Universidade Federal do ABC

Mestrado Profissional em Matemática em Rede Nacional

Vanessa Aparecida de Rezende Possebon

Probabilidade nos Inteiros:
Passeio Aleatório e outros processos

Santo André, 2023

Universidade Federal do ABC

PROFMAT

Vanessa Aparecida de Rezende Possebon

Probabilidade nos Inteiros

Passeio Aleatório e outros processos

Orientador: Prof. Dr. Prof. Rafael de Mattos Grisi

Dissertação de mestrado apresentada ao Centro de Matemática, Computação e Cognição para obtenção do título de Mestre

Este exemplar corresponde a versão final da dissertação defendida pela aluna Vanessa Aparecida de Rezende Possebon, e orientada pelo Prof. Dr. Prof. Rafael de Mattos Grisi.

Santo André, 2023

Sistema de Bibliotecas da Universidade Federal do ABC
Elaborada pelo Sistema de Geração de Ficha Catalográfica da UFABC
com os dados fornecidos pelo(a) autor(a).

de Rezende Possebon, Vanessa Aparecida
Probabilidade nos Inteiros : Passeio Aleatório e Outros Processos / Vanessa Aparecida de Rezende Possebon. — 2023.

123 fls. : il.

Orientador: Rafael de Mattos Grisi

Dissertação (Mestrado) — Universidade Federal do ABC, Mestrado Profissional em Matemática em Rede Nacional - PROFMAT, Santo André, 2023.

1. probabilidade. 2. passeio aleatório simples. 3. ruína do jogador. I. de Mattos Grisi, Rafael. II. Mestrado Profissional em Matemática em Rede Nacional - PROFMAT, 2023. III. Título.

Este exemplar foi revisado e alterado em relação à versão original, de acordo com as observações levantadas pela banca examinadora no dia da defesa, sob responsabilidade única da autora e com a anuência do orientador.

MINISTÉRIO DA EDUCAÇÃO
Fundação Universidade Federal do ABC
Avenida dos Estados, 5001 – Bairro Santa Terezinha – Santo André – SP
CEP 09210-580 · Fone: (11) 4996-0017

FOLHA DE ASSINATURAS

Assinaturas dos membros da Banca Examinadora que avaliou e aprovou a Defesa de Dissertação de Mestrado da candidata, VANESSA APARECIDA DE REZENDE POSSEBON realizada em 14 de Dezembro de 2023:

gov.br Documento assinado digitalmente
EDUARDO GUERON
Data: 15/12/2023 14:36:22-0300
Verifique em https://validar.iti.gov.br

Prof.(a) EDUARDO GUERON
UNIVERSIDADE FEDERAL DO ABC

Prof.(a) GLEICIANE DA SILVA ARAGÃO
UNIVERSIDADE FEDERAL DE SÃO PAULO

gov.br Documento assinado digitalmente
GABRIELA COTRIM DE MORAES
Data: 15/12/2023 12:37:00-0300
Verifique em https://validar.iti.gov.br

Prof.(a) GABRIELA COTRIM DE MORAES
INSTITUTO FEDERAL DE SÃO PAULO

Prof.(a) JERONIMO CORDONI PELLEGRINI
UNIVERSIDADE FEDERAL DO ABC

gov.br Documento assinado digitalmente
RAFAEL DE MATTOS GRISI
Data: 15/12/2023 12:06:53-0300
Verifique em https://validar.iti.gov.br

Prof.(a) RAFAEL DE MATTOS GRISI
UNIVERSIDADE FEDERAL DO ABC - Presidente

* Por ausência do membro titular, foi substituído pelo membro suplente descrito acima: nome completo, instituição e assinatura

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoas de Nível Superior - Brasil (CAPES) - Código de Financiamento 001

AGRADECIMENTOS

A toda a minha amada família e, em especial, ao meu esposo e ao meu filho, por me apoiarem e estarem comigo neste e em outros caminhos dificultosos da vida.

Ao meu orientador, Dr. Rafael de Mattos Grisi, por me guiar com tanta dedicação e conhecimento na elaboração desta dissertação, e me acolher com sua tranquilidade nos momentos em que a ansiedade me sufocava.

Aos professores do programa PROFMAT da Universidade Federal do ABC, os quais estruturaram com excelência a minha trajetória durante o Mestrado em Matemática.

Aos meus colegas de turma, cujo apoio mútuo e troca de ideias enriqueceram nossa jornada acadêmica e nos mantiveram perseverantes quando a vontade de desistir bateu à porta.

Aos meus queridos amigos, que me entretiveram nos momentos em que a fuga dos livros de matemática se fez necessária.

Aos amigos da Escola Estadual Padre Afonso Paschotte, pelo incentivo e pela torcida.

E finalmente, ao patrocínio da Capes durante meus estudos de Mestrado, e a todo aquele que direta ou indiretamente contribuiu para o desenvolvimento deste trabalho.

RESUMO

Neste trabalho, estudamos a teoria dos passeios aleatórios simples e do processo de ramificação de Bienaymé-Galton-Watson, ambos a tempo discreto, e apresentamos uma proposta de aplicação desse estudo em uma sequência didática para o ensino de probabilidade no ensino básico. Inicialmente abordamos ferramentas e conceitos probabilísticos necessários para a análise dos nossos objetos de estudo. Em seguida, foi a vez dos passeios aleatórios simples, introduzidos a partir do Problema da ruína do jogador, caso particular dos passeios aleatórios simples, onde provamos resultados referentes à probabilidade da ruína e à duração esperada do jogo, para depois usá-los para provar outros resultados mais gerais, como a probabilidade e o tempo esperado do retorno à origem dos passeios aleatórios simples. Na sequência, tratamos do processo de ramificação de Bienaymé-Galton-Watson, buscando demonstrar resultados que garantem as condições sobre as quais teremos ou não a extinção do processo. E por fim, apresentamos uma proposta de sequência didática cuja peça central é o jogo O Prisioneiro de Monty Hall, construído com o objetivo de gamificar os problemas da Ruína do Jogador e de Monty Hall.

Palavras-chave: probabilidade, passeios aleatórios simples, processo de Bienaymé-Galton-Watson, extinção, O Problema da Ruína do Jogador, O Problema de Monty Hall.

ABSTRACT

In this work, we study the theory of simple random walks and the Bienaymé-Galton-Watson branching process, both in discrete time, and we present a proposal for applying this study in a didactic sequence for teaching probability in basic education. Initially, we approach probabilistic tools and concepts necessary for the analysis of our study objects. Next, it was the turn of simple random walks, introduced from the Player's Ruin Problem, a particular case of simple random walks, where we prove results referring to the probability of ruin and the expected duration of the game, to later use them to prove other more general results, such as the probability and expected time of return to the origin of simple random walks. Next, we deal with the Bienaymé-Galton-Watson branching process, seeking to demonstrate results that guarantee the conditions under which we will or will not have the extinction of the process. And finally, we present a proposal for a didactic sequence whose centerpiece is the game The Prisoner of Monty Hall, built with the aim of gamifying the problems of the Player's Ruin and Monty Hall.

Keywords: probability, simple random walks, Bienaymé-Galton-Watson process, extinction, The Gambler's Ruin Problem ,The Monty Hall Problem.

CONTEÚDO

1

INTRODUÇÃO

O envelhecimento populacional é um fenômeno mundial que tem despertado crescente preocupação em diversos países. Além do aumento de gastos públicos com o aumento da demanda por sistemas de saúde e pensões, o motivo da preocupação está na escassez da força de trabalho que decorre do declínio da população jovem, impactando diretamente a economia de um país [5]. Ainda segundo a reportagem da BBC NEWS Brasil em 2019, "(...) o Fundo Monetário Internacional (FMI) alertou que a economia do Japão poderia encolher mais de 25% nos próximos 40 anos devido ao envelhecimento da população."

O Japão vem introduzindo desde 1990 políticas para reverter uma das principais causas do aumento proporcional de idosos na população, a queda nas taxas de natalidade. Em 2000, foi a vez da Coreia do Sul. E recentemente, temos a China, famosa pelas rígidas restrições à natalidade adotadas no passado, juntando-se ao clube [13].

Mas essa não é uma preocupação somente dos países asiáticos, O insuficiente número de bebês para manter o número de pessoas estável é também uma realidade de alguns países europeus, como Finlândia, suécia, Noruega, Estônia, Itália e França [14].

Este cenário é um dos mais atuais fatores que evidenciam a necessidade do estudo de ferramentas e modelos matemáticos que possam fundamentar estudos sobre o desenvolvimento de populações e suas consequências. Entre esses modelos matemáticos, encontram-se os processos de ramificação de Bienaymé-Galton-Watson e as cadeias de nascimento e morte, as quais são uma generalização do processo conhecido como passeio aleatório simples. Estes processos são modelos probabilísticos que se estabeleceram como ferramentas importantes para a verificação da sobrevivência ou não de uma população ao longo do tempo.

Portanto, indaga-se: em relação às análises matemáticas, quais condições indicam que uma determinada população está a caminho da extinção?

O objetivo geral do presente trabalho é estudar os processos de ramificação de Bienaymé-Galton-Watson e os passeios aleatórios simples, a fim de responder à questão levantada, assim como fomentar o estudo desses processos entre os estudantes da rede nacional de educação, já que são poucos os livros em português que os abordam, visando a formação de futuros estudiosos brasileiros do assunto e a contribuição destes na formulação de políticas públicas efetivas para essa problemática que está em escalada e atinge cada vez mais países mundo afora.

Para tanto, foram delineados os seguintes objetivos específicos: estudar ferramentas e conceitos da probabilidade; aplicar esse estudo para obter resultados relevantes à problemática apresentada; planejar uma sequência didática para introduzir alguns destes estudos no ensino médio.

Parte-se da hipótese de que os indicadores matemáticos, que evidenciam se uma determinada população está ou não caminhando em direção a sua extinção, podem ser relacionados com a probabilidade de extinção do processo de ramificação de Bienaymé-Galton-Watson, uma vez que a extinção neste contexto significa o momento a partir do qual não serão mais gerados novos elementos no processo. Já nos passeios aleatórios simples, tais indicadores teriam uma correspondência com a probabilidade do passeio nos inteiros não negativos chegar ao zero. Neste contexto, os inteiros representariam o número de indivíduos da população.

Assim, para viabilizar o teste da hipótese, realizamos uma pesquisa de finalidade básica pura, com abordagem quantitativa feita através de procedimentos bibliográficos.

Assim, no próximo e segundo capítulo, começamos abordando as definições básicas da Teoria de Probabilidades, seguidas pelos conceitos de variáveis aleatórias e esperança matemática, e finalizamos aprofundando na Teoria de Probabilidades através das funções geradoras de probabilidade. Todos estes conceitos estão acompanhados por demonstrações de suas principais propriedades e proposições, as quais estruturarão a análise dos resultados centrais dessa pesquisa.

No terceiro capítulo, introduzimos o primeiro objeto central deste estudo, os passeios aleatórios simples, através de um importante problema para a teoria de probabilidades, o problema da ruína do jogador. Nele buscamos provar resultados chaves para o presente trabalho, como a probabilidade da ruína do jogador e a duração esperada do jogo, os quais servirão de fundamento para obtermos resultados mais gerais acerca dos passeios aleatórios simples, tais quais a probabilidade e o tempo esperado do seu retorno à origem.

No quarto capítulo, iniciamos detalhando os processos de ramificação de Bienaymé- Galton-Watson, o nosso segundo ponto central de estudo, visando tornar sua definição em linguagem matemática, apresentada na sequência, mais clara aos possíveis leitores que estudarão esse assunto através deste trabalho e que, porventura, possam achar essa definição de difícil compreensão. Em seguida, provamos algumas proposições importantes acerca das esperanças e das funções geradoras de probabilidades desses processos, as quais foram necessárias para obter outro resultado chave para esta dissertação: a probabilidade de extinção dos processos de ramificação de Bienaymé- Galton- Watson. Nós buscamos deduzi-la tanto algebricamente quanto utilizando representações gráficas, a fim de proporcionar uma melhor compreensão desse importante Teorema da teoria de ramificação.

No quinto e último capítulo, apresentamos uma proposta de sequência didática que utiliza conceitos da Probabilidade e dos passeios aleatórios simples abordados neste trabalho. Essa proposta gira em torno do jogo O Prisioneiro de Monty Hall, que tem como base dois famosos problemas da matemática: o *Problema da Ruína do Jogador e o Problema de Monty Hall*. O primeiro, como já mencionamos, é um dos temas que será discutido no terceiro capítulo, enquanto o segundo é famoso por contrariar a intuição de grande parte das pessoas. Nossa sequência didática é dividida em seis aulas e foi construída com intuito de introduzir conceitos desse estudo no ensino médio, visando as novas demandas, principalmente no campo da Probabilidade, que surgiram com a recente reforma curricular, desencadeada pela publicação da Base Nacional Comum Curricular em 2018 [1].

2

PROBABILIDADE NOS NÚMEROS INTEIROS

Neste capítulo apresentaremos um resumo contendo as definições básicas da Teoria de Probabilidades e Variáveis Aleatórias, cujos resultados e conclusões serão utilizados nos capítulos posteriores.

Na primeira parte traremos conceitos básicos da teoria de probabilidades. Apresentaremos tais conceitos de modo simples, de modo similar ao explorado por diversos autores, como [16, 18].

Na segunda parte do capítulo traremos uma abordagem diferente, e um pouco mais simples, daquela normalmente encontrada na literatura, e nos focaremos em variáveis aleatórias com valores inteiros.

Esta é uma apresentação similar à apresentada por Schinazi em [20], e tem a vantagem de conversar de maneira mais direta com o conteúdo do resto do trabalho, onde estudaremos processos que assumem apenas valores inteiros.

2.1 DEFININDO UMA PROBABILIDADE

Talvez a primeira pergunta que devemos responder seria "o que é probabilidade?". Este é um dos conceitos que, a priori, todos tem a sensação de saber o que é, mas poucos conseguem oferecer algum tipo de definição consistente. Não raramente ouvimos respostas como "é a chance de algo acontecer", que apenas troca um palavra por um sinônimo, sem realmente esclarecer o conceito.

Uma das maneiras de pensar, e que seguiremos neste texto, é tomar probabilidade como um "modelo matemático que tenta medir e representar as incertezas associadas a uma certa observação".

Tal definição, apesar de imprecisa, nos oferece uma intenção clara e um caminho a seguir.

Neste sentido, vejamos algumas definições fundamentais que darão o pontapé inicial na construção do que entenderemos por **Probabilidade**.

Primeiro precisamos modelar a *fonte das incertezas que queremos medir*.

Assim, chamamos de *experimento aleatório* qualquer experimento ou observação que produz um resultado que não pode ser predeterminado com certeza, tais quais: lançamento de uma moeda ou de um dado e observar o resultado da face de cima; selecionar uma carta de um baralho e observar seu naipe; sortear bolas enumeradas em uma urna e observar a sequência numérica obtida, entre outros.

Embora não saibamos com antecedência qual resultado de um experimento aleatório, em geral podemos descrever o conjunto de todos os resultados possíveis que podem ocorrer. Nomeamos esse conjunto por *espaço amostral*, que em geral será representado por Ω.

Neste trabalho, vamos considerar somente os casos com um espaço amostral finito ou infinito enumerável.

Os subconjuntos de Ω são chamados de *eventos*. Diremos que um evento ocorre quando o resultado do experimento aleatório pertence ao evento.

Convencionalmente, representamos um evento por uma letra maiúscula do alfabeto, e dizemos que dois eventos A e B contidos em Ω são disjuntos se, e somente se, $A \cap B = \emptyset$.

Exemplo 2.1. Uma moeda será lançada 3 vezes, e será observado a sequência de caras (ca) e coroas (co). Assim, o espaço amostral associado a esse experimento aleatório é

$$\Omega = \{(ca, ca, ca); (ca, ca, co); (ca, co, ca); (co, ca, ca); (ca, co, co); (co, ca, co); (co, co, ca), (co, co, co)\}$$

Eis alguns eventos pertencentes a Ω:

$A = \{(ca, ca, co); (ca, co, ca); (co, ca, ca)\}$, que corresponde ao evento "ocorrência de exatamente uma coroa” no experimento.

$B = \{(ca, ca, ca); (ca, ca, co); (ca, co, ca); (co, ca, ca)\}$, que corresponde ao evento "ocorrência de pelo menos duas caras" no experimento.

$C = \{(co, co, co)\}$, que corresponde ao evento "ocorrência de três coroas" no experimento.

Observemos que $A \cap C = \emptyset$, já que não é possível obter um resultado no experimento que pertença a ambos os eventos, e que $A \cap B = A$. Logo, os eventos A e C são disjuntos, enquanto os eventos A e B não o são.

Para definir a função de Probabilidade, associaremos a cada evento de $A \subset \Omega$ um valor $P(A) \in \mathbb{R}$, que chamaremos de *probabilidade do evento* A e que traduzirá nossa confiança na capacidade do evento ocorrer.

Mas como definir tal função?? Que propriedades ela deve ter?

Um modo bastante comum, e até útil, de se pensar é o *frequentista*. Para isso repetimos o experimento aleatório uma quantidade grande de vezes, e contamos o total de vezes nas quais o resultado observado pertence a um dado evento A. A ideia então é pensar que a probabilidade de A seria a proporção de vezes que observamos A ao longo destas N repetições.

A expectativa é que tal proporção (que claramente depende de N) se aproxime de um número real quando N tende ao infinito. Tal valor seria então chamado de *probabilidade de* A. Ou seja

$$P(A) := \lim_{N \to \infty} \frac{\#A}{N},$$

onde $\#A$ representa o total de vezes que observamos A.

Para entender melhor essa ideia, pense no lançamento de uma moeda. Se moeda for não viciada, a expectativa é que após um total grande de lançamentos, aproximadamente metade destes mostraram cara enquanto a outra metade mostrou coroa.

Exemplo 2.2. Agora, suponhamos o lançamento de um dado não honesto. Assumindo que o lançamento tenha sido repetido 600 vezes e que em 500 destas vezes se tenha observado a ocorrência do número 1. Dessa forma, a probabilidade de ocorrência do número 1 é dada por sua frequência relativa, ou seja, $P(\{1\}) = \frac{500}{600} = \frac{5}{6}$.

A definição acima, apesar de nos permitir uma compreensão um pouco mais clara do que é probabilidade, ainda carrega uma série de problemas e imprecisões que dificultam sua utilização.

Podemos apontar, por exemplo, a ideia de *repetir o experimento aleatório* N vezes. O que realmente significa isso?? Ou ainda, o problema da existência de tais limites, que não parece clara e nem óbvia.

E mesmo que exista, não sabemos qual o nível de dependência das realizações específicas dos experimentos realizados. Ou seja, o que nos garante que no exemplo acima sempre observaremos o número 1 aproximadamente $5/6$ das vezes sempre que repetirmos o experimento?

E pensar que a chance de tal coisa acontecer é pequena trás uma circularidade fatal para a definição que estamos buscando.

Uma ideia para contornar estes problemas é pensar a partir das **propriedades** que esperamos observar em uma função que mede a probabilidade de um evento. E para isso, a noção frequentista descrita acima, pode nos ser bastante útil.

A partir desta ideia podemos determinar, por exemplo, que a probabilidade $P(A)$ de um evento A deve ser um número entre 0 e 1, representando o nível de confiança que temos que tal evento ocorrerá. Se temos a certeza da não ocorrência do evento A, então $P(A) = 0$. À medida que aumenta a confiança na observação de A no resultado do experimento, $P(A)$ também aumenta, até ao ponto em que $P(A) = 1$ corresponderá à certeza da ocorrência de A.

Fazendo tal análise com cuidado chegamos à seguinte definição.

Definição 2.1. Uma função de probabilidade associa a cada evento A um número $P(A)$, o qual representa a probabilidade de A ocorrer, de forma que:

1. Para todo evento $A \subset \Omega$, devemos ter $0 \leq P(A) \leq 1$;
2. $P(\Omega) = 1$
3. Se $A_1, A_2, A_3, \ldots$ são eventos 2 a 2 disjuntos, então

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$$

Apesar de usarmos aqui a definição canônica de probabilidade, seria possível começar tomando $\Omega = \{x_1, x_2, \ldots\}$ e definir P a partir dos valores $P(\{x_k\}), k \geq 1$, representado as probabilidades de observarmos cada um dos resultados do nosso experimento.

Neste caso definiríamos $P(A)$ pela soma das probabilidades de seus elementos. Ou seja, se $A = \{x_{i_1}, x_{i_2}, \ldots\}$, então

$$P(A) = \sum_k P(\{x_{i_k}\}).$$

Quando todos os eventos unitários de um espaço amostral finito têm a mesma probabilidade de ocorrer, dizemos que o modelo probabilístico é *equiprovável*. Nesse caso, se $\Omega = \{x_1, x_2, \ldots, x_n\}$, temos que $P(\{x_1\}) = P(\{x_2\}) = \ldots = P(\{x_n\}) = k$. Daí, segue que

$$1 = P(\Omega) = P(\{x_1, x_2, \ldots, x_n\} = P(\{x_1\} \cup \{x_2\} \cup \ldots \cup \{x_n\})$$

Notemos que os eventos unitários $\{x_i\}$, para $i = 1, 2, \ldots, n$, são 2 a 2 disjuntos. Logo,

$$1 = P(\{x_1\} \cup \{x_2\} \cup \ldots \cup \{x_n\}) = P(\{x_1\}) + P(\{x_2\}) + \ldots + P(\{x_n\}) = nk$$

$$\Leftrightarrow k = \frac{1}{n}, \quad \text{para} \quad n \neq 0$$

e, portanto, $P(x_1) = P(x_2) = \ldots = P(x_n) = \frac{1}{n}$.

Analogamente, nesse modelo, se um evento $A \subset \Omega$ possui j elementos, então

$$P(A) = j \cdot \frac{1}{n} = \frac{j}{n} = \frac{\text{número de elementos de } A}{\text{número de elementos de } \Omega}.$$

Exemplo 2.3. Um dado honesto será lançado e é observado o número da face de cima. A *honestidade* do dado nos informa que nosso nível de confiança deve ser igual para todos os resultados. Dessa forma, todos os números contidos nas faces do dado são equiprováveis e

$$\Omega = \{1, 2, 3, 4, 5, 6\} \text{ e } P(\{1\}) = P(\{2\}) = P(\{3\}) = P(\{4\}) = P(\{5\}) = P(\{6\}) = \frac{1}{6}$$

Nesse caso, se $A = \{2, 3, 5\}$, então

$$P(A) = \frac{\text{número de resultados em } A}{\text{número de resultados em } \Omega} = \frac{3}{6}$$

que corresponde a probabilidade de ocorrer de um número primo no experimento.

2.1.1 *Algumas propriedades da função de probabilidade*

Adiante, demonstraremos algumas propriedades da probabilidade necessárias para as discussões dos capítulos posteriores.

Proposição 2.2. *Se A e B são eventos de Ω, então:*

i) $P(A^c) = 1 - P(A)$.

ii) $P(\emptyset) = 0$.

iii) $P(A \backslash B) = P(A) - P(A \cap B)$.

iv) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

v) *Se $B \subset A$ então $P(A \cap B^c) = P(A) - P(B)$.*

vi) *Se $B \subset A$ então $P(A) \geq P(B)$.*

Demonstração. De fato,

i) como temos que $P(\Omega) = 1$, $\Omega = A \cup A^c$ e que A e A^c são eventos disjuntos, então

$$1 = P(A \cup A^c) = P(A) + P(A^c) \Rightarrow P(A^c) = 1 - P(A).$$

ii) fazendo $A = \Omega$ no item anterior, temos $A^c = \emptyset$ e

$$P(\emptyset) = 1 - P(\Omega) = 1 - 1 = 0.$$

iii) usando o fato que $P(A) = P((A \backslash B) \cup (A \cap B))$ e que $(A \backslash B)$ e $(A \cap B)$ são disjuntos, então

$$P(A) = P(A \backslash B) + P(A \cap B) \Rightarrow P(A \backslash B) = P(A) - P(A \cap B). \tag{2.1}$$

iv) diante da igualdade $P(A \cup B) = P((A \backslash B) \cup B)$ e de que $(A \backslash B)$ e B são disjuntos, logo, usando a equação (2.1), temos que

$$P(A \cup B) = P(A \backslash B) + P(B) = P(A) + P(B) - P(A \cap B).$$

v) se $B \subset A$, então $A \cap B = B$. Como $A \cap B^c = A \backslash B$ e usando novamente a equação (2.1), segue que

$$P(A \cap B^c) = P(A \backslash B) = P(A) - P(A \cap B) = P(A) - P(B). \tag{2.2}$$

vi) se $B \subset A$, então, por (v) obtemos

$$P(A) - P(B) = P(A \cap B^c) \geq 0 \Rightarrow P(A) \geq P(B),$$

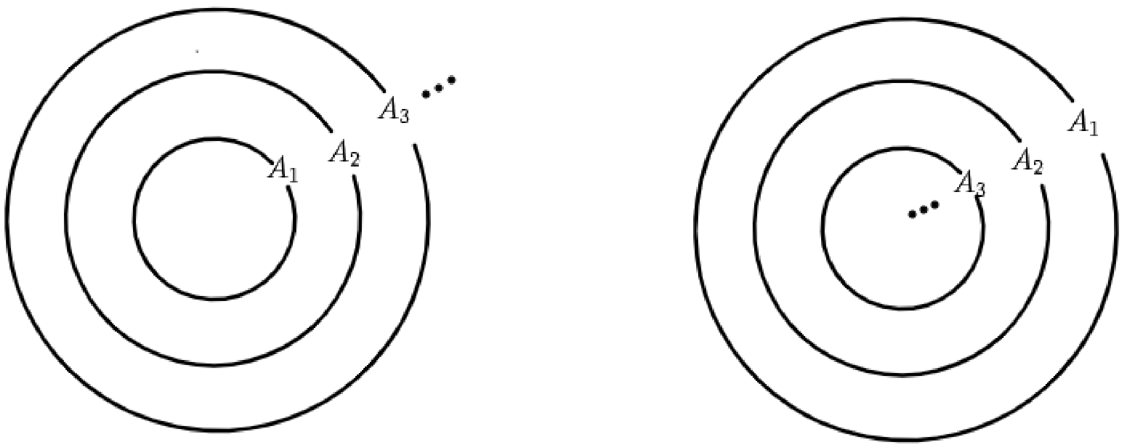
E assim, finalizamos a demonstração da Proposição 2.2.

$\square$

Antes de partirmos para as demonstrações das proposições que envolvem sequências de eventos de Ω, precisamos definir o seguinte conceito.

Dizemos que uma sequência de eventos $A_1, A_2, A_3, \ldots$ é crescente se $A_n \subset A_{n+1}$ para todo $n \geq 1$. E dizemos que é decrescente se $A_n \supset A_{n+1}$ para todo $n \geq 1$, conforme podemos vizualizar na figura 1.

Figura 1: Sequência de eventos de Ω



(a) Crescente. (b) Decrescente.

Dito isso, demonstraremos duas proposições que serão de grande importância ao longo deste trabalho.

Proposição 2.3. *Seja $(A_n)_{n\geq 1}$ uma sequência de eventos de Ω.*

I) Se $(A_n)_{n\geq 1}$ é crescente, então $\lim_{n\to\infty} P(A_n) = P(\bigcup_{n\geq 1} A_n)$.

II) Se $(A_n)_{n\geq 1}$ é decrescente, então $\lim_{n\to\infty} P(A_n) = P(\bigcap_{n\geq 1} A_n)$.

Demonstração. Pois bem,

I) defina $B_1 = A_1$ e $B_n = A_n \backslash A_{n-1}$. Notemos que, se $(A_n)_{n\geq 1}$ é crescente, então

$$\bigcup_{k=1}^{n} B_k = A_n = \bigcup_{k=1}^{n} A_k$$

e que $B_i \cap B_j = \emptyset$ para todo $i \neq j$, ou seja, os eventos $B_1, B_2, \ldots$ são dois a dois disjuntos.

Sendo assim,

$$\lim_{n\to\infty} P(A_n) = \lim_{n\to\infty} P(\bigcup_{k=1}^{n} B_k) = \lim_{n\to\infty} \sum_{k=1}^{n} P(B_k) = \sum_{k=1}^{\infty} P(B_k) = P(\bigcup_{k=1}^{\infty} B_k) = P(\bigcup_{k=1}^{\infty} A_k).$$

Logo,

$$\lim_{n\to\infty} P(A_n) = P(\bigcup_{n\geq 1} A_n),$$

concluindo a demonstração para o caso (i).

II) notemos que, se $(A_n)_{n\geq 1}$ é decrescente, então $(A_n^c)_{n\geq 1}$ é crescente.

Dessa forma e por (i), temos que

$$\lim_{n\to\infty} P(A_n^c) = P(\bigcup_{n\geq 1} A_n^c) = P((\bigcap_{n\geq 1} A_n)^c) = 1 - P(\bigcap_{n\geq 1} A_n)$$

$$\Leftrightarrow \lim_{n\to\infty} P(A_n) = P(\bigcap_{n\geq 1} A_n),$$

concluindo a demonstração para o caso (ii)

□

2.1.2 *Probabilidade Condicional e Independência*

Consideremos novamente a experiência que consiste no lançamento de um dado honesto. Sabemos que a probabilidade de ocorrer um número primo no lançamento é

$$P(A) = P(\{2,3,5\}) = \frac{3}{6} = \frac{1}{2},$$

conforme descrito no exemplo 2.3. Agora suponhamos que, realizada a experiência, alguém nos informe que o resultado foi um número maior que 3, ou seja, que o evento $B = \{4,5,6\}$ ocorreu.

Diante dessa condição, o nosso nível de incerteza na ocorrência do evento A ocorrer é alterada. De fato, o evento A ocorre agora somente se o resultado 5 ocorreu, já que sabemos de antemão que o resultado do lançamento foi 4, 5 ou 6. Assim, com estamos diante de um dado honesto, podemos dizer que a *probabilidade do evento A dado que o evento B ocorreu* é $P(A|B) := \frac{1}{3}$.

Detalhando o argumento acima, notemos que a informação sobre a ocorrência do evento B faz com que o espaço amostral seja restringido para B, com os resultados favoráveis à ocorrência de um evento A pertencentes à $A \cap B$. Logo, como estamos em um espaço amostral Ω equiprovável, podemos expressar

$$P(A|B) = \frac{\#(A \cap B)}{\#(B)},$$

onde $\#(I)$ corresponde ao número de elementos em I.

Ajustando os valores acima em termos da frequência relativa, obtemos que

$$P(A|B) = \frac{\#(A \cap B)}{\#(B)} = \frac{\frac{\#(A\cap B)}{\#(\Omega)}}{\frac{\#(B)}{\#(\Omega)}} = \frac{P(A \cap B)}{P(B)}.$$

Deste modo podemos utilizar o mesmo raciocínio em espaços amostrais não equiprováveis e chegar assim a essa mesma conclusão.

Sendo assim, a definição matemática para a probabilidade condicional é dada da seguinte forma.

Definição 2.4. Seja Ω um espaço amostral e $A, B \subset \Omega$ eventos com $P(B) > 0$. Nestas condições definimos a *probabilidade condicional de A dada a ocorrência de B* por

$$P(A|B) = \frac{P(A \cap B)}{P(B)}. \tag{2.3}$$

Em algumas ocasiões, usaremos a fórmula acima na forma conhecida como *Regra do Produto*:

$$P(A \cap B) = P(A|B) \cdot P(B) \tag{2.4}$$

Analogamente, se $P(A) > 0$, então temos que $P(A \cap B) = P(B|A) \cdot P(A)$.

Exemplo 2.4. Para a realização de uma pesquisa de saúde em uma certa população, sorteamos uma pessoa ao acaso e perguntamos sua idade e se é ou não fumante. Considere que a população em questão tem a seguinte configuração:

	Fumante	Não Fumante
18 anos ou menos	5%	20%
entre 19 e 45 anos	7%	41%
mais de 45 anos	13%	14%

Considerando agora os eventos $A = \{\text{fumante}\}$ e $B = \{\text{mais de 45 anos}\}$ temos

$$P(A) = 0{,}25 \qquad P(B) = 0{,}27 \qquad P(A \cap B) = 0{,}13,$$

e portanto

$$P(A|B) = \frac{0{,}13}{0{,}27} = 0{,}4814 \quad \text{e} \quad P(B|A) = \frac{0{,}13}{0{,}25} = 0{,}52.$$

2.1.3 *Independência*

Vamos introduzir agora um conceito central na teoria da probabilidade: a independência.

Informalmente, podemos pensar que um evento A é independente de outro evento B quando *a ocorrência de B não traz nenhuma informação relevante sobre A*. Mais formalmente diremos que o evento A é independente do evento B, com $P(B) > 0$, quando a ocorrência de B não altera a probabilidade de A, isto é, se

$$P(A|B) = P(A).$$

Exemplo 2.5. Pensemos na rolagem de duas moedas não-viciadas. Este experimento tem um espaço amostral

$$\Omega = \{(cara, cara), (cara, coroa), (coroa, cara), (coroa, coroa)\},$$

e a hipótese de moedas não viciadas pode ser traduzida em modelo equiprovável.

Pensando apenas na realização do experimento, gostaríamos que o evento

$$A_2 = \{\text{observamos cara na segunda moeda}\}$$

fosse independente do evento

$$A_1 = \{\text{observamos cara na primeira moeda}\}.$$

Isso porque a rolagem da segunda moeda é *fisicamente* independente da primeira rolagem, e o resultado da primeira não deveria trazer informações sobre a segunda.

Para verificar isso observe que

$$A_1 = \{(cara, cara), (cara, coroa)\}, \quad A_2 = \{(cara, cara), (coroa, cara)\},$$

e assim

$$P(A_2|A_1) = \frac{\frac{1}{4}}{\frac{1}{2}} = \frac{1}{2} = P(A_2),$$

mostrando que A_2 é independente de A_1.

Ainda no exemplo acima vemos também que

$$P(A_1|A_2) = \frac{\frac{1}{4}}{\frac{1}{2}} = \frac{1}{2} = P(A_1),$$

mostrando que A_1 é também independente de A_2.

Surge então a seguinte questão: vale em geral que se A independe de B, então B independe de A? O que observamos no exemplo anterior é sempre válido, ou apenas consequência das fortes simetrias dos eventos escolhidos?

Percebam que o questionamento é válido, visto que saber que um pai é financeiramente independente de um filho não nos diz nada sobre a independência financeira do filho em relação ao pai.

Pois bem, se A independe de B, temos que $P(A|B) = P(A)$. Supondo que $P(A) > 0$, pelas igualdades em (2.3) e (2.4), segue que

$$\begin{aligned} P(A|B) = P(A) &\Leftrightarrow \frac{P(A \cap B)}{P(B)} = P(A) \\ &\Leftrightarrow P(A \cap B) = P(A) \cdot P(B) \\ &\Leftrightarrow \frac{P(A \cap B)}{P(A)} = P(B) \\ &\Leftrightarrow P(B|A) = P(B). \end{aligned} \tag{2.5}$$

Logo, B independe de A.

Portanto, se A independe B, então B independe de A, e podemos simplesmente dizer que os eventos A e B são independentes.

Mas isso faz a nossa definição original um tanto inadequada. Ela também possui a necessidade desagradável de que $P(B) > 0$.

Com o intuito de abranger essa simetria presente na independência entre eventos, definimos eventos independentes da seguinte forma.

Definição 2.5. Dizemos que dois eventos $A, B \subset \Omega$ são independentes se

$$P(A \cap B) = P(A) \cdot P(B).$$

Notemos que a expressão acima é a equação central no argumento detalhado em (2.5).

É interessante notar que se A, B são eventos independentes então

$$P(A^c \cap B) = P(B) - P(A \cap B) = P(B) - P(A)P(B) = (1 - P(A))P(B) = P(A^c)P(B),$$

de modo que A^c, B são também independentes.

Repetindo o mesmo argumento acima para A, B^c e A^c, B^c chegamos no seguinte resultado.

Proposição 2.6. *Se A, B são eventos independentes, então também são independentes os eventos A^c, B, os eventos A, B^c e os eventos A^c, B^c.*

Dando continuidade, observe que a definição acima parece nos sugerir uma definição mais geral de independência de eventos. De fato, é realmente tentador pensar que $A_1, A_2, \ldots, A_n$ são independentes se

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1) \cdot P(A_2) \cdots P(A_n). \tag{2.6}$$

Mas, infelizmente, tal definição não sobrevive a uma análise mais cuidadosa.

Considere, por exemplo, o experimento onde rolamos um dado de 12 lados, e considere os eventos

$$A = \{1, 2, 3, 4, 5, 6\}$$
$$B = \{3, 4, 5, 6\}$$
$$C = \{6, 7, 8, 9, 10, 11\}.$$

Alguns cálculos simples nos revelam rapidamente que não é razoável pensarmos em tais eventos como independentes. Podemos ver, por exemplo, que

$$P(C|A \cap B) = \frac{1}{4} \neq \frac{1}{2} = P(C)$$

ou ainda que

$$P(A|C) = \frac{1}{6} \neq \frac{1}{2} = P(A).$$

Note que, no entanto, $A \cap B \cap C = \{6\}$ e portanto

$$P(A \cap B \cap C) = \frac{1}{12} = \frac{1}{2} \cdot \frac{1}{3} \cdot \frac{1}{2} = P(A)P(B)P(C).$$

Isso deixa claro que a equação (2.6) não nos fornece uma maneira adequada de definir independência para mais de 2 eventos.

Precisamos então de uma nova definição. Uma tentativa válida seria definir a independência do grupo pela independência dos pares. Ou seja, $A_1, \ldots, A_n$ seriam independentes se A_i, A_j são independentes para todo par $i, j = 1, \ldots, n$, tais que $i \neq j$.

Para ver que tal definição também não funciona consideraremos mais um exemplo. O experimento agora consistirá em lançar um dado 2 vezes. Nesta situação, defina os eventos

$$\begin{aligned} A &= \{\text{observamos 2 no primeiro lançamento}\} \\ B &= \{\text{observamos 5 no segundo lançamento}\} \\ C &= \{\text{a soma dos valores é 7}\}. \end{aligned}$$

Uma conta simples (que ficará de exercício ao leitor) nos mostra que A, B são independentes, A, C são independentes e B, C são independentes. No entanto

$$P(A|B \cap C) = 1 \neq P(A),$$

o que nos mostra que apesar de todo par ser independente, a família A, B, C não pode ser.

A solução para tal dilema é, de algum modo, juntar as duas definições, culminando na seguinte definição.

Definição 2.7. Diremos que os eventos $A_1, A_2, \ldots$ são ditos independentes se para quaisquer índices $i_1, i_2, \ldots, i_k$, distintos, $k \geq 2$

$$P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) = P(A_{i_1}) \cdot P(A_{i_2}) \cdots P(A_{i_k}).$$

2.1.4 *Lei da probabilidade total*

Comecemos a partir da questão: numa urna existem 9 bolas, sendo 2 vermelhas, 3 azuis e 4 amarelas. Ao sortearmos duas bolas sem reposição, qual a probabilidade da segunda bola ser azul?

Estamos interessados em determinar a probabilidade do evento $B = \{2^{\text{a}} \text{ bola é azul}\}$. Antes, consideremos os eventos $A_1 = \{1^{\text{a}} \text{ bola é vermelha}\}$, $A_2 = \{1^{\text{a}} \text{ bola é azul}\}$ e $A_3 = \{1^{\text{a}} \text{ bola é amarela}\}$. Os eventos A_1, A_2 e A_3 formam o que chamamos uma partição do espaço amostral.

Definição 2.8. Dizemos $A_1, A_2, \ldots, A_n$ formam uma partição do espaço amostral Ω quando:

i) $A_i \cap A_j = \emptyset \quad \text{para} \quad i \neq j$

ii) $A_1 \cup A_2 \cup \cdots \cup A_n = \Omega$

De fato, qualquer que seja o resultado do sorteio, teremos a primeira bola vermelha, azul ou amarela, logo $A_1 \cup A_2 \cup A_3 = \Omega$. Além disso, não é possível que a primeira bola sorteada tenha duas cores, o que significa que A_1, A_2 e A_3 são dois a dois disjuntos.

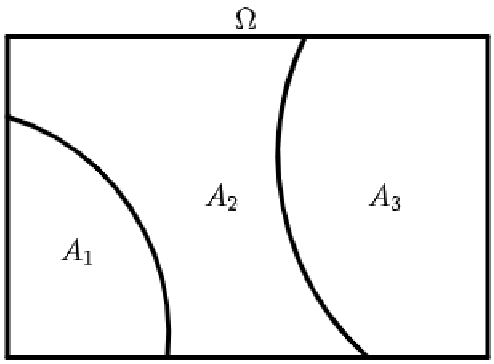


Figura 2: A_1, A_2 e A_3 formam uma partição de Ω.

Agora observemos que $B = (A_1 \cap B) \cup (A_2 \cap B) \cup (A_3 \cap B)$. Visualmente, essa igualdade é representada pela figura 3.

Consequentemente, $P(B) = P((A_1 \cap B) \cup (A_2 \cap B) \cup (A_3 \cap B))$.

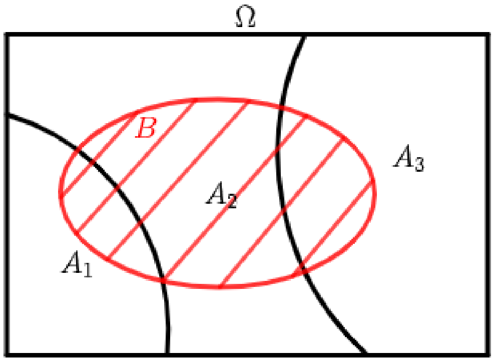


Figura 3: $B = (A_1 \cap B) \cup (A_2 \cap B) \cup (A_3 \cap B)$

Percebamos que, como A_1, A_2 e A_3 são disjuntos dois a dois, então $(A_1 \cap B), (A_2 \cap B)$ e $(A_3 \cap B)$ também são. Sendo assim,

$$P(B) = P(A_1 \cap B) + P(A_2 \cap B) + P(A_3 \cap B).$$

Pela Regra do Produto, segue que se $P(A_i) > 0$ para $i = 1, 2, 3$, então

$$P(B) = P(B|A_1) \cdot P(A_1) + P(B|A_2) \cdot P(A_2) + P(B|A_3) \cdot P(A_3) = \frac{3}{8} \cdot \frac{2}{9} + \frac{2}{8} \cdot \frac{3}{9} + \frac{3}{8} \cdot \frac{4}{9} = \frac{1}{3}.$$

Generalizando a situação descrita acima, seja B um evento qualquer de Ω, logo B pode ser representado pela união das partes que correspondem às intersecções entre B e cada evento da partição $A_1, A_2, \dots, A_n$ de Ω, ou seja,

$$B = (B \cap A_1) \cup (B \cap A_2) \cup \cdots \cup (B \cap A_n)$$

Consequentemente, a probabilidade de B pode ser obtida pela expressão

$$P(B) = P((B \cap A_1) \cup (B \cap A_2) \cup \cdots \cup (B \cap A_n))$$

Como já discutido, teremos$(B \cap A_1), (B \cap A_2), \dots, (B \cap A_n)$ disjuntos dois a dois e, portanto, chegamos ao resultado conhecido como *Lei da probabilidade total*:

$$P(B) = P(B \cap A_1) + P(B \cap A_2) + \cdots + P(B \cap A_n),$$

ou, usando a Regra do Produto e considerando $P(A_i) > 0$ para $i = 1, \dots, n$,

$$P(B) = P(B|A_1) \cdot P(A_1) + P(B|A_2) \cdot P(A_2) + \cdots + P(B|A_n) \cdot P(A_n),$$

onde $A_1, A_2, \dots, A_n$ formam uma partição do espaço amostral Ω.

Esta última igualdade para a Lei da probabilidade estará muito presente em demonstrações nos capítulos posteriores.

2.2 VARIÁVEIS ALEATÓRIAS

Voltemos a 2020, quando a COVID mudou radicalmente a vida de todos ao redor do mundo. Pense em um cientista tentando determinar a relação entre obesidade, idade e possibilidade de infecção por COVID. Para tanto, tal cientista precisaria selecionar ao acaso uma série de pacientes em uma população, e observar para cada paciente sorteado uma série de informações, como idade, índice de massa corpórea, peso, etc.

Observe que o ato de escolher um paciente na população é, por si só, um experimento aleatório. O espaço amostral Ω de tal experimento é *toda a população* da cidade ou região onde foi feita a escolha. Quando recolhe alguma informação de um paciente $\omega \in \Omega$, digamos o peso, o que o pesquisador está fazendo é associar a cada resultado do experimento (paciente) um valor real $M(\omega)$. O mesmo acontece quando ele pergunta a idade $I(\omega)$, e assim por diante. Tais informações numéricas, associadas ao resultados possíveis do experimento, são conhecidas como *variáveis aleatórias*. Neste exemplo podemos ter, por exemplo, as variáveis M =peso, I =idade, X =índice de massa corpórea, ou até $T \in \{0, 1\}$ indicando se o teste de COVID foi positivo ou negativo.

Colocando em um contexto mais claro e didático, consideremos o experimento aleatório no qual dois dados são lançados. Seja X o número correspondente ao resultado da soma das faces obtidas nos dados. Logo, X pode assumir os seguintes valores: 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, conforme mostra a figura 4.

		DADO 1					
	+	**1**	**2**	**3**	**4**	**5**	**6**
DADO 2	**1**	2	3	4	5	6	7
	2	3	4	5	6	7	8
	3	4	5	6	7	8	9
	4	5	6	7	8	9	10
	5	6	7	8	9	10	11
	6	7	8	9	10	11	12

Figura 4: soma das faces obtidas no lançamento de dois dados

Assim, dizemos que X é uma *variável aleatória* com valores no conjunto $\{2,3,4,5,6,7,8,9,10,11,12\}$.

Mais formalmente podemos definir variáveis aleatórias da seguinte forma.

Definição 2.9. Uma variável aleatória, abreviadamente v.a, é uma função X definida no espaço amostral Ω, que associa cada elemento de Ω a um valor real. Isto é, uma variável aleatória X é uma função

$$X : \Omega \longrightarrow \mathbb{R}.$$

Deste ponto em diante denotaremos as variáveis aleatórias por letras maiúsculas e seus valores (imagens possíveis), por letras minúsculas.

Em geral, se uma v.a pode assumir uma quantidade finita ou enumerável de valores possíveis, ela é chamada *variável aleatória discreta*. Como neste trabalho consideramos apenas espaços amostrais finitos ou enumeráveis, segue trivialmente que todas as variáveis aleatórias das quais trataremos são discretas.

Observação 2.2.1. *De fato, a definição apresentada de variáveis aleatórias só está completa no caso de espaços finitos ou enumeráveis. Nos demais casos é necessária uma condição adicional de* mensurabilidade *que foge ao escopo deste trabalho.*

O objetivo do nosso trabalho é apresentar e discutir alguns processos aleatórios que assumem apenas valores inteiros. Sendo assim, os conceitos abordados deste ponto em diante serão feitos a partir de variáveis aleatórias inteiras. Ou seja,

$$X : \Omega \longrightarrow \mathbb{Z}.$$

Antes de seguir, vamos fixar algumas notações. Dada uma variável aleatória X assumindo valores inteiros e $k \in \mathbb{Z}$, denotaremos por $\{X = k\}$ o evento

$$\{X = k\} := \{\omega \in \Omega : X(\omega) = k\}.$$

Da mesma forma

$$\{X \leq k\} := \{\omega \in \Omega : X(\omega) \leq k\},$$

ou ainda

$$\{X \in A\} := \{\omega \in \Omega : X(\omega) \in A\},$$

para $A \subset \mathbb{Z}$.

Com isso, $P(X = k)$ indica a probabilidade da v.a X assumir o valor k em uma dada realização do experimento. Vale algo análogo para $P(X \leq k)$ ou $P(X \in A)$ para $A \subset \mathbb{Z}$.

A *distribuição* de uma variável aleatória discreta X é a sequência de probabilidades $P(X = k)$, $k \in \mathbb{Z}$.

Sendo assim, devemos ter

$$0 \leq P(X = k) \leq 1 \quad \text{e} \quad \sum_{k \in \mathbb{Z}} P(X = k) = 1.$$

É claro que $P(X = k) = 0$ sempre que k não pertence à imagem de X. Com isso, para determinar a distribuição de X basta conhecer $P(X = k)$ para $k \in \mathcal{I}m(X)$.

Exemplo 2.6. Voltemos ao experimento aleatório descrito no início dessa seção: observar o resultado da soma das faces obtidas no lançamento de dois dados. Lembrando que X corresponde ao número obtido nessa soma, e analisando todos os resultados possíveis desse experimento através da figura 4, concluímos que a distribuição da variável aleatória X é dada pela sequência

$$P(X = 2) = \frac{1}{36}; P(X = 3) = \frac{2}{36}; P(X = 4) = \frac{3}{36}; P(X = 5) = \frac{4}{36}; P(X = 6) = \frac{5}{36}; P(X = 7) = \frac{6}{36};$$
$$P(X = 8) = \frac{5}{36}; P(X = 9) = \frac{4}{36}; P(X = 10) = \frac{3}{36}; P(= 11) = \frac{2}{36}; P(X = 12) = \frac{1}{36}.$$

Notemos que

$$\sum_{k=2}^{12} P(X = k) = \frac{1}{36} + \frac{2}{36} + \frac{3}{36} + \frac{4}{36} + \frac{5}{36} + \frac{6}{36} + \frac{5}{36} + \frac{4}{36} + \frac{3}{36} + \frac{2}{36} + \frac{1}{36} = \frac{36}{36} = 1$$

Algumas distribuições aparecem frequentemente em aplicações, e por essa razão recebem nomes especiais.

Para exemplificar três destas distribuições considere a seguinte situação. Vamos realizar um dado experimento aleatório repetidas vezes, de modo independente. Ao longo destas repetições estaremos interessados na observação de um certo evento $A \subset \Omega$. Suponha também que $P(A) = p \in (0, 1)$.

Considere então as seguintes variáveis aleatórias:

- X assume 1 se o evento A foi observado em uma dada repetição, e 0 caso contrário;
- N conta quantas vezes observamos A em n repetições.
- M conta o total de vezes que observamos A^c até observarmos A pela primeira vez.

Encontrar as distribuições de cada uma destas variáveis é parte integrante de qualquer curso básico em probabilidade, e não apresentaremos os cálculos aqui para não nos prolongarmos demais. De todo modo, listaremos as distribuições abaixo, e o leitor interessado pode buscar tais argumentos em [18].

- $P(X = 1) = p$ e $P(X = 0) = 1 - p$;
- $N \in \{0, 1, 2 \cdots, n\}$ com
$$P(N = k) = \binom{n}{k} p^k (1 - p)^{n-k};$$
- $M \geq 0$ com
$$P(M = k) = (1 - p)^k p.$$

A variável X tem distribuição de *Bernoulli* de parâmetro p, N tem distribuição *Binomial* de parâmetros n e p, enquanto M tem distribuição *Geométrica* de parâmetro p.

2.2.1 *Distribuições Conjuntas*

Comentários sobre Processos Estocásticos

Com alguma frequência, como acontece neste trabalho, precisamos estudar não apenas uma, mas uma sequência de variáveis aleatórias. Tais variáveis podem representar, por exemplo, uma característica observada ao longo de várias repetições de um mesmo experimento aleatório, ou ainda a evolução no tempo de algum fenômeno observável.

Tais sequências são conhecidas como *Processos Estocásticos*, e são em geral representados por alguma sequência (X_n) indexadas por algum parâmetro n. Neste trabalho vamos assumir $n \in \mathbb{N} \cup \{0\}$ e $X_n \in \mathbb{Z}$.

Para tornar mais claro, imaginemos que o experimento aleatório "observar o resultado da soma das faces obtidas no lançamento de dois dados" será repetido infinitas vezes. Seja X_n o resultado obtido nessa soma na n-ésima repetição. Agora notemos que esse experimento realizado nas condições descritas gera um processo estocástico a tempo discreto representado pela sequência $(X_1, X_2, X_3, \ldots)$, onde o índice "temporal" n indica simplesmente qual a repetição registrada.

No estudo de tais processos estaremos especialmente interessados no que chamamos de **distribuição conjunta** das variáveis do processo. Ou seja, estaremos interessados em estudar os valores

$$P(X_0 = i_0, X_1 = i_1, \cdots, X_n = i_n),$$

para quaisquer $i_0, \cdots, i_n \in \mathbb{Z}$.

Observação 2.2.2. *Com o intuito de simplificar a notação, em vez da notação $P(\{X = k\} \cap \{Y = i\})$, que corresponde a probabilidade da v.a X assumir o valor k e v.a Y, o valor i, adotamos na definição acima a notação $P(X = k, Y = i)$ e a usaremos deste ponto em diante.*

Exemplo 2.7. Considere um processo simples gerado pela rolagem de um dado $N \geq 2$ vezes. Defina X_n a soma dos $n \leq N$ primeiros resultados.

Neste caso a distribuição conjunta de X_1 e X_2 é dada pela tabela abaixo

$X_1 \backslash X_2$	2	3	4	5	6	7	8	9	10	11	12
1	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	0	0	0	0	0
2	0	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	0	0	0	0
3	0	0	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	0	0	0
4	0	0	0	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	0	0
5	0	0	0	0	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	0
6	0	0	0	0	0	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$

Não nos aprofundaremos demais em distribuições conjuntas. Mais uma vez, aqueles interessados em se aprofundar um pouco mais neste assunto podem verificar [18]. Terminaremos esse breve parêntese comentando apenas uma propriedade simples das distribuições conjuntas.

Note que se X_k é variável inteira então os eventos $\{\{X_k = m\}, m \in \mathbb{Z}\}$ formam uma partição de Ω. Com isso vale que

$$\sum_{m \in \mathbb{Z}} P(X_1 = i_1, \ldots, X_k = m, \ldots, X_n = i_n) = P(X_1 = i_1, \ldots, X_{k-1} = i_{m-1}, X_{k+1} = i_{m+1}, \ldots, X_n = i_n)$$

Neste trabalho nos concentraremos em uma classe de processos conhecida *Cadeias de Markov*. Informalmente falando, em uma cadeia de Markov o valor do processo em um instante $n \geq 1$ pode ser determinado sabendo apenas o valor no instante anterior, podendo ignorar toda a tragetória que levou o processo até ali.

Definição 2.10. Chamamos de *Cadeia de Markov* a tempo discreto um processo estocástico a tempo discreto $(X_n)_{n\geq 0}$ tal que

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

para todo $n \geq 1$ e para todo subconjunto $\{i_0, i_1, \ldots, i_{n-1}, i, j\}$ de valores, também denominados *estados*, das v.a's X_n.

Em outras palavras, nos processos estocásticos Markovianos, os estados anteriores são irrelevantes para a predição do estado seguinte, o qual depende unicamente do estado atual do processo.

Denote agora por $p(i, j)$ como a probabilidade de transição do estado i para estado j em um passo do processo. Suporemos também que tais probabilidades não dependem do instante n observado. Ou seja, as probabilidades de transição de uma cadeia de Markov são dadas por

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

para todo estado $i, j \in \mathbb{Z}$ e para todo $n \geq 0$.

Vale assim que que

$$p(i, j) \geq 0 \quad \text{e} \quad \sum_k p(i, j) = 1.$$

Se a probabilidade $p(i, i) = 1$, dizemos que i é um *estado absorvente*.

Para não nos prolongarmos demais, não daremos exemplos aqui. O resto do trabalho se concentrará em três exemplos importantes de Cadeiras de Markov.

É interessante ressaltar que, para que a cadeia de Markov esteja completamente determinada, é preciso que as probabilidades de transição $p(i, j)$ estejam determinadas, assim como a distribuição do estado inicial X_0. Isto é, dados estados $i_0, i_1, \ldots, i_n$ do processo, para determinar $P(X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n)$ precisamos conhecer $P(X_0 = i_0)$ e $p(i_{k-1}, i_k)$, $1 \leq k \leq n$.

De fato, segue da definição de probabilidade condicional que

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

Logo, a probabilidade $P(X_0 = i_0, X_1 = i_1, \dots, X_{n-1} = i_{n-1}, X_n = i_n)$ é igual a

$$P(X_n = i_n | X_0 = i_0, X_1 = i_1, \dots, X_{n-1} = i_{n-1}) \cdot P(X_0 = i_0, X_1 = i_1, \dots, X_{n-1} = i_{n-1}) \qquad (2.7)$$

Agora, da propriedade da cadeia de Markov, temos que

$$P(X_n = i_n | X_0 = i_0, X_1 = i_1, \dots, X_{n-1} = i_{n-1}) = P(X_n = i_n | X_{n-1} = i_{n-1}) = p(i_{n-1}, i_n) \qquad (2.8)$$

e substituindo (2.8) em (2.7), obtemos

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

Um argumento indutivo nos garante então que

$$P(X_0 = i_0, X_1 = i_1, \dots, X_n = i_n) = p(i_{n-1}, i_n) \cdot p(i_{n-2}, i_{n-1}) \cdots p(i_0, i_1) \cdot P(X_0 = i_0).$$

Assim, se $P(X_0 = i_0) = \alpha_0$, então,

$$P(X_0 = i_0, X_1 = i_1, \dots, X_n = i_n) = p(i_{n-1}, i_n) \cdot p(i_{n-2}, i_{n-1}) \cdots p(i_0, i_1) \cdot \alpha_0.$$

2.2.2 *Variáveis Aleatórias Independentes*

Introduzidas as noções de variáveis aleatórias e suas distribuições, estamos prontos para generalizar um pouco mais a noção de independência, e introduzir a noção de variáveis aleatórias independentes.

Mas o que queremos dizer com "variáveis X, Y independentes”? O que significa, por exemplo, duas rolagens de um dado são independentes?

Informalmente significa que informações sobre uma rolagem (ou uma variável) não afeta as probabilidades associadas à outra rolagem (ou variável).

Mas isso significa que não estamos comparando apenas dois eventos, mas toda uma coleção deles. Mais precisamente, precisaríamos que todo evento associado à uma variável seja independente e qualquer evento associado à outra variável.

Para deixar tais ideias mais claras, considere variáveis X, Y com a seguinte distribuição conjunta.

$X\backslash Y$	-1	0	1	
-1	$1/24$	$1/12$	$1/12$	$5/24$
0	$1/12$	$1/4$	$1/6$	$1/2$
1	$1/12$	$1/6$	$1/24$	$7/24$
	$5/24$	$1/2$	$7/24$	1

Note que

$$P(X = 0, Y = 0) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = P(X = 0)P(Y = 0),$$

e portanto os eventos $\{X = 0\}$ e $\{Y = 0\}$ são independentes.

Por outro lado

$$P(X = 1, Y = 0) = \frac{1}{6} \neq \frac{7}{24} \cdot \frac{1}{2} = P(X = 1)P(Y = 0),$$

de modo que os eventos $\{X = 1\}$ e $\{Y = 0\}$ não são independentes.

Em outras palavras, apesar do evento $\{Y = 0\}$ não nos trazer nenhuma informação sobre o evento $\{X = 0\}$, ele informa algo sobre o evento $\{X = 1\}$.

Com isso as variáveis X, Y não podem ser independentes.

Definição 2.11. Diremos que as variáveis aleatórias $X_1, X_2, \ldots, X_n \in \mathbb{Z}$ são independentes se

$$P(X_1 = i_1, \ldots, X_n = i_n) = P(X_1 = i_1) \cdots P(X_n = i_n),$$

para quaisquer $i_1, \ldots, i_n \in \mathbb{Z}$.

Se além disso temos que

$$P(X_k = i) = P(X_1 = i)$$

para todo $1 \leq k \leq n$ e $i \in \mathbb{Z}$, diremos que $X_1, X_2, \ldots, X_n$ são *independentes e identicamente distribuídas* (iid).

2.3 ESPERANÇA

Suponhamos que uma loja deseja lançar a seguinte promoção: nas compras acima de $R\$100,00$, o cliente seleciona duas bolas ao acaso, sem reposição, de uma urna contendo duas bolas douradas e três vermelhas. O cliente ganha 10 reais por cada bola dourada obtida. Qual o valor médio de desconto por cliente que a loja deve dar participantes no decorrer da promoção?

Pois bem, se considerarmos o desconto recebido por um cliente obtemos os seguintes valores:

- 0 reais de desconto com probabilidade $\frac{3}{5} \cdot \frac{2}{4} = \frac{3}{10}$,
- 10 reais de desconto com probabilidade $\frac{2}{5} \cdot \frac{3}{4} + \frac{3}{5} \cdot \frac{2}{4} = \frac{6}{10}$,
- 20 reais de desconto com probabilidade $\frac{2}{5} \cdot \frac{1}{4} = \frac{1}{10}$.

Heuristicamente, podemos pensar que ao longo da promoção

- $\frac{3}{10}$ dos clientes recebeu 0 reais de desconto;
- $\frac{6}{10}$ dos clientes recebeu 10 reais de desconto;
- $\frac{1}{10}$ dos clientes recebeu 20 reais de desconto;

Assim o desconto médio por cliente deve ser

$$0 \cdot \frac{3}{10} + 10 \cdot \frac{6}{10} + 20 \cdot \frac{1}{10} = 6 + 2 = 8 \text{ reais.}$$

Se X é a variável aleatória que corresponde ao valor do desconto dado a um cliente, diremos neste caso que a *esperança* de X é

$$E(X) = 0 \cdot P(X = 0) + 10 \cdot P(X = 10) + 20 \cdot P(X = 20) = 8.$$

Em geral vale a seguinte definição.

Definição 2.12. Dada uma variável aleatória $X \in \mathbb{Z}$ e $f : \mathbb{Z} \to \mathbb{R}$, definimos a **esperança**, ou o **valor esperado**, ou a **média** da variável $f(X)$ por

$$E(f(X)) = \sum_{k \in \mathbb{Z}} f(k) P(X = k),$$

se a soma acima existir. Em particular

$$E(X) = \sum_{k \in \mathbb{Z}} k P(X = k).$$

Observação 2.3.1. *Ao longo deste trabalho trabalharemos com variáveis aleatórias finitas ou positivas e com f também positiva. No primeiro caso, as somas acima são finitas e estão bem definidas. No segundo caso as somas acima devem ser pensadas como*

$$\sum_{k\in\mathbb{Z}} f(k)P(X=k) = \sum_{k=0}^{\infty} f(k)P(X=k).$$

Neste caso, se a série divergir diremos que $E(f(X)) = \infty$.

Os demais casos podem ser um pouco mais complicados de se trabalhar, e não entraremos em maiores detalhes neste trabalho.

A esperança de X é uma espécie de *medida de posição* da distribuição de X, que corresponde à média ponderada dos possíveis valores de X, ponderados pela probabilidade de tal valor ser observado.

2.3.1 *Propriedades da Esperança*

Seguimos enumerando algumas propriedades da Esperança matemática fundamentais para o nosso estudo.

Proposição 2.13. *Se X e Y são variáveis aleatórias inteiras com esperança finita, então vale que*

1. *$E(c) = c$ e $E(cX) = cE(X)$, para qualquer $c \in \mathbb{R}$.*
2. *$E(X+Y) = E(X) + E(Y)$.*
3. *se X e Y são independentes, então $E(XY) = E(X) \cdot E(Y)$.*

Demonstração. De fato, segue da definição de Esperança que,

1. Considere as funções $f(x) = c$ e $g(x) = cX$, e note que

$$E(c) = E(f(X)) = \sum_k f(k)P(X=k) = \sum_k cP(X=k) = c\sum_k P(X=k) = c\cdot 1 = c,$$

e

$$E(cX) = E(g(X)) = \sum_k g(k)P(X=k) = \sum_k ckP(X=k) = c\sum_k kP(X=k) = cE(x),$$

concluindo a primeira propriedade.

2. Como X e Y são v.a inteira, $X+Y$ também é uma v.a inteira, e além disso

$$\begin{aligned} P(X+Y=k) &= \sum_i P(X=i, X+Y=k) \\ &\sum_i P(X=i, Y=k-i), \end{aligned}$$

e portanto

$$\begin{aligned} E(X+Y) &= \sum_k \sum_k kP(X+Y=k) \\ &= \sum_k \sum_i kP(X=k, Y=k-i), \end{aligned}$$

e fazendo $j=k-i$ temos

$$\begin{aligned} E(X+Y) &= \sum_j \sum_i (i+j)P(X=k, Y=j) \\ &= \sum_i \sum_j iP(X=i, Y=j) + \sum_j \sum_i jP(X=i, Y=j) \\ &= \sum_i i \left(\sum_j P(X=i, Y=j) \right) + \sum_j j \left(\sum_i P(X=i, Y=j) \right) \\ &= \sum_i iP(X=i) + \sum_j jP(Y=j) \\ &= E(X) + E(Y), \end{aligned}$$

e o resultado segue.

3. Assim como no item anterior, como X e Y são v.a inteira, $X \cdot Y$ também é uma v.a inteira, e portanto para $k \neq 0$ vale que

$$\begin{aligned} P(XY=k) &= \sum_i P(X=i, XY=k) \\ &= \sum_i P(X=i, Y=k/i), \end{aligned}$$

Seguindo um argumento análogo ao do item anterior encontramos que

$$E(X \cdot Y) = \sum_i \sum_j i \cdot jP(X=i, Y=j).$$

Como X e Y são v.a independentes, então

$$P(X=k, Y=i) = P(X=k) \cdot P(Y=i),$$

e portanto

$$\begin{aligned}
E(XY) &= \sum_i \sum_j (i \cdot j) P(X = i, Y = j) \\
&= \sum_i \sum_j i \cdot j \cdot P(X = i) \cdot P(Y = j) \\
&= \left(\sum_i i \cdot P(X = i) \right) \left(\sum_i j \cdot P(Y = j) \right) \\
&= E(X)E(Y),
\end{aligned}$$

concluindo a propriedade e a proposição.

□

Na proposição acima foi necessário calcular a esperança de variáveis definidas a partir de outras duas variáveis distintas. Mais especificamente, $X + Y$ e XY. Ao longo do trabalho precisaremos fazer isso com funções um pouco mais complexas, e para isso precisamos deixar claro como pretendemos realizar isso com algumas expressões. Assim, inspirados nos resultados acima temos a seguinte definição.

Definição 2.14. Dadas variáveis aleatórias $X_1, \ldots X_n$ inteiras e $f : \mathbb{Z}^n \to \mathbb{R}$ uma função real, defina

$$E(f(X_1, \ldots, X_n)) = \sum_{k_1} \cdots \sum_{k_n} f(k_1, \ldots, k_n) P(X_1 = k_1, \ldots, X_n = k_n).$$

2.3.2 *Esperança Condicional*

Comecemos novamente usando uma situação como exemplo: duas bolas são selecionadas ao acaso, com reposição, de uma urna contendo 5 bolas: 2 pretas e três brancas.

Se denotarmos por X o total de bolas pretas sorteadas, temos que $X \in \{0, 1, 2\}$ com

$$\begin{aligned}
P(X = 0) &= \frac{3}{5} \cdot \frac{3}{5} = \frac{9}{25} \\
P(X = 1) &= \frac{3}{5} \cdot \frac{2}{5} + \frac{2}{5} \cdot \frac{3}{5} = \frac{12}{25} \\
P(X = 2) &= \frac{2}{5} \cdot \frac{2}{5} = \frac{4}{25},
\end{aligned}$$

e portanto

$$E(X) = 0 \cdot \frac{9}{25} + 1 \cdot \frac{12}{25} + 2 \cdot \frac{4}{25} = \frac{4}{5}.$$

Suponha que recebamos a informação de que a primeira bola sorteada foi preta. Como podemos determinar a nova quantidade esperada de bolas pretas selecionadas nesse experimento, dado agora que a 1ª bola retirada é preta?

Pois bem, se definirmos

$$Y_1 = \begin{cases} 0, & \text{se primeira bola retirada é branca} \\ 1, & \text{se primeira bola retirada é preta,} \end{cases}$$

encontramos que

- $P(X = 0|Y_1 = 1) = 0$, pois, nessa condição, não podemos ter 0 bolas pretas selecionadas.
- $P(X = 1|Y_1 = 1) = \frac{3}{5} = P(\{2^a$ bola retirada é branca$)\}$, pois a 2ª bola retirada deve ser branca para obtermos apenas 1 bola preta selecionada.
- $P(X = 2|Y_1 = 1) = \frac{2}{5} = P(\{2^a$ bola retirada é preta$\}$, já que a 2ª bola retirada deve ser preta para obtermos 2 bolas pretas selecionadas.

Dessa forma, a quantidade esperada de bolas pretas selecionadas dado que a 1ª bola retirada é preta, que denotaremos por $E(X|Y_1 = 1)$ deve ser

$$\begin{aligned} E(X|Y_1 = 1) &= 0 \cdot P(X = 0|Y_1 = 1) + 1 \cdot P(X = 1|Y_1 = 1) + 2 \cdot P(X = 2|Y_1 = 1) \\ &= 0 \cdot 0 + 1 \cdot \frac{3}{5} + 2 \cdot \frac{2}{5} \\ &= \frac{7}{5} \end{aligned} \tag{2.9}$$

e, portanto, a quantidade esperada de bolas pretas selecionadas, dado que a 1ª bola retirada é preta é $\frac{7}{5} = 1,4$.

Diante dessa exemplificação, apresentamos a seguir a definição de *esperança condicional*.

Definição 2.15. Dadas variáveis aleatórias $X, Y \in \mathbb{Z}$ e uma função $f : \mathbb{Z} \to \mathbb{R}$, a esperança condicional de $f(X)$ dado que $Y = y$ é dada por

$$E(f(X)|Y = y) = \sum_k f(k)P(X = k|Y = y)$$

onde a soma é feita sobre todos os valores possíveis de X.

Para melhor entender a definição acima primeiro notemos que $P(\cdot|Y=y)$ é, por si só, uma probabilidade no espaço amostral Ω. De fato é fácil ver que

1. $0 \leq P(A|Y=y) = P(A, Y=y)/P(Y=y) \leq 1$,
2. $P(\Omega|Y=y) = 1$,
3. Se $A_1, A_2, A_3, \ldots$ são 2 a 2 disjuntos, então $A_1 \cap \{Y=y\}, A_2 \cap \{Y=y\}, \ldots$ são 2 a 2 disjuntos e

$$\begin{aligned} P\left(\bigcup_n A_n \middle| Y=y\right) &= \frac{P\left(\bigcup_n A_n \cap \{Y=y\}\right)}{P(Y=y)} \\ &= \sum_n \frac{P(A_n \cap \{Y=y\})}{P(Y=y)} \\ &= \sum_n P(A_n|Y=y). \end{aligned}$$

Essa mudança de probabilidade no espaço Ω trás consigo alterações nas probabilidades associadas à variável X, e com isso um novo valor esperado.

Em outras palavras, a informação sobre a ocorrência do evento $\{Y=y\}$ altera a probabilidade no espaço amostral Ω de $P(\cdot)$ para $P(\cdot|Y=y)$, e $E(X|Y=y)$ é simplesmente o valor esperado de X de acordo com a probabilidade $P(\cdot|Y=y)$.

As propriedades da esperança condicional e suas demonstrações seguem, em grande parte, de forma análoga aos resultados apresentados na subseção 2.13. Sendo assim, enunciaremos algumas destas propriedades sem maiores justificativas.

Proposição 2.16. *Se X, Y e Z são três variáveis aleatórias inteiras, então,*

1. *$E(c|Y=y) = c$ e $E(cX|Y=y) = cE(X|Y=y)$, onde c é uma constante.*
2. *$E(X+Y|Z=z) = E(X|Z=z) + E(Y|Z=z)$*
3. *Se X, Y são independentes, então $E(X \cdot Y|Z=z) = E(X|Z=z) \cdot E(Y|Z=z)$*
4. *Para $f : \mathbb{Z} \to \mathbb{R}$, é verdade que*

$$E(f(X)) = \sum_{k \in \mathbb{Z}} E(f(X)|Y=k)P(Y=k),$$

5. *Se X e Y são independentes, então $E(X|Y=y) = E(X)$ e $E(Y|X=x) = E(Y)$.*

Demonstração. Como comentamos, as 3 primeiras propriedades seguem de modo análogo às suas contrapartidas não condicionais. Focamos na demonstração das duas últimas propriedades.

4. Segue da regra da probabilidade total que

$$\begin{aligned} E(X) &= \sum_{i\in\mathbb{Z}} f(i)P(X=i) \\ &= \sum_{i\in\mathbb{Z}} f(i) \sum_{k\in\mathbb{Z}} P(X=i|Y=k)P(Y=k) \\ &= \sum_{k\in\mathbb{Z}} \left[\sum_{i\in\mathbb{Z}} f(i)P(X=i|Y=k)\right] P(Y=k) \\ &= \sum_{k\in\mathbb{Z}} E(f(X)|Y=k)P(Y=k). \end{aligned}$$

5. Segue da definição de probabilidade condicional que

$$E(X|Y=y) = \sum_k kP(X=k|Y=y) = \sum_k k\frac{P(X=k, Y=y)}{P(Y=y)}$$

Se X e Y são independentes, então

$$E(X|Y=y) = \sum_k k\frac{P(X=k)\cdot P(Y=y)}{P(Y=y)} = \sum_k kP(X=k) = E(X).$$

Analogamente, obtêm-se $E(Y|X=x) = E(Y)$.

□

Seguimos agora com um conceito fundamental para as futuras discussões neste trabalho.

2.4 FUNÇÕES GERADORAS DE PROBABILIDADE

Antes de dar início a esta última parte da nossa "revisão teórica", precisamos fazer alguns comentários sobre o ferramental matemático que utilizaremos a seguir.

Um breve parêntese sobre séries de potência...

Nos resultados que seguem precisaremos usar (e abusar) do conceito de *séries de potência*. A teoria por trás de tais objetos é maior e mais complexa do que conseguiríamos cobrir nestas páginas.

Abaixo traremos a definição e uma breve lista dos resultados relativos a séries de potência que usaremos ao longo do trabalho.

Aos leitores ainda não familiarizados com o conceito, e que pretendam se aprofundar, recomendamos a leitura de [21] para uma abordagem mais operacional e de [19] para uma visão mais formal do assunto.

Rapidamente então, uma série de potências com coeficientes $(a_n)_{n\geq 0}$ em torno de $a \in \mathbb{R}$ é uma série do tipo

$$f(s) = \sum_{k=0}^{\infty} a_k (s-a)^k.$$

A função acima está bem definida sempre que a série à direita for convergente, o que acontece sempre que $|s-a| < R$, onde

$$R = \liminf_{n\to\infty} |a_n|^{-\frac{1}{n}}$$

é conhecido como **raio de convergência da série**. Neste caso temos que $\sum_k ak(s-a)^k$ diverge para $|s| > R$. O comportamento em $|s| = R$ tem que ser definido caso a caso.

Outro resultado importante é que a série

$$f(s) = \sum_{k=0}^{\infty} a_k (s-a)^k$$

possui raio de convergência $R > 0$ então $f(s)$ é diferenciável em $(a-R, a+R)$ e

$$f'(s) = \sum_{k=1}^{\infty} k a_k (s-a)^{k-1},$$

e a série acima tem o mesmo raio de convergência.

Aplicando esta ideia na série $f'(s)$, podemos derivar $f(s)$ quantas vezes quisermos, sem alterar o raio de convergência.

Ou seja, f é infinitamente diferenciável e

$$f^{(n)}(s) = \sum_{k=n}^{\infty} \frac{k!}{(k-n)!} a_k (s-a)^{k-n},$$

para $s \in (a-R, a+R)$.

Para completar, vamos precisar um último resultado. A versão que apresentamos a seguir é uma adaptação de um teorema maior, conhecido como Teorema de Abel, em homenagem ao matemático Norueguês Niels Henrik Abel.

Teorema 2.17 (Teorema de Abel - Adaptado)**.** *Se $a_n \geq 0$ para todo $n \geq 0$ e*

$$f(s) = \sum_{k=0}^{\infty} a_k s^k$$

converge para $|s| < 1$ então

$$\lim_{s \to 1^-} f(s) = \sum_{k=0}^{\infty} a_k.$$

Em particular, se $\sum\limits_k a_k$ converge, então $f(s)$ é contínua em $s = 1$.

Finalizado este breve parêntese, estamos prontos para retornar ao assunto principal desta seção.

Comecemos com uma definição.

Definição 2.18. Seja X uma variável aleatória com valores em $\mathbb{Z}_+$. Chamamos de *função geradora de probabilidade* da variável aleatória X a função definida por

$$f(s) = E(s^X) = \sum_{k=0}^{\infty} s^k P(X = k)$$

Notemos que f é uma série potências e que $|s^k P(X = k)| \leq |s|^k$. Como a série geométrica $\sum_{k \geq 0} |s|^k$ converge para $|s| < 1$, então, pelo teste de comparação [21], f também converge. E, portanto, f está bem definida em $(-1, 1)$.

Agora, notemos que

$$f(s) = E(s^X) = P(X = 0) + s^1 P(X = 1) + s^2 P(X = 2) + s^3 P(X = 3) + \cdots \tag{2.10}$$

e que

$$f(1) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) + \cdots = 1.$$

Logo, f também é definida em $s = 1$.

Já para $f(-1)$, temos que

$$f(-1) = \sum_{k=0}^{\infty} (-1)^k P(X = k).$$

Lembrando que se uma série $\sum a_n$ for absolutamente convergente, então $\sum a_n$ é convergente [21], segue que f também é definida em $s = -1$ e, portanto, o domínio $D(f)$ de uma função geradora de probabilidade f é tal que $[-1, 1] \subset D(f)$.

Exemplo 2.8 (Variáveis Geométricas)**.** Consideremos uma sequência de tentativas idênticas e independentes de um experimento. Estamos interessados na observação de um certo evento A que ocorre com probabilidade p a cada repetição do experimento. Chamaremos cada observação do evento A de sucesso, e falha cada vez que isso não ocorrer. Seja X quantidade de falhas até o primeiro sucesso.

Assim, a distribuição da variável aleatória X é

$$P(X = K) = (1 - p)^k p, \qquad k \geq 0.$$

Variáveis com a distribuição acima são chamadas de variáveis Geométricas (ou com distribuição Geométrica) de parâmento p, e sua função geradora de probabilidade é dada por

$$\begin{aligned} f(s) &= \sum_{k=0}^{\infty} s^k (1-p)^k p \\ &= p \sum_{k=0}^{\infty} [s(1-p)]^k \end{aligned}$$

Como série geométrica converge para $|x| < 1$ com

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x},$$

segue que para $|s| < 1/(1-p)$

$$f(s) = \frac{p}{1 - s(1-p)}.$$

A função f recebe o nome de geradora de probabilidade justamente por nos fornecer a distribuição de probabilidade da variável aleatória X, ou seja, podemos obter a probabilidade de $X = k$ a partir da função $f(s)$.

De fato, como $f(s)$ está definida no intervalo $(-1, 1)$, sabemos que f é infinitamente diferenciável em $(-1, 1)$. Calculando os valores de $f^{(k)}(0)$, onde $f^{(k)}$ corresponde a k-ésima derivada da função $f(s)$, recuperamos os termos independentes de cada uma das séries. Isto é,

$$f(s) = P(X = 0) + s^1 P(X = 1) + s^2 P(X = 2) + s^3 P(X = 3) + \cdots \Rightarrow f(0) = P(X = 0)$$

$$f'(s) = 1 \cdot P(X = 1) + 2s^1 P(X = 2) + 3s^2 P(X = 3) + \cdots \Rightarrow f'(0) = 1!\, P(X = 1)$$

$$f''(s) = 2\cdot 1\cdot P(X=2) + 3\cdot 2s^1 P(X=3) + 4\cdot 3s^2 P(X=4) + \cdots \Rightarrow f''(0) = 2!\,P(X=2)$$

$$f'''(s) = 3\cdot 2\cdot 1\cdot P(X=3) + 4\cdot 3\cdot 2s^1 P(X=4) + 5\cdot 4\cdot 3s^2 P(X=5) + \cdots \Rightarrow f'''(0) = 3!\,P(X=3)$$

Assim, um argumento indutivo nos mostra que

$$f^{(k)}(0) = k!\,P(X=k),$$

e portanto

$$P(X=k) = \frac{f^{(k)}(0)}{k!}.$$

Exemplo 2.9. Seja N uma variável aleatória inteira positiva com função geradora de probabilidade dada por

$$f(s) = e^{-\lambda(1-s)} = e^{-\lambda}e^{\lambda s}.$$

Nestas condições temos

$$f^{(k)}(s) = e^{-\lambda}e^{\lambda s}\lambda^k,$$

e portanto

$$P(N=k) = \frac{f^{(k)}(0)}{k!} = e^{-\lambda}\frac{\lambda^k}{k!},$$

para $k \geq 0$.

Este resultado nos permite concluir a seguinte proposição.

Proposição 2.19. *Sejam X e Y variáveis aleatórias com funções geradoras de probabilidades f_X e f_Y, respectivamente. Nestas condições, se*

$$f_X(s) = f_Y(s),$$

para $|s| < 1$, então X e Y têm a mesma distribuição. Isto é, se

$$P(X=k) = P(Y=k),$$

para todo $k \geq 0$.

Demonstração. Basta observar que para $k \geq 0$

$$P(X=k) = \frac{f_X^{(k)}(0)}{k!} = \frac{f_Y^{(k)}(0)}{k!} = P(Y=k).$$

Para maiores detalhes, veja [19], página 177, Teorema 8.5. □

Dando sequência, vamos agora analisar a função geradora em $s = 1$.

Já vimos anteriormente que $f(1) = 1$, e o teorema 2.17 nos garante que $f(s)$ é contínua em $s = 1$.

Se supormos que a série $\sum_k s^k P(X = k)$ tem raio de convergência $R > 1$, então $f(s)$ é diferenciável em $s = 1$ e

$$f'(1) = \sum_{k=0}^{\infty} kP(X = k) = E(X).$$

Mas mesmo que não consigamos garantir que $R > 1$, ainda tem algo que podemos concluir.

Primeiro lembre que $f(s)$ é infinitamente diferenciável para $|s| < 1$. Em particular sabemos que

$$f'(s) = \sum_{k=1}^{\infty} ks^{k-1} P(X = k).$$

Assim, o teorema de Abel (teorema 2.17) nos garante que

$$\lim_{s \to 1^-} f'(s) = \sum_{k=0}^{\infty} kP(X = k) = E(X).$$

Finalizaremos nossa não tão breve revisão teórica dos conceitos de probabilidade nos números inteiros com outro resultado importante a cerca das funções geradoras.

Proposição 2.20. *Se $X_1, \ldots, X_n$ são n variáveis aleatórias independentes com funções geradoras de probabilidades $f_{X_1}, \ldots, f_{X_n}$, respectivamente, então para $S_n = X_1 + \cdots + X_n$ vale que*

$$f_{S_n}(s) = f_{X_1}(s) \cdots f_{X_n}(s),$$

onde f_{S_n} é a função geradora de probabilidade de S_n.

Demonstração. Vamos fazer aqui apenas o caso $n = 2$. O caso geral segue por indução.

Para isso faça $X_1 = X$ e $X_2 = Y$ e observe que

$$f_{X+Y}(s) = \sum_{k \geq 0} s^k P(X + Y = k),$$

e como X e Y são independentes, temos que

$$P(X + Y = k) = \sum_{n=0}^{k} P(X = n; Y = k - n) = \sum_{n=0}^{k} P(X = n)P(Y = k - n)$$

Logo,

$$f_{X+Y}(s) = \sum_{k \geq 0} s^k P(X + Y = k) = f_{X+Y}(s) = \sum_{k \geq 0} s^k \sum_{n=0}^{k} P(X = n)P(Y = k - n)$$

Com isso temos que

$$
\begin{aligned}
f_{X+Y}(s) = \quad & s^0[P(X=0)P(Y=0)] \\
& + s^1[P(X=0)P(Y=1) + P(X=1)P(Y=0)] \\
& + s^2[P(X=0)P(Y=2) + P(X=1)P(Y=1) + P(X=2)P(Y=0)] \\
& + s^3[P(X=0)P(Y=3) + P(X=1)P(Y=2) + P(X=2)P(Y=1) + P(X=3)P(Y=0)] \\
& + \cdots
\end{aligned}
$$

Reagrupando os termos dessa soma, podemos somar as "colunas” acima primeiro para obtermos que

$$
\begin{aligned}
f_{X+Y}(s) &= P(X=0)\sum_{k\geq 0} s^k P(Y=k) + P(X=1)\sum_{k\geq 1} s^k P(Y=k-1) + \cdots \\
&= \sum_{n=0}^{\infty} P(X=n)\sum_{k\geq n} s^k P(Y=k-n) \\
&= \sum_{n=0}^{\infty} s^n P(X=n)\sum_{k\geq n} s^{k-n} P(Y=k-n) \\
&= \sum_{n=0}^{\infty} s^n P(X=n)\sum_{j\geq 0} s^j P(Y=j) \\
&= \sum_{n=0}^{\infty} s^n P(X=n) f_y(s) \\
&= f_X(s) f_Y(s),
\end{aligned}
$$

como queríamos demonstrar. □

E assim esperamos ter construído um bom alicerce para os principais resultados que constituirão o corpo desse trabalho. E que comece a diversão!

3

PASSEIOS ALEATÓRIOS SIMPLES

Com o intuito de fomentar nosso estudo sobre os Passeios aleatórios simples, iniciaremos esse capítulo apresentando problema clássico na teoria da probabilidade: *O Problema da ruína do Jogador*. Tal problema modela de maneira simples um jogo de apostas, e nele discutiremos sobre as probabilidades do jogador perder todo o seu dinheiro e a duração esperada em determinado jogo de apostas. A partir dessa discussão, o nosso estudo abordará a definição e alguns resultados importantes de um modelo aparentemente simples, mas com grande poder na modelagem estocástica, os passeios aleatórios simples.

3.1 PROBLEMA DA RUÍNA DO JOGADOR

Consideremos um jogador que vai a um cassino para fazer uma série de apostas. Em cada uma delas, ou ele ganha 1 real com probabilidade p, ou perde 1 real com probabilidade $q = 1 - p$. Começando com uma fortuna inicial de $m \geq 0$, ele jogará até perder todo o seu dinheiro, e estar arruinado, ou até acumular um fortuna de $N \geq m$ reais.

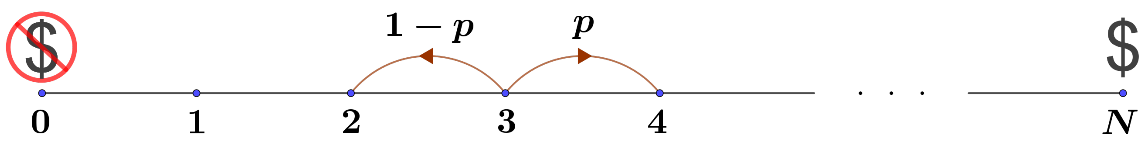


Figura 5: O processo da Ruína do Jogador

Para descrever matematicamente o problema acima, seja $(Y_n)_{n\geq 1}$ uma sequência de variáveis aleatórias independentes e identicamente distribuídas, com distribuição dada por

$$P(Y_n = 1) = p \quad \text{e} \quad P(Y_n = -1) = q = 1 - p$$

com $p \in (0, 1)$, indicando o quanto o jogador ganhou ou perdeu na n-ésima aposta.

O processo será dado então por uma sequência $(X_n)_{n\geq 0}$ de variáveis aleatórias com valores em $\mathbb{Z}$, onde X_n representa a fortuna do jogador na n-ésima aposta, tais que $X_0 = m \geq 0$ e

$$X_n = \begin{cases} 0, & \text{se } X_{n-1} = 0 \\ X_{n-1} + Y_n, & \text{se } 0 < X_{n-1} < N \\ N, & \text{se } X_{n-1} = N \end{cases}$$

para $n \geq 1$.

Assim, para exemplificar, se o jogador começa com 1 real e tem a seguinte sequência de resultados:

vitória, derrota, vitória, vitória, derrota, derrota, derrota,

teremos que

$$Y_1 = 1, Y_2 = -1, Y_3 = 1, Y_4 = 1, Y_5 = -1, Y_6 = -1, Y_7 = -1$$

e o processo seguirá a seguinte trajetória

$$X_0 = 1, X_1 = 2, X_2 = 1, X_3 = 2, X_4 = 3, X_5 = 2, X_6 = 1, X_7 = 0,$$

terminando com a ruína do jogador após 7 jogadas.

É interessante notar que se o jogo demora mais que n jogadas para acabar, então

$$X_n = X_0 + Y_1 + \cdots + Y_n = m + Y_1 + \cdots + Y_n, \tag{3.1}$$

sempre que $X_0 = m$. Esta caracterização nos será útil mais à frente.

Observe que se $X_0 = 0$ ou $X_0 = N$, então $X_n = X_0$ para todo $n \geq 1$, e o processo é trivial. Os casos interessantes são, claramente, aqueles onde $0 < m < N$, mas incluiremos os casos triviais por razões que ficarão claras mais à frente.

Outro ponto importante é que a fortuna do jogador no instante n depende somente da fortuna que ele possui no instante $n-1$. De fato, se $i_0, i_1, i_2, \ldots, i_n$ são tais que $|i_k - i_{k-1}| = 1$, $1 \geq k \geq n$, como Y_{n+1} é independente de X_n, então

$$P(X_{n+1} = j | X_k = i_k, 1 \geq k \geq n) = P(Y_n = j - i_n) = P(X_{n+1} = j | X_n = i_n).$$

Logo, o processo em questão é uma cadeia de Markov com probabilidades de transição dadas por

$$P(X_n = i+1 | X_{n-1} = i) = p \qquad \text{e} \qquad P(X_n = i-1 | X_{n-1} = i) = q = 1-p,$$

para quaisquer $n \geq 1$. Além disso o processo possui dois estados absorventes, 0 e N, como comentamos acima.

A questão agora está em determinar qual a probabilidade do jogador ruir diante das condições apresentadas, que é o que responderemos a seguir.

3.1.1 *A Probabilidade da ruína do jogador*

Nesta seção queremos estudar a ruína do jogador. Mais precisamente, queremos determinar a probabilidade do jogador perder todo seu dinheiro.

Para tal vamos, antes de mais nada tentar descrever o evento $\{$ruína$\}$ usando a descrição do processo dada acima.

Note, antes demais mais nada, que para que ocorra a ruína o jogador deve perder todo o seu dinheiro, antes de ganhar N reais. Esta ideia pode ser formalizada definindo para cada $0 \leq i \leq N$ a variável

$$T_i = \min\{n \geq 0 : X_n = i\},$$

representando o primeiro instante que o processo visita o estado i.

Com isso a ruína ocorre justamente quando $T_0 < T_N$. Ou seja,

$$P(\text{ruína} | X_0 = m) = P(T_0 < T_N | X_0 = m), \quad \text{para } 0 \leq m \leq N.$$

É interessante observar que $P(T_0 < T_N | X_0 = N) = 0$ e $P(T_0 < T_N | X_0 = 0) = 1$, justamente porque 0 e N são estados absorventes do processo.

Podemos fazer também algumas observações simples a respeito da probabiliade de ruína. A primeira é que se $X_0 = m \notin \{0, N\}$ e $Y_1 = \cdots = Y_m = -1$ então $T_0 = m$ e a ruína acontece. Ou seja

$$P(\text{ruína}|X_0 = m) \geq P(Y_1 = -1, \ldots, Y_m = -1) = q^m = (1-p)^m,$$

uma vez que $Y_1, \ldots, Y_m$ são independentes com $P(Y_m = -1) = 1 - p$.

De maneira análoga, se $X_0 = m \notin \{0, N\}$ e $Y_1 = \cdots = Y_{N-m} = 1$, então $T_N = N - m$ e a ruína é evitada. Com isso

$$P(\text{ruína}|X_0 = m) = 1 - P(\text{não ruína}|X_0 = m) \leq 1 - P(Y_1 = 1, \ldots, Y_{N-m} = 1) = 1 - p^{N-m}.$$

Em resumo

$$(1-p)^m \leq P(\text{ruína}|X_0 = m) \leq 1 - p^{N-m}.$$

Mas estas estimativas parecem pouco informativas, especialmente para m grande. De fato, se m for grande então $(1-p)^m$ está próximo a 0, o que diria apenas que a probabilidade de ruína é positiva. Do mesmo modo, se $m < N$ é grande então N é grande e o jogador não se satisfaz com pouco. Se, além disso, a diferença $N - m$ é grande, então $1 - p^{N-m}$ está perto de 1 e, não temos nenhuma informação muito relevante.

Felizmente é possível calcularmos exatamente a probabilidade de ruína, como mostramos na seguinte proposição.

Proposição 3.1. *Com base no Problema da ruína do jogador descrito acima, a probabilidade da ruína é dada por*

$$P(\text{ruína}|X_0 = m) = \begin{cases} \dfrac{\left(\frac{q}{p}\right)^m - \left(\frac{q}{p}\right)^N}{1 - \left(\frac{q}{p}\right)^N}, & \text{se } p \neq q \\ 1 - \dfrac{m}{N}, & \text{se } p = q \end{cases}$$

Demonstração. Primeiro note que para $m \notin \{0, N\}$ a probabilidade da ruína pode ser expressa como

$$\begin{aligned} P(T_0 < T_N|X_0 = m) &= P(T_0 < T_N; X_1 = m-1|X_0 = m) + P(T_0 < T_N; X_1 = m+1|X_0 = m) \\ &= P(T_0 < T_N|X_1 = m-1, X_0 = m)P(X_1 = m-1|X_0 = m) \\ &\quad + P(T_0 < T_N|X_1 = m+1, X_0 = m)P(X_1 = m-1|X_0 = m). \end{aligned}$$

Antes de seguir, fixado $i \in \{-1, 1\}$ queremos entender $P(T_0 < T_N | X_1 = m + i, X_0 = m)$, e para isso temos algumas condições à analisar.

Primeiro, se $X_1 = m + i = 0$, então $T_0 = 1$

$$P(T_0 < T_N | X_1 = m + i, X_0 = m) = 1 = P(T_0 < T_N | X_0 = m + i).$$

Do mesmo modo, de $X_1 = m + i = N$, então $T_N = 1$

$$P(T_0 < T_N | X_1 = m + i, X_0 = m) = 0 = P(T_0 < T_N | X_0 = m + i).$$

Por outro lado, se $X_1 = m + i \notin \{0, N\}$, então $Y_1 = i$ segue da equação (3.1) que para $n \leq \min\{T_0, T_N\}$

$$\begin{aligned} X_n &= m + Y_1 + Y_2 + \cdots + Y_n \\ &= m + i + Y_2 + \cdots + Y_n. \end{aligned}$$

E como $Y_2, \cdots, Y_n$ tem a mesma distribuição conjunta de $Y_1, \ldots, Y_{n-1}$, segue que condicionado à $X_1 = m + i$, o processo tem a mesma distribuição de

$$\begin{aligned} X_{n-1} &= X_0 + Y_1 + \cdots + Y_{n-1} \\ &= m + i + Y_1 + \cdots + Y_{n-1}. \end{aligned}$$

Ou seja, tem a mesma distribuição de X_{n-1} condicionado apenas a $X_0 = m + i$.

Intuitivamente falando, ao mover de m para $m + i$ o processo se renova, considerando um novo ponto de início e descontando o instante que passou. Ou seja

$$P(X_n \in A | X_1 = m + i, X_0 = m) = P(X_{n-1} \in A | X_0 = m + i).$$

Observação 3.1.1. *Este resultado, assim como o que apresentaremos abaixo, é em geral bastante intuitivo, e ao apresentar o tema para estudantes ou para o público em geral, pode ser uma ideia pular tais argumentos, e apresentar diretamente a equação (3.3), que apresentamos abaixo. De fato, a própria relação entre a ruína e o evento $\{T_0 < T_N\}$ pode ser desnecessária.*

De fato, é assim que a encontramos na maior parte da literatura, incluindo [20].

Para ver o efeito disto, note que se $X_1 = m + i = 0$, então para $j \in \{0, N\}$

$$T_j = \min\{n \geq 1 : X_n = j\} = \min\{n \geq 1 : m + i + Y_2 + \ldots + Y_n = j\},$$

que tem a mesma distribuição de

$$\min\{n \geq 1 : m + i + Y_1 + \ldots + Y_{n-1} = j\} = \min\{n \geq 1 : X_{n-1} = j\} = T_j + 1, \tag{3.2}$$

mas agora condicionado apenas à $X_0 = m + i$.

Intuitivamente, ao andarmos de m para $m + i$ o total de jogadas necessárias para terminar o jogo (com ou sem ruína) é a primeira jogada mais o total necessário se nossa fortuna inicial fosse de $m + i$.

Em outras palavras, se $X_1 = m + i \notin \{0, N\}$ então

$$P(T_0 < T_N | X_1 = m + i, X_0 = m) = P(T_0 + 1 < T_N + 1 | X_0 = m + i) = P(T_0 < T_N | X_0 = m + i).$$

E assim obtemos que

$$\begin{aligned} P(T_0 < T_N | X_0 = m) &= P(T_0 < T_N | X_1 = m - 1, X_0 = m) P(X_1 = m - 1 | X_0 = m) \\ &\quad + P(T_0 < T_N | X_1 = m + 1, X_0 = m) P(X_1 = m - 1 | X_0 = m) \\ &= P(T_0 < T_N | X_0 = m - 1) \cdot q + P(T_0 < T_N | X_0 = m + 1) \cdot p. \end{aligned}$$

Colocando em palavras, se o jogador perder a primeira jogada, o que acontece com probabilidade q, a probabilidade de sua ruína passa a ser a mesma de um jogador que começa com um real a menos. Por outro lado, se ele ganhar a primeira jogada, o que acontece com probabilidade p, ele passa a ter a mesma probabilidade de ruína que ele teria se começasse o jogo com $m + 1$ reais.

Denotando por u_m a probabilidade de ruína com fortuna inicial m, temos

$$u_m = P(T_0 < T_N | X_0 = m),$$

e equação acima nos dá que

$$u_m = u_{m-1} \cdot q + u_{m+1} \cdot p \tag{3.3}$$

com $u_0 = 1$ e $u_N = 0$.

Resumindo, encontramos a seguinte relação de recorrência.

$$\begin{cases} u_m = u_{m-1} \cdot q + u_{m+1} \cdot p & \text{, para } 1 \leq m \leq N - 1 \\ u_0 = 1, u_N = 0 \end{cases}$$

Para simplificar a expressão acima observamos que $p + q = 1$, e portanto

$$(p+q)u_m = u_{m-1}q + u_{m+1}p$$

O que nos garante que

$$(u_{m+1} - u_m)p = (u_m - u_{m-1})q.$$

Agora, defina $a_m = u_{m+1} - u_m$, para $1 \leq m \leq N-1$. Assim obtemos

$$a_m p = a_{m-1} q,$$

ou ainda

$$a_m = \frac{q}{p} a_{m-1}.$$

Deste modo a sequência (a_m) é uma progressão geométrica de razão $\frac{q}{p}$ e, portanto

$$a_m = \left(\frac{q}{p}\right)^m a_0$$

Definindo $\lambda = \frac{q}{p}$, temos $u_{m+1} - u_m = \lambda^m (u_1 - u_0)$ e

$$\begin{aligned}
u_m - u_{m-1} &= \lambda^{m-1}(u_1 - u_0) \\
u_{m-1} - u_{m-2} &= \lambda^{m-2}(u_1 - u_0) \\
\vdots \quad & \qquad \vdots \\
u_2 - u_1 &= \lambda^1 (u_1 - u_0) \\
u_1 - u_0 &= \lambda^0 (u_1 - u_0)
\end{aligned}$$

Somando e simplificado as equações acima, temos

$$u_m - u_0 = (1 + \lambda + \lambda^2 + \cdots + \lambda^{m-1}) \cdot (u_1 - u_0) \tag{3.4}$$

$$\Rightarrow u_m = (1 + \lambda + \lambda^2 + \cdots + \lambda^{m-1}) \cdot (u_1 - 1) + 1 \tag{3.5}$$

já que $u_0 = 1$.

Caso 1: $p \neq q$

Pois bem, se $p \neq q$, então $\frac{q}{p} = \lambda \neq 1$. Notemos também que a expressão $(1 + \lambda + \lambda^2 + \cdots + \lambda^{m-1})$ trata-se da soma dos m primeiros termos da progressão geométrica de razão λ e com primeiro termo igual a 1. Assim, temos que

$$1 + \lambda + \lambda^2 + \cdots + \lambda^{m-1} = \frac{\lambda^m - 1}{\lambda - 1}$$

Por (3.5) obtemos

$$u_m = \left(\frac{\lambda^m - 1}{\lambda - 1}\right)(u_1 - 1) + 1 \tag{3.6}$$

Usando isso e o fato que $u_N = 0$ encontramos

$$\begin{aligned} u_N = \left(\frac{\lambda^N - 1}{\lambda - 1}\right)(u_1 - 1) + 1 \Rightarrow 0 = u_1\left(\frac{\lambda^N - 1}{\lambda - 1}\right) - \left(\frac{\lambda^N - 1}{\lambda - 1}\right) + 1 \\ \Rightarrow u_1 = 1 - \left(\frac{\lambda - 1}{\lambda^N - 1}\right) \end{aligned} \tag{3.7}$$

Substituindo (3.7) em (3.6), resulta em

$$u_m = \left(\frac{\lambda^m - 1}{\lambda - 1}\right)\left(1 - \frac{\lambda - 1}{\lambda^N - 1} - 1\right) + 1 = 1 - \frac{\lambda^m - 1}{\lambda^N - 1} = \frac{\lambda^m - \lambda^N}{1 - \lambda^N}$$

Portanto,

$$P(\text{ruína}|X_0 = m) = \frac{\left(\frac{q}{p}\right)^m - \left(\frac{q}{p}\right)^N}{1 - \left(\frac{q}{p}\right)^N}$$

o que conclui a demonstração para o caso 1 da Proposição 3.1

Caso 2: $p = q$

Pois bem, se $p = q$, então $p = q = \frac{1}{2}$ e $\frac{q}{p} = \lambda = 1$. Assim, temos que

$$1 + \lambda + \lambda^2 + \cdots + \lambda^{m-1} = m.$$

E por (3.5), obtemos

$$u_m = m(u_1 - 1) + 1 \tag{3.8}$$

Usando isso e novamente o fato de $u_N = 0$, segue que

$$u_N = N(u_1 - 1) + 1 \Rightarrow 0 = Nu_1 - N + 1 \Rightarrow u_1 = 1 - \frac{1}{N} \tag{3.9}$$

Substituindo (3.9) em (3.8), temos o seguinte resultado

$$u_m = m(1 - \frac{1}{N} - 1) + 1 \Rightarrow P(\text{ruína}|X_0 = m) = 1 - \frac{m}{N}$$

Concluindo a demonstração para o caso 2 da Proposição. □

Exemplo 3.1. Um jogador fará uma série de apostas na roleta, todas no preto. Sua fortuna inicial é de R$ 50,00, e ele jogará até que sua fortuna seja 0 ou R$ 100,00, o legítimo conceito de "o dobro ou nada". As apostas são de R$ 1,00, ou seja, a cada aposta, ou ele ganha R$1,00 com probabilidade $\frac{18}{38} = \frac{9}{19}$, ou perde R$ 1,00 com probabilidade $\frac{20}{38} = \frac{10}{19}$ (lembrando que na roleta temos 18 pretos, 18 vermelhos e 2 verdes). Notemos que $\frac{q}{p} = \frac{10}{9}$. Assim e pela Proposição 3.1, a probabilidade da ruína desse jogador é

$$P(\text{ruína}|X_0 = 50) = \frac{\left(\frac{10}{9}\right)^{50} - \left(\frac{10}{9}\right)^{100}}{1 - \left(\frac{10}{9}\right)^{100}} \approx 0,99$$

Logo, o jogador perderá R$ 50,00 com probabilidade de 99%

Exemplo 3.2. Se em vez de apostar na roleta, o jogador do exemplo anterior fizesse a série de apostas utilizando uma série de lançamentos de uma moeda não viciada, onde ganharia R$ 1,00 se o resultado fosse cara, ou perderia R$ 1,00 caso fosse coroa. Nessa situação, teríamos $p = q = \frac{1}{2}$ e, pela Proposição 3.1, a probabilidade da ruína nesse caso é dada por

$$P(\text{ruína}|X_0 = 50) = 1 - \frac{50}{100} = 0,5$$

Logo, nessa situação, a probabilidade de fracasso do jogador decresceria para 50%. Um cenário bem mais atrativo para as apostas, comparado ao jogo na roleta apresentado no exemplo anterior.

É interessante notar que no exemplo 3.1 a probabilidade de vitória era de $9/19$ que não é muito menor que $1/2$, a probabilidade de vitória no exemplo 3.2. Ainda assim a diferença na probabilidade de ruína é significativa! Isso vai ser um ponto essencial na proposta de atividade feita no capítulo final.

3.1.2 *A Probabilidade da ruína do jogador ganancioso*

Consideremos novamente o Problema da ruína do jogador, mas agora diante de um jogador que nunca estará satisfeito com a fortuna adquirida, e estará sempre em busca de um lucro cada vez maior através da série de apostas. Ele só deixará de apostar caso esteja arruinado. Caso contrário, jogará para sempre. Nesse novo cenário, como fica a probabilidade da ruína do jogador?

Pois bem, observemos que o estado N, que antes era absorvente, agora não existe (ou é infinito), e a ruína agora só ocorre se o tempo T_0 da primeira visita ao estado 0 for finito. Assim, nesse caso, a probabilidade da ruína é dada por

$$P(T_0 < \infty | X_0 = m).$$

Note agora que fixado $m > 0$, então $T_N \geq N - m$ para todo $N > m$. Isso nos dá que

$$\lim_{N \to \infty} T_N = \infty$$

e portanto os eventos $\{T_0 < T_N\}$ também convergem. Ou seja

$$\{T_0 < T_N\} \to \{T_0 < \infty\},$$

e além disso, como $T_N < T_{N+1}$, temos que a sequência de eventos $\{T_0 < T_N\}$ é crescente e pela Proposição 2.3 temos que

$$P(T_0 < \infty | X_0 = m) = \lim_{N \to \infty} P(T_0 < T_N | X_0 = m).$$

Dito isso, seguimos para mais um resultado.

Proposição 3.2. *Com base na descrição do Problema da Ruína do Jogador ganancioso, a probabilidade da ruína neste caso é dada por*

$$P(\textit{ruína do ganancioso} | X_0 = m) = \begin{cases} \left(\frac{q}{p}\right)^m, & \textit{se } \ p > q \\ 1, & \textit{se } \ p \leq q \end{cases}$$

Demonstração. Pois bem, como comentamos acima

$$P(\text{ruína do ganancioso} | X_0 = m) = P(T_0 < \infty | X_0 = m) = \lim_{N \to \infty} P(T_0 < T_N | X_0 = m).$$

E usando a Proposição 3.1, para o caso em que $p \neq q$, temos

$$P(\text{ruína do ganancioso} | X_0 = m) = \lim_{N \to \infty} P(T_0 < T_N | X_0 = m) = \lim_{N \to \infty} \frac{\left(\frac{q}{p}\right)^m - \left(\frac{q}{p}\right)^N}{1 - \left(\frac{q}{p}\right)^N}$$

Nesse caso, temos duas possibilidades a analisar:

- Se $p > q$, então quando N tende a infinito, $(\frac{q}{p})^N$ tende a 0. Logo,

$$P(\text{ruína do ganancioso}|X_0 = m) = \left(\frac{q}{p}\right)^m$$

  e, portanto, nesse cenário, a ruína do jogador não é certa, o que não é surpreendente, já que, em cada aposta, a probabilidade de vencer é maior que a probabilidade de perder. Logo, o jogo tem um viés a favor do jogador. Além disso, observe ainda que a probabilidade da ruína diminui à medida que a fortuna inicial m aumenta.

- Se $p < q$, então quando N tende a infinito, $(\frac{q}{p})^N$ tende a infinito, enquanto $(\frac{p}{q})^N \to 0$. Logo,

$$P(\text{ruína do ganancioso}|X_0 = m) = \lim_{N\to\infty} \frac{\left(\frac{p}{q}\right)^{N-m} - 1}{\left(\frac{p}{q}\right)^N - 1} = 1$$

  e, portanto, a ruína do jogador é inevitável nessa condição. O que também não surpreende, pois agora o jogo segue com um viés contra o jogador (probabilidade de perder maior que a probabilidade de vencer em cada aposta realizada).

  Já para o caso em que $p = q$ e ainda usando a Proposição 3.1, temos

$$P(\text{ruína do ganancioso}|X_0 = m) = \lim_{N\to\infty} P(T_0 < T_N|X_0 = m) = \lim_{N\to\infty} 1 - \frac{m}{N} = 1$$

  e, portanto, mais uma vez temos a certeza que em algum momento o jogador ganancioso estará arruinado. Esta sim pode causar uma certa surpresa à intuição, visto que o jogador não pode esperar nenhum êxito no jogo, mesmo tendo a vitória e a derrota igualmente possíveis em cada aposta.

□

Em um contexto mais biológico, o processo de ruína do jogador ganancioso é um caso particular de um outro processo, conhecido como processo de nascimento e morte, que modela o crescimento de uma população, indivíduo a indivíduo. Neste processo X_n representa o tamanho da população no "instante" $n > 0$, e quando $X_n = k > 0$ o nascimento de um indivíduo ocorre com probabilidade p_k, a morte com probabilidade q_k e o processo permanece no mesmo estado com probabilidade r_k, de modo que $p_k + q_k + r_k = 1$. Dessa forma, o tamanho atual dessa população aumenta uma unidade a cada indivíduo que nasce, e diminui uma unidade a cada indivíduo que morre. Neste contexto, dizemos que a população está extinta quando seu tamanho chegar 0. Ou seja, quando $X_n = 0$. Isso obviamente decorre do fato de que, em $X_n = 0$, teremos 0 indivíduos e, portanto, não serão mais possíveis os nascimentos e as mortes nessa população.

O processo de ruína do jogador ganancioso é recuperado quando $p_k = p$, $q_k = 1 - p$ e $r_k = 0$ para todo $k > 0$.

Neste contexto, uma questão interessante a ser considerada é: qual o tempo esperado para que a extinção de uma população ocorra?

Voltando ao problema da Ruína do Jogador do Ganancioso, o tempo esperado de extinção corresponde a duração esperada do jogo. Nas próximas seções seguimos com o nosso estudo nessa direção.

3.1.3 *Duração esperada do jogo*

Já sabemos quais as probabilidades do nosso jogador ruir em sua saga pelo acaso. Agora, o que podemos esperar em relação à duração do Jogo?

No caso do jogador ganancioso, onde o jogo só chega ao seu fim diante da hipótese da fortuna do jogador zerar, teremos que pensar como determinar a quantidade média de apostas necessárias para que isso ocorra.

Para termos uma ideia do que esperar, vamos fazer uma análise informal e simples usando o comportamento médio do processo. Para isso lembre que antes do final do jogo, vale que

$$\begin{aligned} X_n &= X_0 + Y_1 + Y_2 + \cdots + Y_n \\ \Rightarrow E(X_n) &= E(m + Y_1 + Y_2 + \cdots + Y_n) \\ \Rightarrow E(X_n) &= E(m) + E(Y_1) + E(Y_2) + \cdots + E(Y_n) \\ \Rightarrow E(X_n) &= m + nE(Y_1) \end{aligned}$$

já que $Y_1, Y_2, \ldots, Y_n$ são iid.

Como temos $E(Y_1) = 1 \cdot p + (-1) \cdot q = p - (1 - p) = 2p - 1$, segue que

$$E(X_n) = m + nE(Y) = m + n(2p - 1).$$

Assim, em média, o jogador terá $m + n(2p - 1)$ reais após n jogadas.

Como a ruína ocorre quando o processo atinge 0, espera-se n seja tal que $E(X_n) = 0$. Assim, temos $0 = E(X_n) = m + n(2p - 1)$ e

$$n = \frac{m}{1 - 2p}. \tag{3.10}$$

Essa é uma análise informal, e trás informações limitadas sobre o que devemos esperar do processo.

Observe, por exemplo, que se $p = \frac{1}{2}$, teríamos que

$$E(X_n) = m + n(2p - 1) = 0,$$

e em média o jogador não ganha ou perde nada, o que não nos dá muita informação sobre o tempo de jogo.

Por outro lado, como $n > 0$, precisamos que $1 - 2p > 0$. Isso implica que a estimativa encontrada só é válida para $p < \frac{1}{2}$. Isso acontece por que se $p > \frac{1}{2}$, então o saldo médio por jogada é positivo e o jogador nunca chegaria a ruína. E portanto, mais uma vez, a análise média falha em nos dar alguma informação sobre o tempo de jogo esperado.

De todo modo, para $p < \frac{1}{2}$ devemos esperar um total aproximado de $n = \frac{m}{1-2p}$ apostas.

Para uma melhor análise da quantidade média de apostas do jogador, voltaremos ao Problema inicial da ruína do Jogador, ou seja, ao processo X_n que possui 0 e N como estados absorventes, e portanto o processo se encerra no instante em que o processo atinge 0, o que acontece em T_0, ou quando atingir N, o que acontece em T_N. Com isso, se chamarmos de S_N o tempo total de jogo, temos

$$S_N = \min\{T_0, T_N\}.$$

Com isso estamos prontos para o nosso primeiro resultado relativo à duração esperada do jogo.

Proposição 3.3. *De acordo com o Problema da ruína do jogador (seção 3.1), seja $D_m = E(S_N | X_0 = m)$ a duração esperada do jogo de um apostador com fortuna inicial m. Se ele apostar até que sua fortuna atinja 0 ou N, então,*

$$D_m = \begin{cases} \frac{m}{1-2p} - \left(\frac{N}{1-2p}\right)\left(\frac{1-\left(\frac{q}{p}\right)^m}{1-\left(\frac{q}{p}\right)^N}\right), & se \;\; p \neq q \\ \\ m(N-m), & se \;\; p = q \end{cases}$$

Demonstração. Primeiro observemos que $D_0 = D_N = 0$, já que não haverá apostas se a fortuna inicial do jogador for 0 ou N.

Condicionando o jogo ao resultado da primeira aposta percebemos um detalhe importante. Ao caminhar de m para $m-1$ ou $m+1$, o processo toma um instante de tempo, e portanto, como vimos nos argumentos levando à equação (3.2), condicionado a $X_1 = m+i$ temos que S_N tem a mesma distribuição de

$$\min\{T_0+1, T_N+1\} = S_N + 1,$$

condicionado apenas à $X_0 = m+i$, $i = 1$ ou -1. Ou seja,

$$E(S_N|X_0 = m; X_1 = m+i) = E(S_N + 1|X_0 = m+i),$$

e

$$\begin{aligned}
D_m &= E(S_N|X_0 = m) \\
&= E(S_N|X_0 = m; X_1 = m-1)P(X_1 = m-1|X_0 = m) \\
&\qquad + E(S_N|X_0 = m; X_1 = m+1)P(X_1 = m+1|X_0 = m) \\
&= E(S_N + 1|X_0 = m-1)\cdot q + E(S_N + 1|X_0 = m+1)\cdot p \\
&= q\cdot(1 + D_{m-1}) + p\cdot(1 + D_{m+1}) \\
&= p + q + p\cdot D_{m+1} + q\cdot D_{m-1} \\
&= 1 + p\cdot D_{m+1} + q\cdot D_{m-1}.
\end{aligned}$$

Ou seja, a quantidade esperada de apostas para o jogo com fortuna inicial m será igual à soma das quantidades esperadas de apostas dos jogos iniciados com fortuna $m+1$ e com fortuna $m-1$, cada uma adicionada a um instante de tempo, referente à primeira aposta, multiplicadas por suas respectivas probabilidades de ocorrência. A partir disso, obtemos

$$D_m = 1 + p\cdot D_{m+1} + q\cdot D_{m-1} \tag{3.11}$$

ou ainda

$$p\cdot D_{m+1} - D_m + q\cdot D_{m-1} = -1.$$

Observe agora que o equação acima é uma equação linear não-homogênea, com equação homogênea associada dada por

$$h_m = p\cdot h_{m+1} + q\cdot h_{m-1},$$

que é exatamente a equação que encontramos para a probabilidade de ruína, dada em (3.3).

As soluções gerais são dadas em (3.6) e (3.8), e com uma simples manipulação algébrica encontramos que

$$\begin{cases} C_1 + C_2 \left(\frac{q}{p}\right)^m, & \text{se } p \neq q, \\ C_1 + C_2 m, & \text{se } p = q. \end{cases}$$

A busca por soluções particulares difere nos dois casos, e é possível verificar que $t_m = \frac{m}{1-2p}$ é solução particular da equação (3.11) para $p \neq q$, enquanto $t_m = -m^2$ é solução particular quando $p = q$.

Com isso temos que a solução geral é dada por

$$\begin{cases} C_1 + C_2 \left(\frac{q}{p}\right)^m + \frac{m}{1-2p}, & \text{se } p \neq q, \\ C_1 + C_2 m - m^2, & \text{se } p = q. \end{cases} \tag{3.12}$$

A seguir calcularemos a solução específica de cada caso, lembrando que $D_0 = D_N = 0$.

Caso 1: $p \neq q$

Neste caso, fazendo $D_0 = D_N = 0$ em (3.12) obtemos que

$$\begin{cases} C_1 + C_2 = 0 \\ C_1 + C_2 \left(\frac{q}{p}\right)^N + \frac{N}{1-2p} = 0, \end{cases}$$

e portanto $C_2 = -C_1$ e

$$C_1 = -\left(\frac{N}{1-2p}\right)\left(\frac{1}{1-\left(\frac{q}{p}\right)^N}\right).$$

Segue assim que

$$D_m = \frac{m}{1-2p} - \left(\frac{N}{1-2p}\right)\left(\frac{1-\left(\frac{q}{p}\right)^m}{1-\left(\frac{q}{p}\right)^N}\right)$$

Caso 2: $p = q = \frac{1}{2}$

Procedendo da mesma forma que no caso 1, usando que $D_0 = D_N = 0$, encontramos que

$$\begin{cases} C_1 = 0 \\ C_2 N - N^2 = 0 \end{cases}$$

e assim $C_2 = N$ e

$$D_m = Nm - m^2 = m(N - m),$$

concluindo a prova. $\square$

Exemplo 3.3. Voltemos à situação descrita no exemplo 3.1. Pela Proposição 3.3, a duração esperada da série de apostas realizadas na roleta será

$$D_{50} = \frac{50}{1 - 2 \cdot \frac{9}{19}} - \left(\frac{100}{1 - 2 \cdot \frac{9}{19}}\right)\left(\frac{1 - \left(\frac{10}{9}\right)^{50}}{1 - \left(\frac{10}{9}\right)^{100}}\right) \approx 940$$

Portanto, espera-se que ocorra aproximadamente 940 apostas até o final do jogo.

Exemplo 3.4. Não podemos também de deixar de voltar ao exemplo 3.2 para determinar a duração esperada da série de apostas com base na cara e na coroa da moeda. Daí e também pela Proposição 3.3, temos que

$$D_{50} = 50(100 - 50) = 2500$$

Portanto, é esperado um total de 2500 apostas.

Apesar do jogo realizado na moeda ter se mostrado mais vantajoso ao jogador, conforme discutimos no exemplo 3.2, ele exigirá mais paciência por parte dos envolvidos, já que espera-se um número bem maior de apostas para finalizá-lo do que em relação ao jogo na roleta.

3.1.4 *Duração esperada do jogo com um jogador ganancioso*

A partir dos resultados obtidos sobre a duração esperada do jogo com estados de absorção inferior e superior para fortuna inicial m, podemos dizer o que esperar em relação à duração do jogo com o jogador ganancioso.

Lembrando que o jogador ganancioso jogará para sempre, exceto se sua fortuna zerar decretando sua ruína, percebemos que o jogo só acaba quando o processo atinge 0, e portanto, se denotarmos por S o tempo de jogo do jogador ganancioso, temos simplesmente que

$$S = T_0.$$

Lembremos agora que o tempo de jogo do jogador "comedido"é

$$S_N = \min\{T_0, T_N\},$$

e que $T_N \to \infty$ quando $N \to \infty$, e assim

$$S = T_0 = \lim_{N\to\infty} \min\{T_0, T_N\} = \lim_{N\to\infty} S_N.$$

De posse dessas informações, estamos prontos para o próximo resultado.

Proposição 3.4. *A duração esperada D_m^g do jogo em A Ruína do Jogador Ganancioso é dado por*

$$D_m^g = \begin{cases} \frac{m}{1-2p}, & se \;\; p < q \\ \infty, & se \;\; p \geq q \end{cases}$$

Demonstração. Pois bem, sendo assim, temos que se $D_m^g = E(S|X_0 = m)$ corresponde à duração esperada do jogo com um jogador ganancioso de fortuna inicial m, então

$$D_m^g = E(S|X_0 = m) = \lim_{N\to\infty} E(S_N|X_0 = m) = \lim_{N\to\infty} D_m.$$

Analisando caso a caso obtemos o seguinte.

caso 1: $p \neq q$

Pela Proposição 3.3, segue que

$$D_m^g = \lim_{N\to\infty} D_m = \lim_{N\to\infty} \left[\frac{m}{1-2p} - \left(\frac{N}{1-2p}\right) \left(\frac{1-\left(\frac{q}{p}\right)^m}{1-\left(\frac{q}{p}\right)^N} \right) \right]$$

$$= \frac{m}{1-2p} - \frac{1-\left(\frac{q}{p}\right)^m}{1-2p} \lim_{N\to\infty} \frac{N}{1-\left(\frac{q}{p}\right)^N}$$

Nesse caso, temos duas possibilidades para analisar:

- Se $p > q$, temos que $\left(\frac{q}{p}\right)^N$ tende a 0 quando N vai ao infinito. Assim,

$$\lim_{N\to\infty} \frac{N}{1-\left(\frac{q}{p}\right)^N} = \infty.$$

Como $p > q \Rightarrow p > \frac{1}{2} \Rightarrow 1-2p < 0$, então

$$D_m^g = \frac{m}{1-2p} - \frac{1-\left(\frac{q}{p}\right)^m}{1-2p} \lim_{N\to\infty} \frac{N}{1-\left(\frac{q}{p}\right)^N} = \frac{m}{1-2p} + \infty = \infty$$

Lembrando que, nesse caso, o jogo tem um viés a favor do jogador. Portanto, sua ruína pode até ocorrer, porém, é esperado que leve muito tempo para que isso aconteça.

- Se $p < q$, temos que $\left(\frac{q}{p}\right)^N$ tende ao infinito com N. Assim,

$$\lim_{N\to\infty} \frac{N}{1-\left(\frac{q}{p}\right)^N} = 0,$$

visto que, diante dessa hipótese, o crescimento exponencial de $\left(\frac{q}{p}\right)^N$ é mais rápido que o crescimento polinomial de N.

Logo,

$$D_m^g = \frac{m}{1-2p} - \frac{1-\left(\frac{q}{p}\right)^m}{1-2p} \lim_{N\to\infty} \frac{N}{1-\left(\frac{q}{p}\right)^N} = \frac{m}{1-2p} - 0 = \frac{m}{1-2p}$$

Vimos que a ruína do jogador ganancioso nesse caso ocorre com probabilidade 1. Portanto, ela é certa e com duração esperada de $\frac{m}{1-2p}$.

Notemos que este resultado corrobora o que discutimos na equação (3.10).

Caso 2: $p = q$

Nessa hipótese e ainda pela Proposição 3.3, obtemos

$$D_m^g = \lim_{N\to\infty} D_m = \lim_{N\to\infty} m(N-m) = \infty$$

Este talvez seja o mais interessante destes resultados. Isso por que vimos que a ruína do jogador ganancioso nesse caso é certa, mas apesar disso, o tempo esperado de jogo é infinito! Assim, o jogador vai perder, mas esperará muito tempo por isso, alimentando expectativas, e deixando a derrota ainda mais amarga! $\square$

Observação 3.1.2. *O resultado $D_m^g = \infty$ no caso $p > q$, onde o viés está do lado do jogador, poderia ter sido rapidamente encontrado de outro modo. Basta notar que da Proposição 3.3, temos*

$$P(T_0 < \infty | X_0 = m) = \left(\frac{q}{p}\right)^m < 1,$$

e portanto

$$P(T_0 = \infty | X_0 = m) = 1 - P(T_0 < \infty | X_0 = m) > 0,$$

nos dando $E(S|X_0 = m) = E(T_0|X_0 = m) = \infty$.

Exemplo 3.5. Voltemos à situação em que um jogador fará uma série de apostas na roleta, todas no preto, começando com uma fortuna de R$ 50,00. Só que dessa vez, ele abandonará as apostas somente em caso de ruína. Lembrando que, ou ele ganha R$ 1,00 com probabilidade $p = \frac{18}{38} = \frac{9}{19}$, ou perde R$ 1,00 com probabilidade $q = \frac{20}{38} = \frac{10}{19}$. Notemos que $p < q$. Assim e conforme discutido na probabilidade da ruína do jogador ganancioso, a ruína do jogador ocorre com probabilidade 1, e a duração esperada para que isso ocorra, de acordo com a discussão feita acima, é dada por

$$D_{50}^g = \frac{50}{1 - 2 \cdot \frac{9}{19}} = 950 \text{ apostas}$$

Exemplo 3.6. Se o jogador ganancioso do exemplo anterior preferisse realizar suas apostas utilizando lançamentos de uma moeda não viciada, onde ganharia R$ 1,00 se o resultado fosse cara, ou perderia R$ 1,00 caso fosse coroa, teríamos também a ruína desse jogador inevitavelmente, já que temos $p = q = \frac{1}{2}$. Porém, ainda segundo as demonstrações feitas na probabilidade da ruína e na duração esperada do jogo, ambas com um jogador ganancioso, a duração esperada nesse caso é $D_{50}^g = \infty$, ou seja, pode levar muito tempo para a ruína ocorrer.

3.2 PASSEIOS ALEATÓRIOS SIMPLES

Suponhamos que queremos modelar o seguinte fenômeno: uma partícula se move nos sítios de $\mathbb{Z}$. No instante de tempo $n = 0$, a partícula está no sítio X_0. A cada instante de tempo discreto, ou a partícula salta para o sítio vizinho à direita com probabilidade p, ou salta para o sítio vizinho à esquerda com probabilidade $q = 1 - p$. Na figura 6 ilustramos esse fenômeno partindo de $X_0 = 0$.

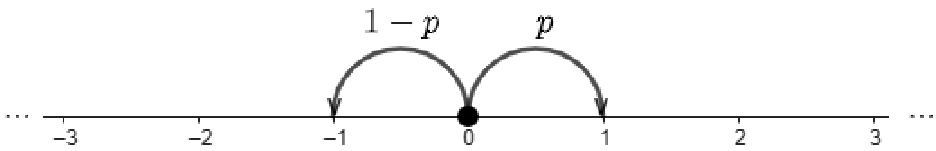


Figura 6: Passeio aleatório simples iniciado em $X_0 = 0$

Seja X_n a posição da partícula no n-ésimo salto. Dessa forma, a sequência $(X_n)_{n\geq 0}$ estabelece um *Passeio aleatório simples* (PAS) ou, simplesmente, *um Passeio aleatório em* $\mathbb{Z}$.

Definição 3.5. Passeio aleatório simples a tempo discreto é um processo estocástico $(X_n)_{n\geq 0}$ com valores em $\mathbb{Z}$ tal que

$$X_n = X_{n-1} + Y_n, \quad \text{para} \ \ n \geq 1$$

onde $Y_1, Y_2, \dots$ é uma sequência de variáveis aleatórias independentes com valores em $\{-1, 1\}$ e identicamente distribuídas, com distribuição dada por

$$P(Y = 1) = p \quad \text{e} \quad P(Y = -1) = q = 1 - p$$

Também podemos expressar $X_n = X_0 + Y_1 + Y_2 + \cdots + Y_n$.

Quando temos $p = q = \frac{1}{2}$, dizemos que o processo é um *Passeio aleatório simples simétrico* (PASS).

Observemos que a posição da partícula no instante n depende somente da posição que ela ocupa no instante $n - 1$. Matematicamente isso pode ser expresso por

$$P(X_n = k | X_0 = i_0, \dots, X_{n-2} = i_{n-2}, X_{n-1} = i) = P(X_n = k | X_{n-1} = i),$$

o que nos diz que os passeios aleatórios simples são cadeias de Markov.

Notemos ainda que o Problema da Ruína do Jogador Ganancioso pode ser visto como um caso particular dos PAS, mais especificamente, dos passeios aleatórios em $\mathbb{Z}$, com início em $X_0 = m > 0$ (fortuna inicial), onde acompanhamos o processo até o instante T_0 de primeira visita ao sítio 0. Neste paralelo cada aposta vitoriosa representa um salto para o sítio imediatamente à direita, assim como, cada aposta perdida, um salto para o sítio imediatamente à esquerda.

Até T_0 os dois processos são indistinguíveis, e assim podemos aplicar os resultados obtidos a partir do problema da ruína do jogador nos passeios aleatórios, como veremos mais a frente.

A partir desses resultados iniciais sobre os Passeios aleatórios simples, iremos agora aprofundar um pouco mais nosso estudo sobre esse processo. Nesse sentido, uma discussão interessante a ser feita é: como saber se a partícula retornará ou não ao ponto de origem em seu PAS, após um intervalo de tempo finito?

3.2.1 *Retorno à origem*

Para responder à questão levantada acima e simplificar a exposição, tomaremos o passeio aleatório simples de uma partícula cuja posição inicial é $X_0 = 0$ e $p \in (0,1)$. Usando novamente a definição

$$T_i = \min\{n \geq 1 : X_n = i\},$$

a discussão rodará entorno da probabilidade da partícula retornar ao estado 0 pela primeira vez após finitos saltos, que pode ser expressado por

$$P(T_0 < \infty | X_0 = 0)$$

onde $T_0 = \min\{n \geq 1 : X_n = 0\}$.

Neste sentido é útil uma rápida definição antes de continuarmos.

Definição 3.6. Dizemos que o PAS é recorrente quando $P(T_0 < \infty | X_0 = 0) = 1$. E dizemos que ele é transiente quando $P(T_0 < \infty | X_0 = 0) < 1$.

Em outras palavras, a recorrência significa que o passeio retornará ao sítio 0 em algum momento com probabilidade 1.

Notemos que, assim como acontece com o modelo da ruína do jogador, o passeio é renovado a cada retorno ao 0. Ou seja, a partir do instante em que o passeio chega a 0, o processo recomeça de forma independente da trajetória que o levou até ali. Deste modo, quando o PAS é recorrente, o sítio 0 será visitado infinitas vezes.

Por outro lado, se o PAS for transiente, a probabilidade de retorno pode ser positiva, mas a cada visita existe uma probabilidade positiva dele nunca mais retornar à origem. Com isso, assim como acontece com uma moeda viciada, existe um momento em que o passeio visita a origem pela última vez, fazendo o total de visitas ser finito.

Podemos inclusive determinar a distribuição da quantidade de vezes que um PAS transiente visita o sítio 0 em sua trajetória.

Para isso, consideremos um passeio aleatório simples transiente que retorna à origem com probabilidade $\theta < 1$ e, não retorna com probabilidade $1 - \theta > 0$. Agora seja W a v.a que corresponde à quantidade de visitas à origem deste passeio durante sua trajetória. Assim, considerando que o passeio se renova a cada retorno à origem, temos que

- $P(W = 1) = 1 - \theta$,

  ou seja, a probabilidade do passeio nunca mais retornar é igual a probabilidade do passeio visitar a origem apenas 1 vez: o momento inicial.

- $P(W = 2) = \theta \cdot (1 - \theta)$,

- $P(W = 3) = \theta \cdot \theta \cdot (1 - \theta) = \theta^2 \cdot (1 - \theta)$,

  $\vdots \qquad \vdots \qquad \vdots$

- $P(W = k) = \theta^{k-1} \cdot (1 - \theta) \qquad$ para $k = 1, 2, 3, \ldots$

Notemos que

$$P(W < \infty) = \sum_{i=1}^{\infty} P(W = i) = \sum_{i=1}^{\infty} \theta^{i-1} \cdot (1 - \theta) = 1,$$

e que a esperança de W é dada por

$$\begin{aligned}
E(W) &= \sum_{i=1}^{\infty} i \cdot \theta^{i-1} \cdot (1 - \theta) \\
\Leftrightarrow E(W) &= \sum_{i=1}^{\infty} (i - 1 + 1) \cdot \theta^{i-1} \cdot (1 - \theta) \\
\Leftrightarrow E(W) &= \sum_{i=1}^{\infty} (i - 1) \cdot \theta^{i-1} \cdot (1 - \theta) + \sum_{i=1}^{\infty} \theta^{i-1} \cdot (1 - \theta) \\
\Leftrightarrow E(W) &= \sum_{i=1}^{\infty} (i - 1) \cdot \theta^{i-1} \cdot (1 - \theta) + 1.
\end{aligned}$$

Fazendo $j = i - 1$, segue

$$E(W) = \sum_{j=0}^{\infty} j \cdot \theta^j \cdot (1 - \theta) + 1$$

$$\begin{aligned}
&\Leftrightarrow E(W) = \theta \sum_{j=1}^{\infty} j \cdot \theta^{j-1} \cdot (1-\theta) + 1 \\
&\Leftrightarrow E(W) = \theta \cdot E(W) + 1 \Leftrightarrow E(W) - \theta \cdot E(W) = 1 \\
&\Leftrightarrow E(W) \cdot (1-\theta) = 1 \\
&\Leftrightarrow E(W) = \frac{1}{1-\theta} \qquad (3.13)
\end{aligned}$$

Portanto, em um passeio aleatório simples transiente que retorna à origem com probabilidade θ, espera-se que ele visite a origem $\frac{1}{1-\theta}$ vezes.

O próximo resultado caracteriza a recorrência/transiência de um PAS.

Teorema 3.7. *Se $p = q$ então Passeio aleatório simples é recorrente. Caso contrário, é transiente.*

Demonstração. Seja $(X_n)_{n\geq 0}$ um passeio aleatório simples com $X_0 = 0$.

De fato, temos que

$$P(T_0 < \infty | X_0 = 0) = P(T_0 < \infty; X_1 = 1 | X_0 = 0) + P(T_0 < \infty; X_1 = -1 | X_0 = 0)$$

Assim como fizemos com o modelo da ruína do jogador, podemos condicionar ao primeiro passo para encontrar que

$$\begin{aligned}
P(T_0 < \infty | X_0 = 0) &= P(T_0 < \infty | X_1 = 1) P(X_1 = 1 | X_0 = 0) + P(T_0 < \infty | X_1 = -1) P(X_1 = -1 | X_0 = 0) \\
&= P(T_0 < \infty | X_0 = 1) \cdot p + P(T_0 < \infty | X_0 = -1) \cdot q
\end{aligned}$$

Como comentamos anteriormente, o PAS começando de $X_0 = 1$ é indistinguível do modelo da ruína do jogador ganancioso até, e incluindo, o instante T_0. Deste modo $P(T_0 < \infty | X_0 = 1)$ é o mesmo que calculamos para a ruína do jogador.

Para o jogador ganancioso mostramos que se $p \leq q$ então $P(T_0 < \infty | X_0 = m) = 1$, e se $p > q$ então $P(T_0 < \infty | X_0 = m) = \left(\frac{q}{p}\right)^m$, para $m > 0$. Logo,

- para $p \leq q \Rightarrow p \leq \frac{1}{2}$, temos

$$\begin{aligned}
P(T_0 < \infty | X_0 = 0) &= P(T_0 < \infty | X_0 = 1) \cdot p + P(T_0 < \infty | X_0 = -1) \cdot q \\
&= p + P(T_0 < \infty | X_0 = -1) \cdot q \qquad (3.14)
\end{aligned}$$

- para $p > q \Rightarrow p > \frac{1}{2}$, temos

$$\begin{aligned} P(T_0 < \infty | X_0 = 0) &= P(T_0 < \infty | X_0 = 1) \cdot p + P(T_0 < \infty | X_0 = -1) \cdot q \\ &= \frac{q}{p} \cdot p + P(T_0 < \infty | X_0 = -1) \cdot q \\ &= q + P(T_0 < \infty | X_0 = -1) \cdot q \end{aligned} \tag{3.15}$$

Agora, defina $X_n^* = -X_n$ e $T_0^* = min\{n \geq 1 : X_n^* = 0\}$. Dessa forma, temos que $X_n^* = 0 \Leftrightarrow X_n = 0$ e, portanto,

$$T_0^* = \min\{n \geq 1 : X_n^* = 0\} = \min\{n \geq 1 : X_n = 0\} = T_0$$

Além disso, observe que

$$P(X_{n+1}^* = i+1 | X_n^* = i) = P(-X_{n+1} = i+1 | -X_n = i) = P(X_{n+1} = -i-1 | X_n = -i) = q$$

e

$$P(X_{n+1}^* = i-1 | X_n^* = i) = P(-X_{n+1} = i-1 | -X_n = i) = P(X_{n+1} = -i+1 | X_n = -i) = p.$$

Assim, se $(X_n)_{n \geq 0}$ é um PAS com probabilidade p de saltar para a direita, então $(X_n^*)_{n \geq 0}$ é um PAS que salta para a esquerda com a mesma probabilidade. Ou seja, se $p > \frac{1}{2}$, $(X_n)_{n \geq 0}$ tem um viés para a direita, enquanto $(X_n^*)_{n \geq 0}$ tem um para a esquerda, conforme ilustrado na figura 7 a seguir.

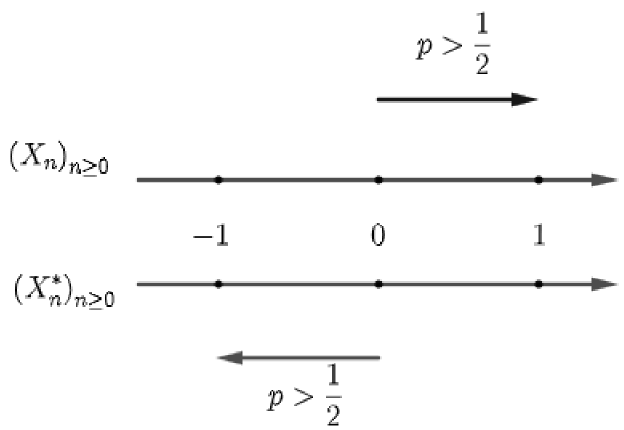


Figura 7: $(X_n)_{n \geq 0}$ com viés para a direita implicando em $(X_n^*)_{n \geq 0}$ com viés para a esquerda

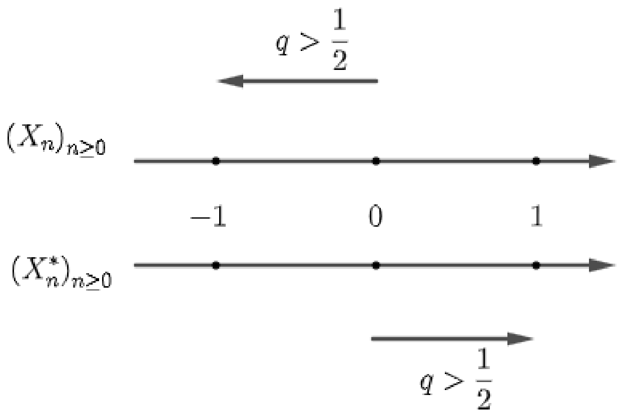


Figura 8: $(X_n)_{n\geq 0}$ com viés para a esquerda implicando em $(X_n^*)_{n\geq 0}$ com viés para a direita

Analogamente, com probabilidade q, $(X_n)_{n\geq 0}$ salta para a esquerda enquanto $(X_n^*)_{n\geq 0}$ para a direita. Assim, se $q > \frac{1}{2}$, o viés de $(X_n)_{n\geq 0}$ é para a esquerda e de $(X_n^*)_{n\geq 0}$ é para a direita (figura 8).

E para $X_0^* = 1$, temos o passeio aleatório simples representado na figura 9.

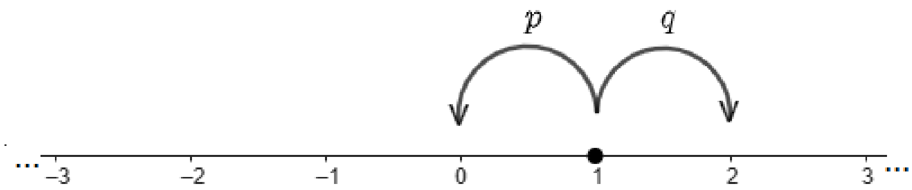


Figura 9: Passeio aleatório simples $(X_n^*)_{n\geq 0}$ com $X_0^* = 1$

Percebamos que podemos adaptar os resultados obtidos referentes às probabilidades da ruína do jogador ganancioso para deduzir as probabilidades do retorno à origem do passeio aleatório $(X_n^*)_{n\geq 0}$ iniciado nos inteiros positivos, apenas trocando p por q e vice-versa.

Lembrando, se X_n é o processo de ruína do jogador ganancioso com probabilidade p de ganhar 1 real, então

$$P(\text{ruína do ganancioso}|X_0 = m) = P(T_0 < \infty|X_0 = m) = \begin{cases} \left(\frac{q}{p}\right)^m, & \text{se } \ p > q \\ 1, & \text{se } \ p \leq q. \end{cases}$$

Assim, para $(X_n^*)_{n\geq 0}$ com $X_0^* = 1$, temos que

$$P(T_0 < \infty | X_0 = -1) = P(T_0^* < \infty | X_0^* = 1) = \begin{cases} \frac{p}{q}, & \text{se } q > p \\ 1, & \text{se } q \leq p \end{cases}$$

Lembrando também que, pela hipótese, temos $p \in (0,1) \Rightarrow q \in (0,1)$. Logo,

- por (3.14), se $p = q \Rightarrow p = \frac{1}{2}$, então

$$P(T_0 < \infty | X_0 = 0) = p + P(T_0 < \infty | X_0 = -1) \cdot q = p + 1q = p + q = 1$$

  e, portanto, $(X_n)_{n\geq 0}$ é recorrente.

- também por (3.14), se $p < q \Rightarrow p < \frac{1}{2}$, então

$$P(T_0 < \infty | X_0 = 0) = p + P(T_0 < \infty | X_0 = -1) \cdot q = p + \frac{p}{q} q = p + p = 2p < 1$$

  e, portanto, $(X_n)_{n\geq 0}$ é transiente.

- por (3.15), se $p > q \Rightarrow p > \frac{1}{2}$ e $q < \frac{1}{2}$, então

$$P(T_0 < \infty | X_0 = 0) = q + P(T_0 < \infty | X_0 = -1) \cdot q = q + 1q = 2q < 1$$

  e, portanto, mais uma vez $(X_n)_{n\geq 0}$ é transiente.

$\square$

Assim, concluímos que, se $p = q$, o passeio aleatório simples visitará a origem infinitas vezes. Para o caso em que $p \neq q$, segundo (3.13), a quantidade esperada de visitas à origem é $\frac{1}{1-2p}$ quando $p < q$, e $\frac{1}{1-2q}$ visitas para $p > q$.

3.2.2 *Tempo esperado para o retorno à origem*

A discussão na subseção anterior nos permite dizer sob quais probabilidades um passeio aleatório simples retorna à origem, e ainda nos permite ter a certeza das infinitas vezes da ocorrência desse retorno no caso dos passeios aleatórios simples simétricos.

Pois bem, agora estamos interessados em descobrir o tempo esperado para que esse retorno aconteça, correspondente à quantidade esperada de saltos. Portanto, queremos determinar

$$E(T_0 | X_0 = 0)$$

onde $T_0 = \min\{n \geq 1 : X_n = 0\}$.

E nessa busca, nos deparamos com o intrigante resultado a seguir.

Teorema 3.8. *Começando em 0, o tempo esperado para que o passeio aleatório simples retorne a 0 é infinito.*

Demonstração. Seja $S_m = E(T_0|X_0 = m)$. Condicionaremos o passeio ao resultado do primeiro salto para obter a seguinte expressão:

$$S_m = (1 + S_{m+1})p + (1 + S_{m-1})q$$

Em palavras, o tempo esperado da primeira visita ao sítio 0 após o instante 0 de um PAS iniciado no sítio m corresponde à soma dos tempos esperados da primeira visita ao sítio 0 após o instante 0 referentes a um PAS iniciando no sítio $m+1$ e a um PAS iniciando no sítio $m-1$, cada um deles adicionado a um instante de tempo, correspondente ao salto necessário para ir de m aos sítios vizinhos, multiplicados pela probabilidade de ocorrência desse salto.

Daí, segue que

$$S_m = (1 + S_{m+1})p + (1 + S_{m-1})q = p + q + pS_{m+1} + qS_{m-1} = 1 + pS_{m+1} + qS_{m-1}$$

Logo,

$$E(T_0|X_0 = 0) = S_0 = 1 + pS_1 + qS_{-1} \tag{3.16}$$

Agora, observemos que o PAS que começa nos inteiros positivos precisa passar pelo sítio 0 antes de chegar nos inteiros negativos. Então, a partir da discussão feita na subseção 3.1.4, podemos notar que o tempo esperado da primeira visita ao sítio 0 após o instante 0 de um PAS iniciado no sítio $m = 1$ corresponde à duração do jogo com um jogador ganancioso de fortuna inicial $m = 1$.

Sendo assim e pela Proposição 3.4,

$$S_1 = D_1^g = \begin{cases} \infty, & \text{se } \ p \geq q \\ \frac{1}{1-2p}, & \text{se } \ p < q \end{cases}$$

Usando ainda a definição $X_n^* = -X_n$ e os resultados obtidos para $(X_n^*)_{n\geq 0}$ na subseção anterior, temos que

$$S_{-1} = E(T_0|X_0 = -1) = E(T_0^*|X_0^* = 1)$$

Lembrando que $(X_n^*)_{n\geq 0}$ é um PAS que salta para direita com probabilidade q e salta para esquerda com probabilidade p, podemos adaptar os resultados obtidos para a duração do jogo com um jogador ganancioso com fortuna inicial $m = 1$, trocando p por q e vice-versa, para determinar o tempo esperado da primeira visita ao sítio 0 após o instante 0 do PAS iniciado em $X_0^* = 1$.

Sendo assim,

$$S_{-1} = E(T_0^*|X_0^* = 1) = \begin{cases} \infty, & \text{se } \ q \geq p \\ \frac{1}{1-2q}, & \text{se } \ q < p \end{cases}$$

Logo, voltando à equação (3.16), obtemos que,

- se $p = q = \frac{1}{2}$, então

$$S_0 = 1 + pS_1 + qS_{-1} = 1 + p \cdot \infty + q \cdot \infty = \infty$$

- se $p < q \Rightarrow p < \frac{1}{2} \Rightarrow 1 - 2p > 0$, então

$$S_0 = 1 + pS_1 + qS_{-1} = 1 + p \cdot \frac{1}{1-2p} + q \cdot \infty = \infty$$

- se $p > q \Rightarrow q < \frac{1}{2} \Rightarrow 1 - 2q > 0$, então

$$S_0 = 1 + pS_1 + qS_{-1} = 1 + p \cdot \infty + q \cdot \frac{1}{1-2q} = \infty$$

Portanto, $S_0 = E(T_0|X_0 = 0) = \infty$ em qualquer um dos casos acima. Isto é, o tempo esperado de retorno à origem de um passeio aleatório simples é infinito. E assim completamos a prova do Teorema 3.8 e o nosso estudo sobre os passeios aleatórios simples.

$\square$

4

PROCESSOS DE BIENAYMÉ-GALTON-WATSON

Neste capítulo, abordaremos um modelo matemático que gera resultados muito interessantes a cerca da probabilidade de extinção de uma população, o processo de Bienaymé-Galton-Watson (BGW). Esse processo foi introduzido em meados do século XIX de forma independente por Bienaymé, e por Galton e Watson durante o estudo sobre a sobrevivência do sobrenome de uma família quando esse é passado de pai para filho, assumindo que cada homem tem k filhos com probabilidade a_k, $k = 0, 1, 2, 3, 4, 5$ [17]. Desde então, esta teoria passou a ter aplicações importantes na modelagem e no estudo do crescimento de populações, entre outros campos.

Começaremos com uma definição detalhada do processo para, em seguida, discutirmos sobre a probabilidade desse processo extinguir-se ao longo do tempo.

4.1 DEFININDO O PROCESSO

De modo geral e com enfoque no crescimento populacional, podemos descrever um processo BGW da seguinte forma: vamos modelar o crescimento de uma população contando o total de indivíduos em cada geração, e considerando que cada indivíduo dá origem a novos indivíduos. Começa-se o processo com um conjunto inicial de indivíduos, os quais formam a geração 0. Cada indivíduo dessa geração gera novos indivíduos em quantidades aleatórias, independentes e identicamente distribuídas. Somando todos esses novos indivíduos gerados por cada indivíduo da geração 0, formaremos a primeira geração. Repetindo o mesmo procedimento para os indivíduos da primeira geração, encontramos o tamanho da segunda geração, e assim sucessivamente. Em geral, os indivíduos da geração n geram novos indivíduos da geração $n+1$, para $n \in \mathbb{N} \cup \{0\}$. Denotamos por Z_n à variável aleatória que representa a quantidade de indivíduos da n-ésima geração. Notemos que a quantidade de indivíduos da geração $n+1$ depende somente da quantidade de indivíduos da geração n, e portanto $(Z_n)_{n\geq 0}$ é uma cadeia de Markov.

Para encontrar uma representação mais formal do modelo descrito acima, consideremos que cada indivíduo dê origem a um total de novos indivíduos com a mesma distribuição de uma variável aleatória X, com distribuição de probabilidades $(p_k)_{k\geq 0}$ definida por:

$$P(X=k) = p_k, \quad \text{para} \quad k = 0, 1, 2, \ldots \tag{4.1}$$

Em outras palavras, um indivíduo gera k indivíduos com probabilidade p_k independente da geração que ele pertença e do total de filhos de outros membros da população. Reforçando o fato que a quantidade de descendentes diretos de cada indivíduo em todas as gerações é escolhida de forma independente e de acordo com uma única distribuição de probabilidades, tomemos $(X_{i,n})_{i\geq 1, n\geq 0}$ uma sequência de variáveis aleatórias independentes e identicamente distribuídas (i.i.d) com distribuição $(p_k)_{k\geq 0}$. A variável $X_{i,n}$ representará a quantidade de descendentes diretos do i-ésimo indivíduo da n-ésima geração. Logo, se $Z_n = j$ temos que

$$Z_{n+1} = \sum_{r=1}^{j} X_{i,n}. \tag{4.2}$$

Na figura 1 ilustramos uma possível realização deste processo.

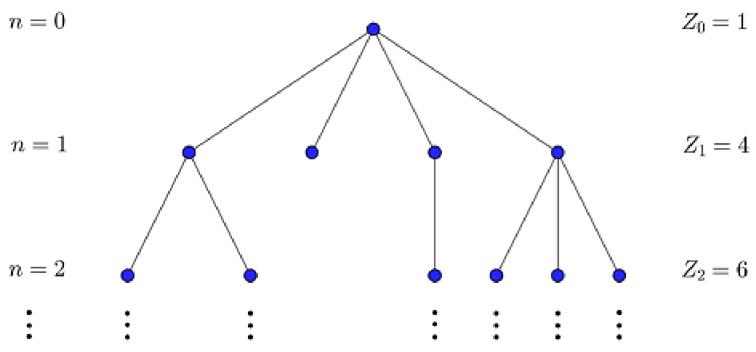


Figura 10: Exemplificando o processo de BGW

A partir do exposto acima, formalizaremos a definição do processo Bienaymé-Galton-Watson (BGW).

Definição 4.1. Chamamos processo de Bienaymé- Galton- Watson à cadeia de Markov $(Z_n)_{n\geq 0}$ com valores no conjunto $\{0,1,2,\ldots\}$ e probabilidades de transição dadas por:

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

onde $X_{1,n}, \ldots, X_{i,n}, i \geq 1, n \geq 0$ são variáveis i.i.d com mesma distribuição de uma variável X com $P(X = k) = p_k, k \geq 0$.

Os processos de BGW também são conhecidos por processos de ramificação. O motivo é facilmente percebido ao observarmos a figura 10.

Com o intuito de simplificar a exposição desse trabalho, consideraremos apenas os processos de BGW iniciados sempre por um único indivíduo, ou seja, $P(Z_0 = 1) = 1$.

Notemos que o estado 0 é um estado absorvente para $(Z_n)_{n>0}$. De fato, se não existir nenhum indivíduo na $n-$ésima geração, então não existirá indivíduos em nenhuma geração posterior, causando o que chamamos de extinção do processo. O ponto central do nosso estudo nesse capítulo é demonstrar sob quais condições ela pode ou não ocorrer. A próxima seção será dedicada inteiramente a isso.

Por ora, exporemos alguns resultados que necessitaremos acerca das variáveis aleatórias X e Z_n do processo de BGW.

Seja m o valor esperado para o número de descendentes diretos de cada indivíduo, ou seja,

$$m = E(X) = \sum_{k=0}^{\infty} k p_k.$$

Denote também por $f(s)$ a função geradora de probabilidade do total de filhos de um indivíduo. Ou seja

$$f(s) = E(s^X) = \sum_{k=0}^{\infty} p_k s^k, \quad |s| \leq 1$$

conforme discutido na seção 2.4.

Proposição 4.2. *Seja $E(X) = m$, então $E(Z_n|Z_0 = 1) = m^n$, $n \geq 0$*

Demonstração. A demonstração segue pelo princípio indução sobre n.

De fato, observe que

$$E(Z_1|Z_0 = 1) = E(X) = m = m^1.$$

Logo, a proposição é válida para $n = 1$

Suponha que a ela valha também para n. Ou seja, suponha como hipótese de indução que

$$E(Z_n|Z_0 = 1) = m^n.$$

Agora usaremos a propriedade de Markov para obter

$$\begin{aligned} E(Z_{n+1}|Z_0 = 1) &= \sum_{k\geq 1} E(Z_{n+1}|Z_n = k, Z_0 = 1)P(Z_n = k|Z_0 = 1) \\ &= \sum_{k\geq 1} E(Z_{n+1}|Z_n = k)P(Z_n = k|Z_0 = 1). \end{aligned}$$

Note que

$$E(Z_{n+1}|Z_n = k) = E\left(\sum_{r=1}^{k} X_{r,n}\right) = E(X_{1,n}) + E(X_{2,n}) + \cdots + E(X_{k,n}) = km$$

Sendo assim

$$E(Z_{n+1}|Z_0 = 1) = \sum_{k\geq 1} kmP(Z_n = k|Z_0 = 1) = mE(Z_n|Z_0 = 1)$$

E pela hipótese indução, temos que

$$E(Z_{n+1}|Z_0 = 1) = mm^n = m^{n+1}$$

Portanto, pelo princípio de indução finita, a proposição é verdadeira para todo $n \geq 0$, já que $E(Z_0|Z_0 = 1) = 1 = m^0$.

Em outras palavras, se o valor esperado para o número de descendentes diretos de cada indivíduo é m, então o valor esperado para a quantidade de indivíduos na n-ésima geração será m^n, dado que o processo tenha sido iniciado por um único indivíduo.

□

Proposição 4.3. *Seja $f_1 = f$ e $f_{n+1} = f \circ f_n$, para $n \geq 1$. Seja também f a função geradora de probabilidade da variável aleatória X. Então, a função geradora de probabilidade de Z_n condicionada a $Z_0 = 1$ é f_n.*

Demonstração. A demonstração segue pelo princípio indução sobre n.

Seja g_n a função geradora de probabilidade de Z_n dado que $Z_0 = 1$.

De fato, como $Z_1 = X_{1,1}$, temos

$$g_1(s) = E(s^{Z_1}|Z_0 = 1) = E(s^X) = f(s) = f_1(s)$$

Logo, a proposição é válida para $n = 1$.

Suponha que ela valha também para n. Ou seja, suponha que

$$g_n(s) := E(s^{Z_n}|Z_0 = 1) = \sum_{k=0}^{\infty} P(Z_n = k|Z_0 = 1)s^k = f_n(s)$$

Agora mostraremos que, se a proposição vale para n, então vale também para $n+1$.

Pois bem, da propriedade das cadeias de Markov, temos

$$\begin{aligned} g_{n+1} &= E(s^{Z_{n+1}}|Z_0 = 1) \\ &= \sum_{k=0}^{\infty} E(s^{Z_{n+1}}|Z_n = k, Z_0 = 1)P(Z_n = k|Z_0 = 1) \\ &= \sum_{k=0}^{\infty} E(s^{Z_{n+1}}|Z_n = k)P(Z_n = k|Z_0 = 1) \end{aligned}$$

Mas

$$E(s^{Z_{n+1}}|Z_n = k) = E(s^{\sum_{r=1}^{k} X_{r,n}}) = E(s^{X_{1,n}}s^{X_{2,n}} \dots s^{X_{k,n}}).$$

Como $X_{1,n}, \dots X_{i,n}$ são independentes e identicamente distribuídas com a mesma distribuição da variável aleatória X, então segue da equação acima que

$$E(s^{Z_{n+1}}|Z_n = k) = E(s^{X_{1,n}})E(s^{X_{2,n}}) \dots E(s^{X_{k,n}}) = (E(s^X))^k = (f(s))^k$$

Sendo assim,

$$g_{n+1} = \sum_{k=0}^{\infty} P(Z_n = k|Z_0 = 1)(f(s))^k = g_n(f(s)) = g_n \circ f$$

E pela hipótese de indução, temos que

$$g_{n+1} = f_n \circ f = f_{n+1}$$

Portanto, pelo princípio de indução finita, a proposição é verdadeira para todo $n \geq 1$.

Sendo assim,

$$E(s^{Z_n}|Z_0 = 1) = \sum_{k\geq 0} s^k \; P(Z_n = k|Z_0 = 1) = f_n(s).$$

Ou seja, a função geradora de probabilidade da variável aleatória Z_n, dado que $Z_0 = 1$, é a n-ésima interação da função de probabilidade f da variável aleatória X.

□

4.2 A PROBABILIDADE DE EXTINÇÃO DO PROCESSO

Como comentamos, o BGW modela o total de indivíduos em uma certa geração de uma população. Um dos fenômenos de interesse neste tipo de modelo é o da *extinção*. Ou seja, o momento em que aquela população deixa de existir.

No caso de modelos estocásticos, como o BGW, algumas perguntas importantes são: o que podemos dizer sobre a probabilidade de extinção do processo? Sob que condições a extinção é certa? Existe alguma condição para a qual a sobrevivência é certa?

Para responder estas perguntas, vamos antes estabelecer alguns fatos importantes.

Antes de mais nada, é importante notar que a extinção do processo de BGW acontece quando uma dada geração não tem mais filhos. De forma mais formal, se $(Z_n)_{n\geq 0}$ é um processo de BGW, a extinção do processo é representada pelo evento

$$\{Z_n = 0 \text{ para algum } n \geq 1\} = \bigcup_{n=1}^{\infty} \{Z_n = 0\}.$$

Como 0 é um estado absorvente de Z_n temos que se $Z_n = 0$, então $Z_{n+1} = 0$. Ou seja

$$\{Z_n = 0\} \subseteq \{Z_{n+1} = 0\} \quad \text{para todo } n \geq 1.$$

Isso mostra que a sequência de eventos $(\{Z_n = 0\})_{n\geq 1}$ é crescente e, pela Proposição 2.3, se q é a probabilidade de extinção do processo, então

$$\begin{aligned} q &= P(\text{extinção}|Z_0 = 1) \\ &= P(Z_n = 0 \;\text{ para algum }\; n \geq 1|Z_0 = 1) \\ &= P\left(\cup_{n\geq 1}\{Z_n = 0\}|Z_0 = 1\right) \\ &= \lim_{n\to\infty} P(Z_n = 0|Z_0 = 1). \end{aligned} \tag{4.3}$$

O evento complementar da extinção é chamado de *sobrevivência*, e ocorre se existir pelo menos um indivíduo em cada geração n, ou seja, $Z_n \geq 1$ para todo $n \geq 0$. Com isso a sobrevivência é representado pelo evento

$$\{Z_n \geq 1 \text{ para todo } n \geq 1\} = \bigcap_{n=1}^{\infty}\{Z_n \geq 1\}.$$

Como $\{Z_n \geq 1\}$ é o complementar de $\{Z_n = 0\}$, a sequência $(\{Z_n \geq 1\})_{n\geq 1}$ é decrescente e a Proposição 2.3 nos dá, mais uma vez, que a probabilidade de sobrevivência é dada por

$$\begin{aligned} 1 - q &= P(\text{sobrevivência}|Z_0 = 1) \\ &= P(Z_n \geq 1 \ \text{ para todo } \ n \geq 0|Z_0 = 1) \\ &= \lim_{n\to\infty} P(Z_n \geq 1|Z_0 = 1). \end{aligned} \tag{4.4}$$

Posto isso vamos tirar da frente alguns casos mais simples.

Primeiro vamos considerar o caso em que $p_0 = P(X = 0) = 1$. Neste caso o primeiro indivíduo não pode ter filhos e o processo se extingue imediatamente. Ou seja,

$$P(Z_1 = 0|Z_0 = 1) = 1.$$

Como $\{Z_1 = 0\} \subseteq \{Z_n = 0\}$ para todo $n \geq 1$, segue que

$$P(Z_n = 0|Z_0 = 1) = 1 \qquad \text{para todo } n \geq 1$$

e portanto $q = 1$.

Outro caso trivial ocorre quando $p_0 = P(X = 0) = 0$. Neste caso todo indivíduo tem ao menos um filho e a população nunca se extingue. Ou seja,

$$P(Z_1 \geq 1|Z_0 = 1) = 1$$

e

$$P(Z_{n+1} \geq 1|Z_n \geq 1) = 1 \qquad \text{para todo } n \geq 1,$$

de modo que $P(Z_n \geq 1|Z_0 = 1) = 1$ para todo $n \geq 1$ e

$$1 - q = P(\text{sobrevivência}|Z_0 = 1) = 1,$$

fazendo $q = 0$.

Outro caso interessante, mas não tão simples quanto os anteriores, acontece quando $p_0 + p_1 = P(X = 0) + P(X = 1) = 1$ com $p_1 = p > 0$ e $p_0 = 1 - p > 0$.

Neste caso cada indivíduo pode ter no máximo 1 filho. Com isso, como começamos com apenas um indivíduo, a população terá apenas um membro por geração, até que encontremos o primeiro indivíduo sem filhos.

Este processo é similar ao lançamento de sucessivas moedas viciadas, com probabilidade $1 - p$ de cara, onde seguimos com o lançamento até que observemos a primeira cara. Neste caso sabemos que eventualmente observaremos uma cara, finalizando o experimento.

No caso do processo BGW, isso nos diria que $q = 1$.

Para formalizar um pouco melhor a ideia acima, vamos adiantar um argumento que será utilizado no caso geral.

Primeiro lembre que a Proposição 4.3 nos diz que a função geradora de probabilidades de Z_n é $f_n = f \circ \cdots \circ f$, que é dada pela composta de n cópias de f, a geradora de probabilidades de X.

Com isso temos que

$$P(Z_n = 0 | Z_0 = 1) = f_n(0). \tag{4.5}$$

Mas como $p_0 = 1 - p$ e $p_1 = p$ temos

$$\begin{aligned}
f(s) &= (1-p) + ps \\
f_2(s) &= (1-p) + p[(1-p) + ps] = (1-p) + (1-p)p + p^2 s \\
f_3(s) &= (1-p) + p[(1-p) + (1-p)p + p^2 s] = (1-p) + (1-p)p + (1-p)p^2 + p^3 s \\
&\vdots \qquad\qquad\qquad \vdots \qquad\qquad\qquad\qquad\qquad \vdots \\
f_n(s) &= (1-p) + (1-p)p + (1-p)p^2 + \cdots + (1-p)p^{n-1} + p^n s
\end{aligned}$$

Um argumento por indução nos mostra então que

$$f_n(s) = (1-p) + (1-p)p + (1-p)p^2 + \cdots + (1-p)p^{n-1} + p^n s,$$

de onde segue que

$$f_n(0) = (1-p) + (1-p)p + (1-p)p^2 + \cdots + (1-p)p^{n-1}.$$

Notemos que $f_n(0)$ é dada pela soma dos n primeiros termos de uma progressão geométrica de razão p e iniciada pelo termo $(1-p)$. Logo,

$$f_n(0) = (1-p)\cdot\frac{1-p^n}{1-p} = 1-p^n$$

Segue assim que

$$q = \lim_{n\to\infty} P(Z_n = 0|Z_0 = 1) = \lim_{n\to\infty} f_n(0) = \lim_{n\to\infty}(1-p^n) = 1$$

E assim temos novamente a certeza da extinção do processo.

Agora podemos lançar a seguinte questão: além dos casos mencionados, quais condições levam à extinção do processo de BGW? E mais, é possível determinar a probabilidade q dela ocorrer?

E assim, chegamos ao principal resultado desse capítulo.

Teorema 4.4. *Seja $(Z_n)_{n\geq 0}$ um processo de BGW com $p_0 + p_1 < 1$ e $E(X) = m$.*

i. *Se $m \leq 1$ então $q = 1$;*

ii. *Se $m > 1$ então $q < 1$.*

Além disso, se $m > 1$ então q é a única solução em $[0,1)$ da equação $f(s) = s$.

Demonstração. Vamos dividir a demonstração em três partes.

1º parte: demonstraremos que se $m < 1$ então $q = 1$

De fato, como X é uma variável aleatória inteira não negativa, temos que

$$E(X) = \sum_{k\geq 0} kp_k \geq \sum_{k\geq 1} p_k = P(X\geq 1)$$

Como Z_n também é uma uma variável aleatória inteira não negativa, usaremos a desigualdade anterior para obter

$$E(Z_n|Z_0 = 1) \geq P(Z_n \geq 1|Z_0 = 1)$$

Usando a Proposição 4.2, temos

$$P(Z_n \geq 1|Z_0 = 1) \leq E(Z_n|Z_0 = 1) = m^n$$

Por hipótese, $0 \leq m < 1$. Logo, a equação (4.4) nos garante que

$$1 - q = \lim_{n\to\infty} P(Z_n \geq 1|Z_0 = 1) \leq \lim_{n\to\infty} m^n = 0$$

Ou seja, se o valor esperado m para a quantidade de descendentes diretos de cada indivíduo for menor que 1, a probabilidade de sobrevivência do processo BGW dada por $1 - q$ é nula. Assim temos $q = 1$ e a extinção do processo é certa.

Antes de continuar para os demais casos, vamos mostrar uma relação importante entre a probabilidade de extinção q e a função geradora f.

Como argumentamos anteriormente (ver equação (4.5)), temos que f_n é a função geradora de probabilidades da variável aleatória Z_n condicionada a $Z_0 = 1$, e assim podemos afirmar que

$$P(Z_n = 0|Z_0 = 1) = f_n(0).$$

Além disso, a equação (4.3) nos garante que

$$q = \lim_{n\to\infty} P(Z_n = 0|Z_0 = 1) = \lim_{n\to\infty} f_n(0).$$

Por outro lado, sabemos que $f_{n+1}(0) = f(f_n(0))$ e que f é continua em $[0, 1]$, e portanto

$$q = \lim_{n\to\infty} f_{n+1}(0) = \lim_{n\to\infty} f(f_n(0)) = f(q).$$

Portanto, q é solução da equação $f(s) = s$. Quando isso ocorre, dizemos que q é um ponto fixo da função f.

É importante salientar que a discussão acima garante que $\lim_{n\to\infty} f_n(0)$ sempre existe e está no intervalo $[0, 1]$, e é ponto fixo de f neste intervalo.

Com este resultado em mente, sigamos para os demais casos.

2º parte: demonstraremos que se $m = 1$ então $q = 1$.

Primeiramente observemos que

$$f'(s) = \sum_{k\geq 1} ks^{k-1}p_k < \sum_{k\geq 1} kp_k = m, \quad \text{para } s \in [0, 1)$$

Então, como $m = 1$, temos que

$$f'(s) < 1 \text{ para todo } s \in [0, 1) \tag{4.6}$$

Assim, pelo Teorema do Valor Médio (ver [22]), dado $0 \leq s < 1$, existe $c \in (s, 1)$ tal que

$$f(1) - f(s) = f'(c)(1 - s)$$

Como temos $c \in (s, 1)$ e $s \in [0, 1)$, então $c \in [0, 1)$ e, portanto, pela desigualdade (4.6), temos $f'(c) < 1$. Assim, segue que

$$f(1) - f(s) < 1 - s$$

Lembremos agora que, da discussão feita na Seção 2.4, temos $f(1) = 1$. Diante disso,

$$1 - f(s) < 1 - s \Rightarrow f(s) > s.$$

Portanto, se $m = 1$ temos $f(s) > s$ para todo $s \in [0, 1)$. Isto significa que a equação $f(q) = q$ não tem solução no intervalo $[0, 1)$. Porém, como $f(1) = 1$ temos que $s = 1$ é única solução de $f(s) = s$ em $[0, 1]$. Concluímos assim que $q = 1$.

Em outras palavras, se o valor esperado m para a quantidade de descendentes diretos de cada indivíduo for exatamente igual a 1, teremos mais uma vez a ocorrência da extinção do processo com probabilidade 1.

3º parte: demonstraremos que se $m > 1$ então $q < 1$.

Pois bem. Primeiro lembremos que, por hipótese, temos que $p_0 + p_1 < 1$. Assim, existe pelo menos um $k \geq 2$ tal que $p_k > 0$.

Logo,

$$f''(s) = \sum_{k \geq 2} k(k-1) p_k s^{k-2} > 0$$

no intervalo $(0, 1)$. Isso implica em f' é estritamente crescente em $(0, 1)$.

Notemos também que

$$\lim_{s \to 1^-} f'(s) = \lim_{s \to 1^-} \sum_{k \geq 1} k p_k s^{k-1} = \sum_{k \geq 1} k p_k = m$$

Logo, $\lim_{s \to 1^-} f'(s) > 1$, já que por hipótese temos $m > 1$.

Diante disso, temos que $f'(s) > 1$ para s próximo de 1. Em outras palavras, existe $\eta > 0$ tal que se $1 - \eta < s < 1$, então $1 < f'(s) < m$, já que temos f' crescente no intervalo $(0, 1)$.

Sendo assim, novamente pelo Teorema do Valor Médio, existe $c \in (s, 1)$ tal que

$$f(1) - f(s) = f'(c)(1 - s)$$

Como temos $c \in (s, 1)$ e $s \in (1 - \eta, 1)$, então $1 - \eta < c < 1$. Logo, $1 < f'(c) < m$.

Portanto, como $f(1) = 1$, vale que

$$1 - f(s) > 1 - s \Rightarrow f(s) < s \tag{4.7}$$

para $s \in (1 - \eta, 1)$.

Agora consideremos a função $g(x) = x - f(x)$. Notemos que g também é contínua em $[0, 1]$, já que g é dada pela diferença de duas funções contínuas.

Daí, segue da desigualdade em (4.7) que

$$g(s) > 0,$$

para $s \in (1 - \eta, 1)$.

Como $g(0) = 0 - f(0) = 0 - p_0 \leq 0$, segue do Teorema do Valor Intermediário (ver [22]), que existe $s_0 \in [0, 1 - \eta] \subset [0, 1)$ tal que $g(s_0) = 0$. Ou ainda, que $f(s_0) = s_0$.

Assim mostramos que existe $s_0 \in [0, 1)$ solução da equação $f(s) = s$.

Provaremos agora que s_0 é único. Para isso utilizaremos a concavidade de $f(s)$ ou, mais diretamente, a segunda derivada de f.

Por contradição vamos supor a existência de outra raiz s_1 para g em $[0, 1)$. Sem perda de generalidade, consideraremos $s_0 < s_1$.

Como temos $f(1) = 1$, então $g(1) = 1 - f(1) = 1 - 1 = 0$. Logo, $s_0 < s_1 < 1$ são três raízes de g em $[0, 1]$.

Assim, pelo Teorema de Rolle [22], existe $\xi_1 \in (s_0, s_1)$ e $\xi_2 \in (s_1, 1)$ tais que $g'(\xi_1) = g'(\xi_2) = 0$.

Logo, como $g'(x) = 1 - f'(x)$, temos

$$1 - f'(\xi_1) = 1 - f'(\xi_2) = 0,$$

o que nos dá que $f'(\xi_1) = f'(\xi_2) = 1$.

Mas isso é um absurdo, já que temos $\xi_1 < \xi_2$ e $f'(\xi_1) = 1 = f'(\xi_2)$ e f' é estritamente crescente no intervalo considerado.

Portanto, s_0 é a única raiz de g e a única solução da equação $f(s) = s$ no intervalo $[0, 1)$. Assim, temos somente duas soluções s_0 e 1 para a equação em $[0, 1]$.

Diante do fato que $q \in [0, 1]$ também satisfaz $f(s) = s$, então ou $q = s_0$ ou $q = 1$. Demonstraremos que $q = s_0$

De fato, por contradição, suponhamos que $q = 1$. Seguiria daí que

$$\lim_{n\to\infty} f_n(0) = q = 1.$$

Portanto, existiria n_0 suficientemente grande tal que $1 - \eta < f_n(0) < 1$ para todo $n \geq n_0$. Mas da desigualdade (4.7) obtemos que

$$f_{n+1}(0) = f(f_n(0)) < f_n(0).$$

Mas isso é um absurdo, já que $\{Z_n = 0\} \subset \{Z_{n+1} = 0\}$ e

$$f_{n+1}(0) = P(Z_{n+1} = 0|Z_0 = 1) \geq P(Z_n = 0|Z_0 = 1) = f_n(0).$$

Segue assim que $q = s_0 < 1$.

E portanto, a probabilidade do processo sobreviver para sempre é positiva somente se o valor esperado m para a quantidade de descendentes diretos de cada cada indivíduo for maior que 1, concluindo a demonstração. □

Enfim, em resposta aos questionamentos levantados no começo dessa seção, temos o Teorema 4.4.

Para facilitar a compreensão da demonstração do Teorema acima, podemos usar as informações contidas nela para construir o gráfico da função $f(s)$ e, a partir dele, analisar as raízes da equação $f(s) = s$, com o intuito de determinar a probabilidade de extinção q geometricamente.

Pois bem, na figura 11 temos três possíveis comportamentos da função f, dependendo de m. Para a análise dos gráficos, é importante lembrarmos que $f'(1) = \sum_{k\geq 1} kp_k = m$.

Como q é raiz da equação $f(s) = s$, então:

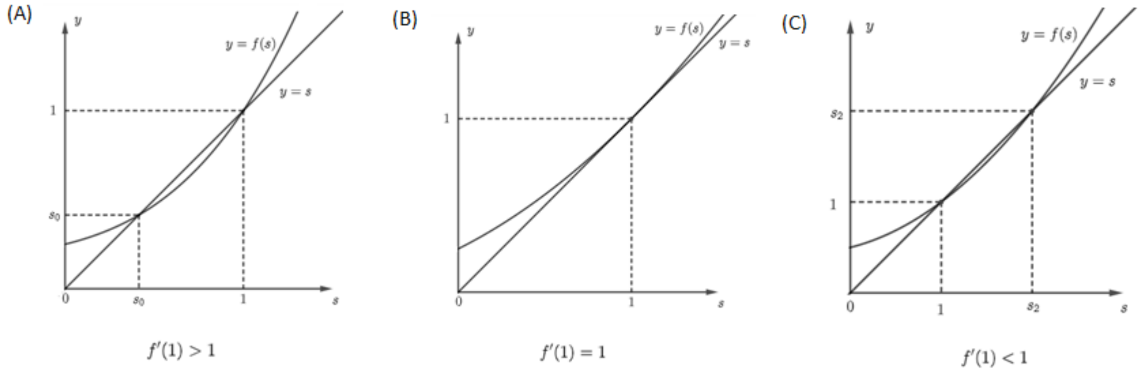


Figura 11: Comportamento da função $f(s)$ e as raízes da equação $f(s) = s$

- de (A) ou $q = s_0 < 1$ ou $q = 1$,
- de (B) $q = 1$,
- de (C) ou $q = 1$ ou $q = s_2 > 1$.

Concluímos de (B) e (C) que $q = 1$ para $f'(1) = m \leq 1$, já que não podemos ter $q > 1$. Para obtermos uma conclusão no caso (A), definiremos

$$q_n = P(Z_n = 0 | Z_0 = 1) = f_n(0), \quad \text{para todo} \quad n \geq 0$$

Dessa forma, como $\{Z_n = 0\} \subset \{Z_{n+1} = 0\}$, temos

$$0 = q_0 \leq q_1 \leq q_2 \leq \ldots \leq q_n \leq \ldots \leq 1$$

Logo a sequência $(q_n)_{n \geq 1}$ converge e

$$\lim_{n \to \infty} q_n = q.$$

Segue ainda que

$$q_n = f_n(0) = f(f_{n-1}(0)) \Rightarrow q_n = f(q_{n-1}),$$

para todo $n \geq 1$.

O que mostramos analiticamente na demonstração do Teorema é que, de fato,

$$q = \lim_{n \to \infty} q_n = s_0 < 1,$$

onde $s_0 \in [0, 1)$ é ponto fixo de f, como ilustrado na parte (A) da figura 11.

Tal convergência pode ser ilustrada na figura 12 abaixo.

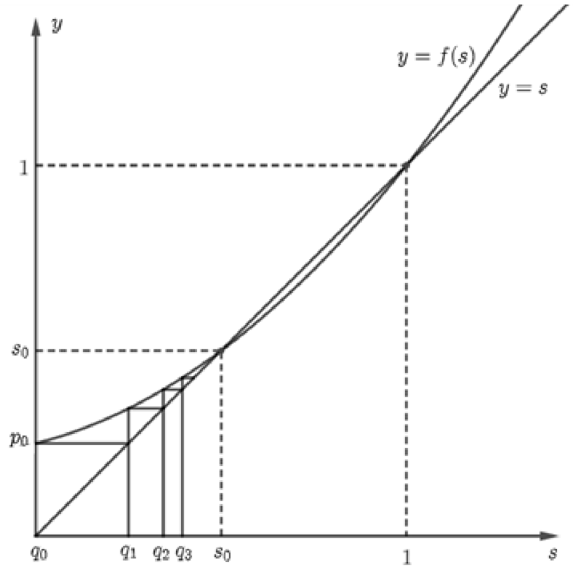


Figura 12: Convergência dos valores q_i à s_0

Exemplo 4.1. Considere um processo BGW iniciado por um indivíduo, com distribuição de descendentes diretos de cada indivíduo dado por $p_0 = \frac{1}{4}$, $p_1 = \frac{1}{2}$ e $p_2 = \frac{1}{4}$.

Calculando o valor esperado m para a quantidade de descendentes diretos por indivíduo, temos

$$m = 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} = 1$$

Pelo Teorema 4.4, temos $q = 1$ e, portanto, a extinção desse processo é certa. Em particular,

$$f(s) = \frac{1}{4} + \frac{1}{2}s + \frac{1}{4}s^2,$$

e fazendo $f(s) = s$, obtemos

$$s^2 - 2s + 1 = 0,$$

o que nos dá uma única raiz $s = 1$ de multiplicidade 2, corroborando com que já havíamos determinado, visto que q é raiz de $f(s) = s$.

Exemplo 4.2. Considerando a mesma situação do exemplo anterior, mas agora com $p_0 = \frac{1}{6}$, $p_1 = \frac{1}{6}$ e $p_2 = \frac{2}{3}$

Neste caso, temos

$$m = 0 \cdot \frac{1}{6} + 1 \cdot \frac{1}{6} + 2 \cdot \frac{2}{3} = \frac{3}{2} > 1$$

E novamente pelo Teorema 4.4, temos $q < 1$. Especificamente,

$$f(s) = \frac{1}{6} + \frac{1}{6}s + \frac{2}{3}s^2$$

e fazendo $f(s) = s$ obtemos

$$4s^2 - 5s + 1 = 0,$$

o que nos dá raízes $s = \frac{1}{4}$ e $s = 1$. Como nesse caso $q \in [0, 1)$, então temos $q = \frac{1}{4}$.

Portanto, a probabilidade do processo sobreviver para sempre é

$$P(\text{sobrevivência}|Z_0 = 1) = 1 - q = 1 - \frac{1}{4} = \frac{3}{4}$$

Exemplo 4.3. Para concluir, vejamos o caso onde o total de filhos de um indivíduo tem distribuição Geométrica de parâmetro $p \in (0, 1)$. Ou seja,

$$p_k = (1-p)^k p, \quad k \geq 0.$$

Neste caso temos

$$\begin{aligned}
m &= \sum_{k=1}^{\infty} k(1-p)^k p \\
&= \sum_{k=1}^{\infty} \sum_{i=1}^{k} (1-p)^k p \\
&= \sum_{i=1}^{\infty} \sum_{k=i}^{\infty} (1-p)^k p \\
&= \sum_{i=1}^{\infty} p(1-p)^i \sum_{k=i}^{\infty} (1-p)^{k-i} \\
&= \sum_{i=1}^{\infty} (1-p)^i \\
&= \frac{1}{1-(1-p)} - 1 \\
&= \frac{1-p}{p},
\end{aligned}$$

e portanto temos que a probabilidade de extinção é $q = 1$ para $p \geq 1/2$, e $q < 1$ quando $p < 1/2$.

Para encontrar a probabilidade de extinção, consideramos $p < 1/2 < 1 - p$ e voltamos para exemplo 2.8 para lembrar que a função geradora de probabilidade de uma distribuição geométrica é dada por

$$f(s) = \sum_{k=0}^{\infty} s^k (1-p)^k p = \frac{p}{1 - s(1-p)},$$

e resolvendo a equação $s = f(s)$ encontramos

$$\frac{p}{1 - s(1-p)} = s,$$

e portanto $s = 1$ e $s = \frac{p}{1-p} < 1$, mostrando que

$$q = \frac{p}{1-p}.$$

E assim terminamos o nosso estudo sobre os processos de ramificação de Bienaymé- Galton-Watson.

5

SEQUÊNCIA DIDÁTICA

Neste capítulo, apresentaremos uma proposta de sequência didática para o estudo da probabilidade no ensino básico utilizando conceitos abordados neste trabalho. Nosso intuito é contribuir com a incessante busca pelo aperfeiçoamento da prática docente e com a produção de atividades pedagógicas sobre esse tema matemático, visando atender às novas demandas de conhecimento nesta área, principalmente as que surgiram com a reforma curricular estabelecida no Brasil, desencadeada em 2018 pela publicação da BNCC, Base Nacional Comum Curricular [1]. Como discutido em [4], esse documento normativo para toda rede de ensino nacional ampliou o espaço curricular estocástico ao instituir Probabilidade e Estatística como uma das cinco unidades temáticas no ensino de Matemática e ao antecipar o trabalho de habilidades específicas dessa área na vida escolar do discente, ao mesmo tempo que superou obstáculos contidos nas diretrizes educacionais anteriores, como a insistência em conceitos pouco realistas da equiprobabilidade e a limitação à perspectiva clássica/laplaciana.

Com base na Proposta Curricular do Estado de São Paulo, o qual possui a maior rede de educação básica do Brasil (ver [2]), consideramos que esta sequência didática é mais indicada para o trabalho com alunos do 3º ano do ensino médio, devido ao fato dos conteúdos e habilidades necessárias para a realização das atividades serem contempladas apenas nessa fase do ensino básico paulista. Porém, isso dependerá da definição do currículo em cada unidade da federação, já que a BNCC permite que cada estado e município elaborem seus currículos considerando suas especificidades e contextos locais.

Antes de iniciarmos com a proposta, precisamos apresentar e discutir *O Problema de Monty Hall*, peça fundamental em nossa sequência didática.

5.1 O PROBLEMA DE MONTY HALL

Tornar as aulas de Matemática mais dinâmicas, visando estimular o interesse dos aluno pelos conteúdos a serem trabalhados, tem sido um grande e crescente desafio para os professores.

Diante disso, os jogos constituem uma importante e difundida ferramenta pedagógica, pois, além do aspecto lúdico bastante atrativo às crianças e aos adolescentes, propiciam aos alunos a oportunidade de aplicar ou validar conceitos matemáticos de forma prática e significativa, conforme ressalta Grando [7] "A busca por um ensino que considere o aluno como sujeito do processo, que seja significativo para o aluno, que lhe proporcione um ambiente favorável à imaginação, à criação, à reflexão, enfim, à construção e que lhe possibilite um prazer em aprender, não pelo utilitarismo, mas pela investigação, ação e participação coletiva de um "todo"que constitui uma sociedade crítica e atuante, leva-nos a propor a inserção do jogo no ambiente educacional, de forma a conferir a esse ensino espaços lúdicos de aprendizagem". E mais, “Inserido neste contexto de ensino-aprendizagem, o jogo assume um papel cujo objetivo transcende a simples ação lúdica do jogo pelo jogo, para se tornar um jogo pedagógico, com um fim na aprendizagem matemática – construção e/ou aplicação de conceitos” [6].

Os jogos ganham ainda mais destaque quando o assunto é Probabilidade, visto que sua origem como disciplina matemática se dá a partir das discussões realizadas na tentativa de quantificar as possibilidades de se ganhar em jogos de azar, concomitantemente com as tentativas de quantificar os riscos associados a sinistros (naufrágios, morte, etc.) em meados do século XV e XVI [24].

Nessas bases e fugindo dos tradicionais problemas de probabilidade envolvendo jogos com dados e baralhos, nossa proposta de sequência didática girará em torno do Jogo "O Prisioneiro de Monty Hall” cujo molde foi criado com base em um famoso problema matemático conhecido como "O Problema de Monty Hall”, que por sua vez é inspirado no jogo "Let’s Make a Deal” de um programa de televisão dos Estados Unidos exibido na década de 1970 e apresentado, obviamente, por Monty Hall [3].

Eis o problema:

> "Imagine que você está de frente para três portas numeradas: 1, 2 e 3, e o apresentador diz: – Atrás de uma dessas portas tem um carro; mas atrás de cada

> uma das outras duas tem um bode. Escolha uma porta e leve para casa o que estiver atrás dela.
>
> Você vai lá e escolhe uma das três portas; mas antes que você possa abri-la, o apresentador (que sabe exatamente onde está o carro) pede para você esperar e ele abre uma das portas não escolhidas, mostrando um dos bodes. Nesse momento ele faz a seguinte pergunta a você: – Você quer ficar com a porta que você escolheu ou quer trocá-la pela outra porta fechada?
>
> Qual é a estratégia mais lógica? Ficar com a porta escolhida inicialmente ou mudar de porta?"

Comumente a resposta dada para essa questão é "tanto faz trocar ou não trocar a porta, pois cada uma delas nesse momento detêm 50% de chance de estar com o carro atrás". Aliás, em um passado não tão distante, muitos matemáticos e físicos chegaram a garantir que o correto era não trocar a porta inicialmente escolhida [3].

Iremos provar que essa análise feita sobre o problema de Monty Hall está equivocada e que a melhor estratégia é fazer a troca das portas.

Pois bem, inicialmente consideremos os eventos:

- $E_1 = \{$ o carro está atrás da porta 1$\}$
- $E_2 = \{$ o carro está atrás da porta 2$\}$
- $E_3 = \{$ o carro está atrás da porta 3$\}$

Notemos que e os eventos E_1, E_2 e E_3 são dois a dois disjuntos e que, inicialmente, o carro pode estar atrás de qualquer uma das três portas, logo, as probabilidades $P(E_1), P(E_2)$ e $P(E_3)$ são iguais, ou seja,

$$P(E_1) = P(E_2) = P(E_3) = \frac{1}{3} \tag{5.1}$$

Agora seja o evento $A = \{\text{ganhar o carro}\}$. Dessa forma e pela Lei da probabilidade total, temos

$$P(A) = P(A|E_1) \cdot P(E_1) + P(A|E_2) \cdot P(E_2) + P(A|E_3) \cdot P(E_3)$$

Substituindo (5.1) na equação acima, obtemos

$$P(A) = \frac{P(A|E_1) + P(A|E_2) + P(A|E_3)}{3} \tag{5.2}$$

Sem perda de generalidade, suponhamos que a porta escolhida inicialmente seja a porta 1. Agora, para determinar as probabilidades envolvidas, devemos considerar dois casos.

Caso 1: a porta escolhida inicialmente não é trocada.

Como as hipótese são: "porta 1 escolhida inicialmente” e "porta escolhida inicialmente não é trocada", temos que

- Se o carro estiver atrás da porta 1, o apresentador abrirá a porta 2 ou 3 e levamos o carro para casa. Logo, $P(A|E_1) = 1$. Note que a escolha do apresentador não altera o fato de ganharmos o carro com probabilidade 1.
- Se o carro estiver atrás da porta 2, o apresentador abrirá a porta 3 e não haverá chance alguma de levarmos o carro para casa. Assim, temos $P(A|E_2) = 0$.
- Se o carro estiver atrás da porta 3, o apresentador abrirá a porta 2 e também não haverá nenhuma possibilidade de levarmos o carro. Logo, $P(A|E_3) = 0$.

Substituindo as probabilidades acima em (5.2), resulta

$$P(A) = \frac{P(A|E_1) + P(A|E_2) + P(A|E_3)}{3} = \frac{1+0+0}{3} = \frac{1}{3} \approx 0,33$$

Portanto, caso a porta escolhida inicialmente não for trocada, a probabilidade de ganharmos o carro é de aproximadamente 33%.

Caso 2: a porta escolhida inicialmente é trocada.

Como as hipótese são: "porta 1 escolhida inicialmente” e "porta escolhida inicialmente é trocada", temos que

- Se o carro estiver atrás da porta 1, o apresentador abrirá a porta 2 ou 3. Se a porta 2 for aberta, trocaremos a porta 1 pela porta 3 e perderemos o carro, inevitavelmente. Se o apresentador decidir abrir a porta 3, trocaremos a porta 1 pela porta 2 e obteremos o mesmo fim. Portanto, nessas condições é impossível levarmos o carro para casa, implicando em $P(A|E_1) = 0$.
- Se o carro estiver atrás da porta 2, o apresentador abrirá a porta 3, faremos a troca da porta 1 pela porta 2 e, infalivelmente, ganhamos o prêmio. Logo, $P(A|E_2) = 1$.
- Se o carro estiver atrás da porta 3, o apresentador abrirá a porta 2, ficaremos com a porta 3 ao invés da porta 1 e bingo! O carro é nosso. Isto é, $P(A|E_3) = 1$.

Substituindo as probabilidades acima em (5.2), segue que

$$P(A) = \frac{P(A|E_1) + P(A|E_2) + P(A|E_3)}{3} = \frac{0+1+1}{3} = \frac{2}{3} \approx 0,67$$

Concluímos, então, que se a troca das portas for realizada, dobramos as chances de levarmos o prêmio para casa, ou seja, a probabilidade de ganharmos o carro aumenta para aproximadamente 67%. Assim, mostramos que é o um equívoco apontar 50% como resultado dessa probabilidade ou, ainda pior, considerar não trocar a porta como melhor estratégia.

5.2 O PRISIONEIRO DE MONTY HALL

A união do Problema da Ruína do Jogador com o problema de Monty Hall, juntamente com a necessidade de ferramentas lúdicas de aprendizagem, deu origem ao jogo *O Prisioneiro de Monty Hall*, criado e desenvolvido em conjunto com o Professor Rafael Grisi.

Este é um jogo em formato digital, disponível em `https://www.geogebra.org/m/dfrv3uqd`, cujo objetivo é, basicamente, alcançar o nível mais alto da torre e se livrar completamente do risco de cair na prisão.

Detalhamos a seguir seu funcionamento:

- Ao iniciar o jogo, o jogador será apresentado à interface mostrada na figura 13.

Figura 13: O prisioneiro de Monty Hall: interface inicial

- O jogador começa no nível 1. A torre vai do nível 0, onde se encontra a prisão, até o nível 7, onde se encontra o prêmio, conforme aponta o indicador na figura 14

Figura 14: O prisioneiro de Monty Hall: indicador de nível

- As portas dão acesso aos níveis vizinhos. Cada porta esconde uma escada. Das três escadas escondidas, somente uma levará ao nível vizinho superior. As outras duas levarão ao nível vizinho inferior. As escadas são sempre dispostas de forma aleatória.
- Na tentativa de subir de nível, o jogador passará pela sequência descrita no Problema de Monty Hall: o jogador escolhe uma porta (figura 15); abre-se uma porta entre as duas não escolhidas que contém uma escada que leva ao nível inferior (figura 16); dá-se ao jogador a chance de trocar a porta inicialmente escolhida (figura 16); revela-se a escada atrás da porta finalmente escolhida (figura 17); o jogador entra na porta finalmente escolhida e irá para um dos níveis vizinhos (figura 17). Esses passos acontecem à medida que clicamos nos "botões" na tela (figura 18).

Figura 15: O prisioneiro de Monty Hall: o jogador escolhe uma porta.

Figura 16: O prisioneiro de Monty Hall: abre-se uma porta entre as duas não escolhidas.

Figura 17: O prisioneiro de Monty Hall: porta escolhida revelada.

Figura 18: O prisioneiro de Monty Hall: botões de comando.

Figura 19: O prisioneiro de Monty Hall: jogo perdido e encerrado.

- Nos demais níveis essa sequência se repete, exceto no nível 0, onde o jogador é aprisionado e perde jogo (figura 19), e no nível 7, onde o jogador escapa definitivamente da prisão e vence o jogo (figura 20).

Figura 20: O prisioneiro de Monty Hall: jogo vencido e encerrado.

- Encerrado o jogo, o jogador pode recomeçar o jogo clicando em "jogar novamente" (figuras 19 e 20).

Abrimos essa seção dizendo que o Prisioneiro de Monty Hall é uma união dos problemas da Ruína do Jogador com o de Monty Hall. Pois bem, a relação com o Problema de Monty Hall é evidente, mas e com o Problema da Ruína do Jogador? É evidente? Caso não seja, notemos que podemos associar o Prisioneiro de Monty Hall com o Problema da Ruína (seção 3.1) no qual o jogador começa com uma fortuna inicial de $X_0 = m = 1$ (o prisioneiro começa no nível 1), e jogará até ruir quando sua fortuna atingir o valor $X_n = 0$ (a prisão no nível 0) ou até que sua fortuna atinja o valor $X_n = N = 7$ (o prêmio no nível 7).

Sendo assim, imaginemos que o prisioneiro de Monty Hall adote a estratégia de nunca trocar a porta inicialmente escolhida. Logo, a probabilidade dele ir para o nível vizinho superior é $p = \frac{1}{3}$ e dele ir para o nível vizinho inferior é $q = \frac{2}{3}$, em consonância com a demonstração feita na seção 5.1 . Nesse cenário e pela Proposição 3.1, a probabilidade dele ir para a prisão é dada por

$$P(\text{prisão}|X_0 = 1) = P(\text{ruína}|X_0 = 1) = \frac{\left(\frac{q}{p}\right)^m - \left(\frac{q}{p}\right)^N}{1 - \left(\frac{q}{p}\right)^N} = \frac{(2)^1 - (2)^7}{1 - (2)^7} \approx 0,99,$$

e pela Proposição 3.3, a duração esperada do jogo é dada por

$$D_1 = \frac{m}{1-2p} - \left(\frac{N}{1-2p}\right)\left(\frac{1-\left(\frac{q}{p}\right)^m}{1-\left(\frac{q}{p}\right)^N}\right) = \frac{1}{1-\frac{2}{3}} - \left(\frac{7}{1-\frac{2}{3}}\right)\left(\frac{1-(2)^1}{1-(2)^7}\right) \approx 2,8.$$

Ou seja, o prisioneiro de Monty Hall, que adotar a estratégia de nunca trocar a porta escolhida no primeiro momento, vai para a prisão com probabilidade de 99% em 2,8 lances de escada em média.

Agora, se estratégia adotada pelo prisioneiro for a de sempre trocar a porta escolhida inicialmente, ele vai para o nível vizinho superior com probabilidade $p = \frac{2}{3}$, e para o nível vizinho inferior com probabilidade $q = \frac{1}{3}$, também conforme discutido na seção 5.1. Assim, a probabilidade dele ir para a prisão passa a ser

$$P(\text{prisão}|X_0 = 1) = P(\text{ruína}|X_0 = 1) = \frac{\left(\frac{q}{p}\right)^m - \left(\frac{q}{p}\right)^N}{1 - \left(\frac{q}{p}\right)^N} = \frac{\left(\frac{1}{2}\right)^1 - \left(\frac{1}{2}\right)^7}{1 - \left(\frac{1}{2}\right)^7} \approx 0,5,$$

e a duração esperada do jogo passa a ser

$$D_1 = \frac{m}{1-2p} - \left(\frac{N}{1-2p}\right)\left(\frac{1-\left(\frac{q}{p}\right)^m}{1-\left(\frac{q}{p}\right)^N}\right) = \frac{1}{1-\frac{4}{3}} - \left(\frac{7}{1-\frac{4}{3}}\right)\left(\frac{1-\left(\frac{1}{2}\right)^1}{1-\left(\frac{1}{2}\right)^7}\right) \approx 7,6.$$

Isto significa que, se o prisioneiro de Monty Hall sempre trocar a porta escolhida no primeiro momento, então a probabilidade dele ir para a prisão é reduzida para 50% e a duração média do jogo aumenta para aproximadamente 8 lances escadas. Uma condição bem mais favorável ao jogador como podemos notar.

5.3 PROPOSTA DE SEQUÊNCIA DIDÁTICA

Após a descrição do jogo O Prisioneiro de Monty Hall, avançaremos com o planejamento da nossa proposta de sequência didática que terá este jogo como elemento fundamental.

O plano abaixo pressupõe 6 encontros, em nível crescente de dificuldade.

O primeiro é apenas para discutir os conceitos matemáticos que usaremos, como probabilidade condicional. Se o público já conhece estes conceitos, ou se o professor julgar desnecessário este passo, pode iniciar na segunda aula.

Os últimos 2 encontros serão usados para explorar o problema da ruína do jogador, e calcular as probabilidades envolvidas no jogo do Prisioneiro de Monty Hall. Sugerimos só desenvolver estas partes da atividade com um grupo de alunos realmente interessados no problema e com alguma afinidade matemática.

5.3.1 *Plano da Proposta.*

- **Tema**: Probabilidade.
- **Público-alvo**: alunos do 3ª ano do ensino médio.
- **Objetivos**: Refletir sobre a importância da probabilidade na tomada de decisões; aplicar conceitos de probabilidade na resolução de problemas.
- **Conhecimentos prévios**: Como trata-se de uma sequência didática que visa a aplicação de conceitos, espera-se que os alunos já tenham estudado os conteúdos relativos ao cálculo de probabilidades, probabilidade condicional e à lei da probabilidade total, para que a interação com o experimento seja bastante significativa. Será também preciso utilizar em uma prova matemática os conceitos relativos às Progressões geométricas e à soma de seus termos.
- **Objetos de conhecimento**: Probabilidade simples e condicional; Lei da probabilidade total.
- **Habilidades da BNCC a serem desenvolvidas**:

  (EM13MAT106) Identificar situações da vida cotidiana nas quais seja necessário fazer escolhas levando-se em conta os riscos probabilísticos.

  (EM13MAT312) Resolver e elaborar problemas que envolvem o cálculo de probabilidade de eventos em experimentos aleatórios sucessivos.
- **Tempo de execução**: 6 aulas de 45 minutos.
- **Recursos didáticos**:

  ✓Caderno, caneta, lápis e borracha para os alunos realizarem registros e resoluções.

  ✓Lousa e giz (ou material similar) para as explanações do professor.

  ✓Laboratório de informática.

  ✓Jogo "O Prisioneiro de Monty Hall”. Disponível: `https://www.geogebra.org/m/dfrv3uqd`

  ✓kit multimídia para projeção de apresentação em slides.

- **Desenvolvimento**:

  1ª *aula*.

  Como ponto de partida, propomos que a primeira aula seja destinada a revisão dos conhecimentos prévios descritos acima, os quais são indispensáveis para a execução dessa sequência didática. Contudo, cabe ao professor avaliar quais pontos exatamente precisam ser revisados e/ou aprofundados considerando particularmente os conhecimentos de sua turma, assim como adequar esse planejamento com a quantidade de aulas necessárias para esse momento.

  2ª *aula*.

  1º momento: o professor apresentará o famoso Problema de Monty Hall. Com o intuito de captar um maior interesse por parte dos alunos e dar ao problema um destaque que corrobora com seu papel central na sequência didática, sugerimos uma apresentação em slides usando um datashow para que possamos realçá-la com ilustrações e animações. Elaboramos uma sequência de slides como sugestão, mas nada impede que o professor use sua criatividade e/ou outras formas para aprimorar esse momento. Os slides mencionados encontram-se disponíveis nos Apêndices A desta dissertação.

  2º momento: os alunos devem ser divididos em grupos para que discutam entre si sobre a questão levantada pelo problema de Monty Hall: *qual é a melhor estratégia? Ficar com a porta escolhida inicialmente, mudar de porta ou tanto faz?*. Peça aos alunos que formulem uma conclusão justificada.

  3º momento: agora cada grupo deve socializar com o restante da turma e com o professor a conclusão a qual chegaram no 2º momento com sua respectiva justificativa.

  Sugerimos que os grupos sejam formados por 4 alunos no máximo.

  Observação:

  – A expectativa é que a conclusão "tanto faz" seja a mais citada entre os grupos. De qualquer forma, o professor não deve dizer ainda que a melhor estratégia é a de trocar a porta escolhida no primeiro momento, pois as etapas da sequência didática farão com que os alunos cheguem a ela empiricamente.

  3ª *aula*.

1º momento: no laboratório de informática, o professor apresentará o jogo "O Prisioneiro de Monty Hall” e uma breve explicação sobre seu funcionamento. Espera-se que os alunos notem sua relação com o problema de Monty Hall. Caso isso não aconteça, o professor deverá explanar sobre as semelhanças e diferenças do jogo com o problema discutido na aula anterior.

2º momento: chegamos ao momento em os alunos jogarão O Prisioneiro de Monty Hall. Cada aluno joga uma única vez e todos devem jogar seguindo a estratégia de **nunca** trocar a porta inicialmente escolhida em todos os níveis do jogo. O professor deve enfatizar que todos devem seguir à risca esse comando para não comprometer a comparação de resultados que iremos fazer futuramente. O ideal é que cada aluno assuma um computador. Caso isso não seja possível e o professor tenha que formar grupos de alunos em cada computador, é importante que cada aluno, um por vez, faça a sua jogada. Após todos os alunos terem feito sua jogada, o professor deve fazer o registro da frequência relativa da quantidade de derrotas entre os alunos.

Lembramos o professor que deve-se esperar que 99% dos alunos tenham ido parar na prisão, conforme discutido na seção 5.2.

3º momento: o professor repetirá o 2º momento desta aula, porém, com uma única diferença: agora os alunos devem jogar seguindo a estratégia de **sempre** trocar a porta inicialmente escolhida em cada nível do jogo.

Novo lembrete ao professor: agora, a expectativa é que a quantidade de alunos que foram para na prisão tenha sido reduzido para 50%, também conforme discutido na seção 5.2.

Observações:

– A quantidade de alunos na sala de aula é importante para os resultados obtidos nesta aula. Quanto mais alunos, melhor. Caso a sala tenha uma quantidade pequena de alunos, o professor deverá fazer adaptações como: mais jogadas por aluno, junção de salas, etc.

– Agora caso o professor disponha de uma quantidade insuficiente de computadores e tenha que agrupar os alunos nos equipamento, fazendo com que o 2º momento desta aula se estenda, será necessário adiar o 3º momento desta para a próxima aula. Assim, aumentando a quantidade total de aulas desta sequência didática para 7. Outra possibilidade é o uso dos aparelhos de celular disponibilizados pelos próprios alunos.

4ª *aula*.

1º momento: De volta à sala de aula, o professor expõe os resultados obtidos na aula anterior (percentual de derrotas em ambos os casos). Na sequência repete a pergunta feita no 2º momento da 2ª aula desta sequência: *qual é a melhor estratégia? Ficar com a porta escolhida inicialmente, mudar de porta ou tanto faz?*.

Com base nos resultados, todos os alunos devem conjecturar que trocar a porta é a melhor estratégia.

Caso isso não aconteça, abra tempo para discussão. Deixe os alunos discutirem e tentarem convencer os céticos do resultado descoberto.

2º momento: agora é o momento do professor transformar a conjectura em fato. Para isso, ele deve começar fazendo a demonstração da solução do Problema de Monty Hall, como exibimos na seção 5.1.

3º momento: o professor pedirá aos alunos que façam uma pesquisa extraclasse que deve ser entregue na próxima aula sobre "O Problema da Ruína do Jogador".

Observação:

– O intuito da pesquisa é fazer com que os alunos estejam familiarizados com o tema para facilitar a compreensão da demonstração que será realizada na próxima aula.

5ª *aula*.

1º momento: o professor recolhe as pesquisas e, em seguida, apresenta oficialmente aos alunos o Problema da Ruína do Jogador (seção 3.1).

2º momento: o professor deverá demonstrar, conforme discutido neste trabalho, o caso 1 da Proposição 3.1, na qual estabelece-se a probabilidade da ruína do jogador em que $p \neq q$. Neste ponto, o argumento que leva à equação de recorrência pode ser pouco rígido, se valendo apenas de uma descrição informal das possibilidades após o primeiro passo.

3º momento: o professor solicitará novamente aos alunos uma atividade extraclasse que deve ser entregue na próxima aula. A atividade consistirá na demonstração do caso 2 da Proposição 3.1, na qual estabelece-se a probabilidade da ruína do jogador em que $p = q$. O professor pode ressaltar que a demonstração solicitada segue a mesma estrutura da demonstração para caso 1 feita no 2º momento desta aula, necessitando apenas de alguns ajustes, e ainda, expor mais algumas dicas caso julgue necessário.

6ª *aula*.

1º momento: o professor estabelecerá a conexão entre o jogo O Prisioneiro de Monty Hall e o Problema da ruína do jogador, conforme discutimos na seção 5.2.

2º momento: os alunos deverão calcular a probabilidade de prisão no jogo O Prisioneiro de Monty Hall aplicando a Proposição 3.1 demonstrada na aula anterior, conforme também discutimos na seção 5.2. Primeiramente, os alunos devem efetuar esse cálculo considerando um jogador que usa a estratégia de nunca realizar a troca da porta escolhida inicialmente em cada nível do jogo e, em seguida, calcular novamente essa mesma probabilidade considerando dessa vez um jogador que sempre realiza a troca.

3º momento: o professor deverá fazer a correção na lousa dos cálculos feitos pelos alunos no momento acima e compará-los com resultados obtidos empiricamente na 3ª aula através do jogo. Isto é, comparar a probabilidade de prisão calculada com o percentual de prisões no decorrer do jogo. O intuito dessa etapa é validar resultados experimentais.

- **Avaliação**: sugerimos que a avaliação seja composta por vários critérios: participação e comprometimento em todas as etapas da sequência; apresentação realizada no 3º momento da 2ª aula; as atividades extraclasses solicitadas na 4ª e 5ª aula.

Assim, fechamos a proposta de sequência didática para o ensino de matemática dentro da mais nova unidade temática instituída para o currículo do ensino básico no País: Probabilidade e Estatística. Ressaltamos ainda que, na elaboração dessa proposta, buscamos a associação entre a investigação e a validação matemática de resultados obtidos através da experiência auxiliada pelo emprego de recursos tecnológicos, em concordância com a BNCC, que determina como uma das competências na área de matemática e suas tecnologias "Investigar e estabelecer conjecturas a respeito de diferentes conceitos e propriedades matemáticas, empregando estratégias e recursos, como observação de padrões, experimentações e diferentes tecnologias, identificando a necessidade, ou não, de uma demonstração cada vez mais formal na validação das referidas conjecturas" [1].

6

CONSIDERAÇÕES FINAIS

Este trabalho teve como objetivo geral estudar os processos de ramificação de Bienaymé-Galton-Watson e os passeios aleatórios simples para que pudéssemos determinar as condições, dentro desses modelos, que possam indicar se determinada população está ou não a caminho de sua extinção. Com base nos resultados encontrados no desenvolvimento da pesquisa, pode-se indicar que o objetivo proposto foi alcançado.

Dentre os principais resultados, observou-se que, ao modelar o desenvolvimento populacional a partir dos processos de ramificação de Bienaymé-Galton-Watson, a extinção está relacionada à esperança para a quantidade de descendentes diretos de cada indivíduo pertencente ao processo. Se essa esperança for menor ou igual a 1, a extinção é um evento certo. Caso contrário, a população terá uma probabilidade positiva de sobrevivência e a probabilidade de extinção será dada pela menor raiz não negativa da equação $f(s) = s$, em que $f(s)$ é a função geradora de probabilidades da variável aleatória correspondente ao número de descendentes diretos de cada indivíduo do processo.

No que tange ao modelo estabelecido a partir dos passeios aleatórios simples, constatou-se que a extinção de uma determinada população é inevitável quando a probabilidade de um indivíduo morrer for maior ou igual a probabilidade de um indivíduo nascer dentro da população em análise. Além disso, foi obtida a duração esperada, definida pelo total entre nascimentos e mortes, para que essa extinção ocorra.

Outro objetivo geral pretendido com este trabalho estava na intenção de fomentar o estudo desses modelos entre os estudantes da rede nacional de educação. Na tentativa de atingir futuramente esse objetivo, foi apresentado uma proposta para introduzir conceitos desse estudo já no ensino médio das escolas de ensino básico.

Dessa forma, este trabalho pode contribuir como base para a formação de pesquisadores, que poderão por sua vez contribuir na formulação de políticas públicas efetivas para a redução do risco de extinção de populações.

Acerca das limitações presentes nesta pesquisa, pode-se destacar a falta da aplicação da proposta apresentada de sequência didática com alunos do 3º ano do ensino médio e, consequentemente, a ausência dos resultados relevantes relacionados à aprendizagem dessa aplicação. Essa limitação não pôde ser superada devido aos prazos estabelecidos para o cumprimento desta dissertação. Outra limitação constatada foi a pouca variedade que encontramos de livros publicados sobre os processos aqui estudados.

Posto isso, em relação às futuras investigações, recomenda-se a aplicação e a avaliação da sequência didática sugerida, como também aperfeiçoá-la a partir dos resultados obtidos. Além disso, recomenda-se aprofundar o conhecimento sobre os passeios aleatórios simples através das cadeias de nascimento e morte, e explorar a aplicabilidade dos processos de ramificação em estudos no campo da Biologia.

A

SLIDES: O PROBLEMA DE MONTY HALL

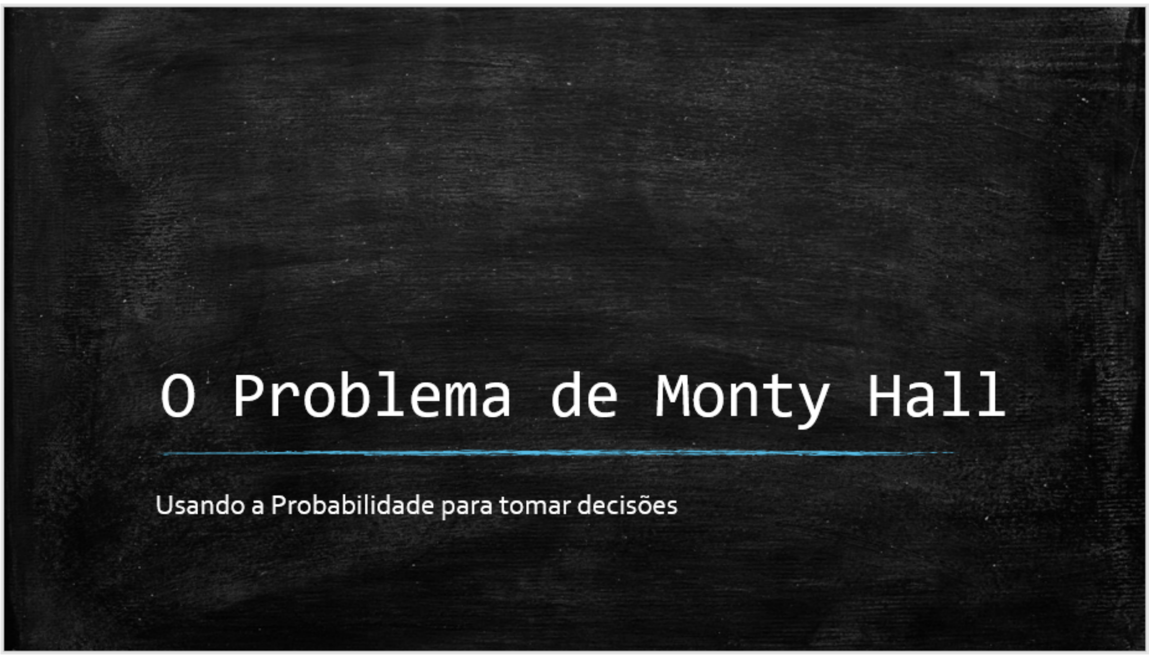


Figura 21: Slide 1

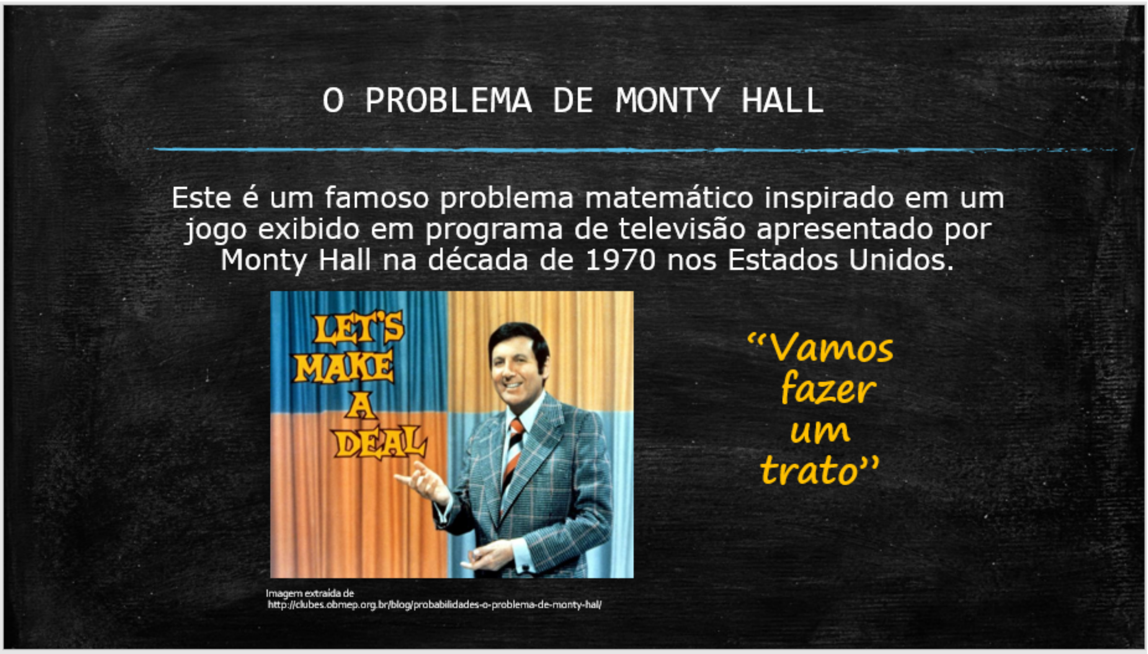


Figura 22: Slide 2

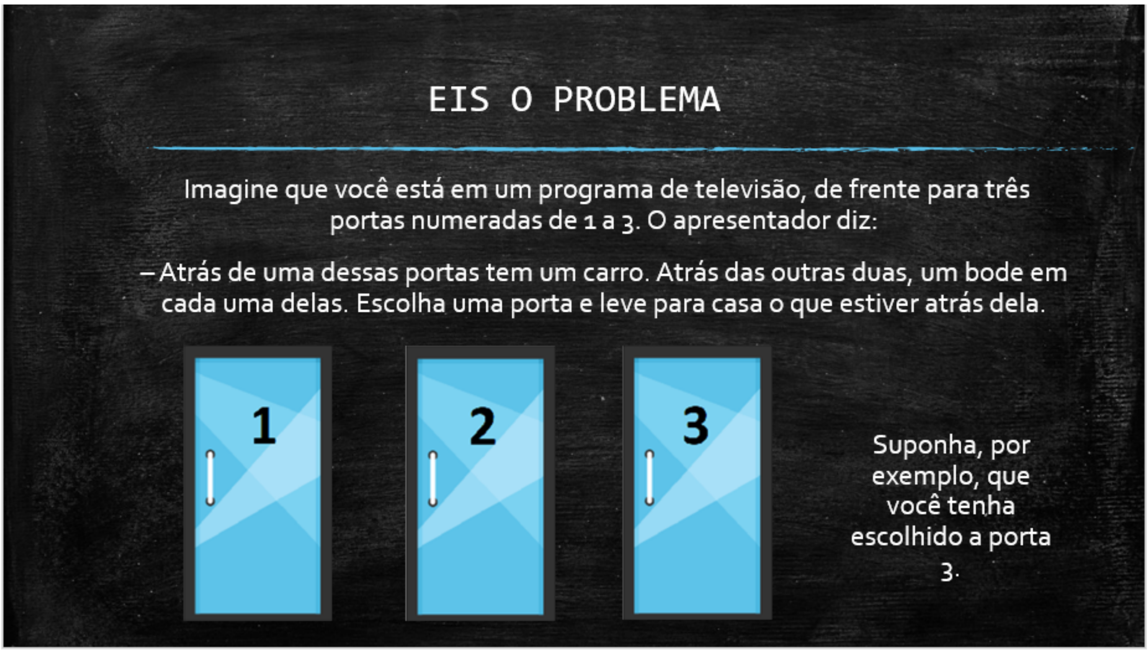


Figura 23: Slide 3

Figura 24: Slide 4

Figura 25: Slide 5

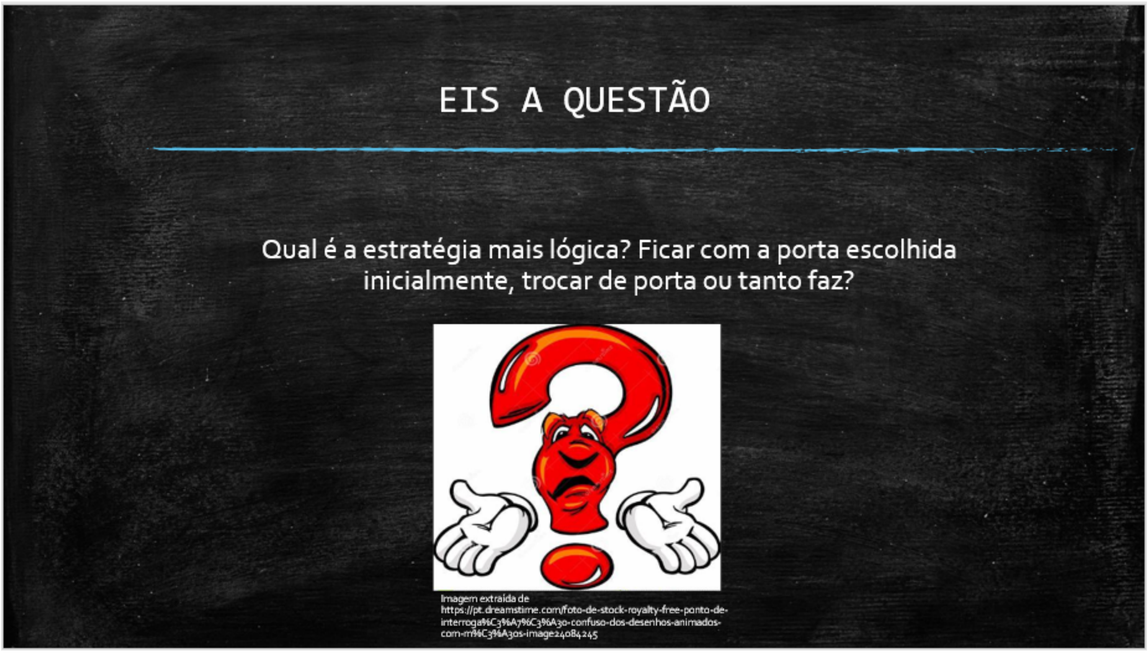


Figura 26: Slide 6

BIBLIOGRAFIA

[1] Brasil, *Ministério da Educação. Base Nacional Comum Curricular*, (2018).

[2] Secretaria de educação do Estado de São Paulo, *A Secretaria*, https://www.educacao.sp.gov.br/institucional/a-secretaria/, acesso em 21-08-2023.

[3] Clube de Matemática da OBMEP, *Probabilidades: o problema de Monty Hall*, http://clubes.obmep.org.br/blog/probabilidades-o-problema-de-monty-hal/, acesso em 16-08-2023.

[4] Samya de Oliveira Lima et al., *Ensino de Estatística, Probabilidade e Combinatória na Educação Básica: os novos desafios da BNCC*, Revista Baiana de Educação Matemática (2022).

[5] Fernando Duarte, *Pela 1ª vez, mundo tem 'mais avós do que netos'*, BBC News Brasil (2019).

[6] Regina Celia Grando, *O jogo [e] suas possibilidades metodológicas no processo ensino-aprendizagem da matemática*, Tese de Doutoramento, Universidade Estadual de Campinas, 1995.

[7] Regina Célia Grando et al., *O conhecimento matemático e o uso de jogos na sala de aula*, (2000), 239.

[8] Samuel Hazzan, *Fundamentos de Matemática Elementar: combinatória, probabilidade*, 7ª ed., vol. 5, Atual, São Paulo, 2004.

[9] Valdivino V Junior, Tiago M Vargas e Divaldo Portilho F Junior, *Passeios Aleatórios Simples: Uma Aplicaçao em Jogos*.

[10] Fátima Aparecida Kian et al., *O ensino de Probabilidade e o novo Ensino Médio: reflexões a partir da BNCC e do Currículo Paulista*, RIPEM (2021).

[11] Augusto César Morgado et al., *Análise Combinatória e Probabilidade*, 9ª ed., SBM, Rio de Janeiro, 2006.

[12] ______, *Matemática Discreta*, 2ª ed., SBM, Rio de Janeiro, 2015.

[13] Mariko Oi, *As medidas bilionárias de países asiáticos para reverter queda na natalidade*, BBC News Brasil (2023).

[14] João Perassolo, *Países têm bebês em falta e dão dinheiro para aumentar população em mundo de 8 bilhões*, Folha de S.Paulo (2022).

[15] João Ismael D. Pinheiro et al., *Probabilidade e Estatística: quantificando a incerteza*, Elsevier, Rio de Janeiro, 2012.

[16] Laura Rifo, *Probabilidade e Estatística: aspectos de tomada de decisões e incerteza para o ensino fundamental e médio*, 1ª ed., SBM, Rio de Janeiro, 2021.

[17] Pablo Martín Rodríguez, *Processos de ramificação: teoria e aplicações*, II Colóquio de Matemática da Região Sul, Universidade Estadual de Londrina, 2012.

[18] Sheldon Ross, *Probabilidade: um curso moderno com aplicações*, 8ª ed., Bookman, Porto Alegre, 2010.

[19] Walter Rudin, *Principles of Mathematical Analysis*, 3ª ed., McGraw-Hill, New York, 1976.

[20] Rinaldo B. Schinazi, *Classical and Spatial Stochastic Processes: with applications to biology*, 2ª ed., Birkhauser, New York, 2014.

[21] James Stewart, *Cálculo*, 5ª ed., vol. 2, Pioneira Thomson Learning, São Paulo, 2006.

[22] ______, *Cálculo*, 6ª ed., vol. 1, Cengage Learning, São Paulo, 2011.

[23] Joan Jesus Amaya Triana et al., *Processos de ramificação e aplicações em modelos de transmissão de informação*, (2018).

[24] Lorí Viali, *Algumas considerações sobre a origem da teoria da probabilidade*, Revista Brasileira de História da Matemática **8** (2008), nº 16, 143–153.