+1} = 1 | N_{i+1} = n, X_t = s, N_t = k) + \\ & \mathbb{P}(H_{i+1} = h | N_{i+1} = n, X_t = s, N_t = k, \omega_{i+1} = 0) \mathbb{P}(\omega_{i+1} = 0 | N_{i+1} = n, X_t = s, N_t = k). \end{aligned}$$

Usando as propriedades de ω_m e a hipótese de indução (a primeira delas).

$$\begin{aligned} & \mathbb{P}(H_{i+1} = h | N_{i+1} = n, X_t = s, N_t = k, \omega_{i+1} = 0) = \mathbb{P}(H_i = h | N_i = n, X_t = s, N_t = k) \\ & = \mathbb{P}(H_i = h | N_i = n, N_t = k) = \mathbb{P}(H_{i+1} = h | N_{i+1} = n, N_t = k, \omega_{i+1} = 0). \end{aligned}$$

Usando a regra de transição do acoplamento e as propriedades de ω_m .

$$\begin{aligned} & \mathbb{P}(H_{i+1} = h | N_{i+1} = n, X_t = s, N_t = k, \omega_{i+1} = 1) \\ & = \sum_r \left[\mathbb{P}(H_{i+1} = h | X_t = s, N_i = n - 1, H_i = r, \omega_{i+1} = 1, N_t = k) \times \right. \\ & \left. \mathbb{P}(H_i = r | X_t = s, N_i = n - 1, \omega_{i+1} = 1, N_t = k) \right] \\ & = \sum_r \mathbb{P}(H_{i+1} = h | N_i = n - 1, H_i = r, \omega_{i+1} = 1, N_t = k) \mathbb{P}(H_i = r | N_i = n - 1, \omega_{i+1} = 1, N_t = k) \\ & = \mathbb{P}(H_{i+1} = h | N_{i+1} = n, N_t = k, \omega_{i+1} = 1). \end{aligned}$$

Ou seja, provamos que $\forall t$ e $\forall i$.

$$\begin{aligned} & \mathbb{P}(X_t = s | N_i = n, H_i = h, N_t = k) = \mathbb{P}(X_t = s | N_i = n, N_t = k) \\ & \iff \\ & \mathbb{P}(H_i = h | N_i = n, X_t = s, N_t = k) = \mathbb{P}(H_i = h | N_i = n, N_t = k). \end{aligned}$$

Com isso, note $i < t$

$$\begin{aligned} & \mathbb{P}(X_t = s | H_i = h, N_t = k) = \sum_n \mathbb{P}(X_t = s | H_i = h, N_t = k, N_i = n) \mathbb{P}(N_i = n | H_i = h, N_t = k) \\ & \stackrel{*}{=} \sum_n \mathbb{P}(X_t = s | N_t = k, N_i = n) \mathbb{P}(N_i = n | N_t = k) = \mathbb{P}(X_t = s | N_t = k) \end{aligned}$$

*- Note, esse argumento é falso, pois em geral $\mathbf{P}(N_i = n|H_i = h, N_t = k) \neq \mathbf{P}(N_i = n|N_t = k)$.

Mas suponha $i > t$

$$\begin{aligned} \mathbf{P}(H_i = h|X_t = s, N_t = k) &= \sum_n \mathbf{P}(H_i = h|X_t = s, N_t = k, N_i = n)\mathbf{P}(N_i = n|X_t = s, N_t = k) \\ &= \sum_n \mathbf{P}(H_i = h|N_t = k, N_i = n)\mathbf{P}(N_i = n|N_t = k) = \mathbf{P}(H_i = h|N_t = k) \end{aligned}$$

$$\begin{aligned} \mathbf{P}(N_i = n|X_t = s, N_t = k) &= \frac{\mathbf{P}(N_t = k, X_t = s, N_i = n)}{\mathbf{P}(X_t = s, N_t = k)} \\ &= \frac{\mathbf{P}(\sum_{m=t+1}^i \omega_m = n - k, X_t = s, N_t = k)}{\mathbf{P}(X_t = s, N_t = k)} \\ &= \mathbf{P}\left(\sum_{m=t+1}^i \omega_m = n - k\right) \\ &= \frac{\mathbf{P}(\sum_{m=t+1}^i \omega_m = n - k)\mathbf{P}(N_t = k)}{\mathbf{P}(N_t = k)} \\ &= \frac{\mathbf{P}(\sum_{m=t+1}^i \omega_m = n - k, N_t = k)}{\mathbf{P}(N_t = k)} \\ &= \frac{\mathbf{P}(N_t = k, N_i = n)}{\mathbf{P}(N_t = k)}. \end{aligned}$$

Provando que H_i, X_t são independentes dado N_t apenas quando $i \geq t$.

QED