

Propriedades Estatísticas de um Processo Multiplicativo em um Espaço Amostral Discreto

Amanda Alves de Oliveira¹

Telles Timóteo da Silva²

Resumo:

O objetivo desse trabalho é apresentar alguns resultados necessários para a compreensão do comportamento assintótico dos processos multiplicativos aleatórios. Estes processos têm sido relativamente pouco estudados, devido a não existir teoremas de limite regendo sua distribuição, em comparação com os processos aditivos que estão sujeitos, dentro de certas condições, ao Teorema Central do Limite. Vamos focar o estudo nos processos multiplicativos binomiais e, mais geralmente, nos processos multiplicativos multinomiais. Uma proposta de atividade para ser realizada em sala de aula será discutida. Essa atividade visa levar os alunos a perceberem o comportamento extremo que os processos multiplicativos podem ter.

Palavras-chave: Processos Multiplicativos Aleatórios, Passeio Aleatório, Teorema Central do Limite.

¹Aluna de Mestrado Profissional em Matemática, Turma 2018

Instituição: Universidade Federal de São João del-Rei - UFSJ - Campus Alto Paraopeba

E-mail: oliveira.amandinha@yahoo.com.br

²Orientador do Trabalho de Conclusão de Curso

Departamento de Estatística, Física e Matemática - DEFIM, UFSJ

E-mail: timoteo@ufsj.edu.br

Abstract:

The objective of this work is to present some results necessary to understand the asymptotic behavior of random multiplicative processes. These processes have been relatively less studied, because there are no limit theorems governing their distribution, in comparison with the additive processes that are subject, under certain conditions, to the Central Limit Theorem. We will focus the study on the binomial multiplicative processes and, more generally, on the multinomial multiplicative processes. A proposal for an activity to be carried out in the classroom will be discussed. This activity aims to make students realize the extreme behavior that multiplicative processes can have.

Keywords: Random Multiplicative Processes, Random Walk, Central Limit Theorem.

Sumário

| | | |
|----------|----------------------------------------------------------------------|-----------|
| 1 | Introdução | 5 |
| 2 | Conceitos Básicos | 6 |
| 2.1 | Notações | 6 |
| 2.2 | Análise combinatória | 7 |
| 2.3 | Probabilidade | 8 |
| 2.4 | Variáveis aleatórias e esperança matemática | 10 |
| 2.5 | Distribuição Normal e Teorema Central do Limite | 14 |
| 2.5.1 | Teorema Central do Limite | 14 |
| 2.5.2 | Passeios aleatórios | 15 |
| 3 | Distribuições Binomial e Multinomial | 17 |
| 3.1 | Distribuição Binomial | 17 |
| 3.2 | Distribuição Multinomial | 22 |
| 4 | Processos Multiplicativos Aleatórios | 23 |
| 4.1 | Processo multiplicativo binomial | 24 |
| 4.1.1 | Esperança condicional | 29 |
| 4.2 | Processos multiplicativos multinomiais | 31 |
| 4.2.1 | Processo multiplicativo trinomial | 31 |
| 4.2.2 | Processos multiplicativos multinomiais | 33 |
| 5 | Proposta de Aplicação em Sala de Aula | 34 |
| 5.1 | Jogo do par ou ímpar | 34 |
| 5.1.1 | 1ª etapa (processo aditivo): | 34 |
| 5.1.2 | 2ª etapa (processo multiplicativo): | 36 |
| 5.1.3 | Etapa final | 36 |
| 5.1.4 | Cálculo da probabilidade do Jogo par ou ímpar excluindo “zero dedos” | 37 |
| 6 | Conclusão | 38 |

1 Introdução

O passeio aleatório é um modelo probabilístico que descreve a trajetória de uma partícula após uma sequência de passos aleatórios. Em um passeio aleatório (que é um processo aditivo), com um número suficientemente grande de variáveis aleatórias, podemos aplicar, dentro de determinadas condições, um teorema fundamental da teoria das probabilidades, o Teorema Central do Limite (TCL) [6], e inferir sobre o comportamento assintótico dessa soma de variáveis aleatórias. Assim os processos aditivos, em geral, são tema de vários estudos [12].

Por outro lado, os processos multiplicativos, apesar de importantes, têm sido pouco estudados, pois não possuem uma versão do TCL que permita estudar seu padrão assintótico e fazer inferências. Nota-se, entretanto, que processos multiplicativos surgem de modelos matemáticos simples, tais como equações de recorrência a tempo discreto da forma

$$y_{i+1} = a_i y_i$$

onde y_i é a sequência a ser determinada e a_i é uma sequência de variáveis aleatórias, para $i = 0, 1, \dots$. A solução da equação no tempo n fica

$$y_n = \left[\prod_{i=0}^{n-1} a_i \right] y_0$$

em que se observa o produto das variáveis aleatórias.

Esse tipo de processo multiplicativo aleatório, com um número grande de variáveis aleatórias, pode conter sequências “pouco prováveis” mas com valores extremos. Essas sequências tendem a influenciar demasiadamente o valor médio do processo. Uma estratégia para tentar entender o comportamento de y_i é calcular seu logaritmo, assim $\ln y_i$ se torna um passeio aleatório. Porém, mesmo que possamos aplicar o TCL para inferir sobre o comportamento do $\ln y_i$, há uma perda de informações sobre os momentos mais extremos dos produtos e o comportamento do processo frente a eventos raros fica ofuscado. Essa abordagem foge do escopo do trabalho, para mais detalhes ver na referência [12].

Esse trabalho almeja apresentar alguns resultados necessários para a compreensão do comportamento assintótico dos processos multiplicativos aleatórios. Vamos focar o estudo nos processos multiplicativos binomiais e, mais geralmente, nos processos multiplicativos multinomiais.

Para desenvolver o trabalho, na Seção 2 apresentamos alguns conceitos matemáticos básicos necessários para o desenvolvimento do tema principal, tais como, análise combinatória, espaços de probabilidade e o teorema central do limite. Na Seção 3 discorreremos sobre a distribuição binomial e a relacionamos com um passeio aleatório. Também abordamos a distribuição multinomial. Os processos multiplicativos aleatórios são apresentados na Seção 4, onde o enfoque são os processos multiplicativos binomiais. São deduzidas algumas de suas propriedades estatísticas tais como a média e a moda, além de ser abordada a generalização para os processos multiplicativos multinomiais. Na Seção 5 sugerimos uma proposta de atividade para ser trabalhada em sala de aula envolvendo um processo multiplicativo aleatório. Essa atividade visa levar os alunos a perceberem o comportamento extremo que esses processos podem ter. As considerações finais do trabalho estão na Seção 6.

2 Conceitos Básicos

Nessa seção vamos revisar alguns conceitos básicos necessários para a compreensão do desenvolvimento dos temas abordados nesse trabalho. Se leitor desejar aprofundar-se nos conteúdos poderá consultar as referências [4], [6], [8], [10] e [11].

2.1 Notações

\mathbb{N} denota o conjunto dos números naturais $\{1, 2, \dots\}$.

\mathbb{R} denota o conjunto dos números reais.

\mathbb{Z} denota o conjunto dos números inteiros $\{\dots, -1, 0, 1, 2, \dots\}$.

$\mathbb{E}[X]$ denota o valor esperado da variável aleatória X .

$\mathbb{M}(X)$ denota a moda da variável aleatória X .

$\mathbb{P}(A)$ denota a medida de probabilidade do evento A .

$\mathbb{P}(A, B)$ denota a probabilidade da interseção dos eventos A e B , isto é, $\mathbb{P}(A \cap B)$.

$[X \leq a]$ denota o evento $\{\omega \in \Omega; X(\omega) \leq a\}$. Notações similares podem ser usadas para os eventos $[X = a]$, $[X < a]$, etc.

$i!$ denota o fatorial de $i \in \mathbb{N}$, isto é, $i(i-1) \cdots 2 \cdot 1$. Convencionamos que $0! = 1$.

$\binom{n}{i}$ denota a combinação de n elementos, tomados i a i , isto é, $\frac{n!}{i!(n-i)!}$.

2.2 Análise combinatória

Permutar é o mesmo que trocar elementos de posição. Se uma dada coleção de objetos possui n elementos diferentes, o número de agrupamentos ordenados distintos que podemos obter pela permutação dos elementos é dada por

$$P_n = n!.$$

Quando numa coleção de objetos ocorrem elementos repetidos, o número de agrupamentos ordenados distintos que podemos obter pela permutação dos elementos é dada pela fórmula da permutação com repetição

$$PR_n^{i_1, i_2, \dots, i_k} = \frac{n!}{i_1! i_2! \dots i_k!}$$

onde n é o número de objetos, k é o número de elementos distintos e $n = i_1 + i_2 + \dots + i_k$, sendo i_1 a quantidade de vezes que o primeiro objeto se repete, i_2 a quantidade de vezes que o segundo objeto se repete, e assim por diante.

Em particular, no caso em que $k = 2$, ao fazer $i_1 = i$, temos

$$PR_n^{i, n-i} = \frac{n!}{i!(n-i)!} = \binom{n}{i}.$$

O desenvolvimento do Binômio de Newton e o caso geral do Teorema Multinomial serão usados na Seção 3.

Teorema 2.1 (Teorema multinomial) *Se x_1, x_2, \dots, x_k são k números reais e n é um número natural, então*

$$(x_1 + x_2 + \dots + x_k)^n = \sum_{i_1=0}^n \sum_{i_2=0}^{n-i_1} \dots \sum_{i_{k-1}=0}^{n-i_1-\dots-i_{k-2}} \frac{n!}{i_1! i_2! \dots i_{k-1}! (n-i_1-\dots-i_{k-1})!} x_1^{i_1} x_2^{i_2} \dots x_{k-1}^{i_{k-1}} x_k^{n-i_1-\dots-i_{k-1}}.$$

Ver a demonstração na referência [11].

Quando $k = 2$, temos o caso especial do **Binômio de Newton**:

$$(x + y)^n = \sum_{i=0}^n \binom{n}{i} x^{n-i} y^i.$$

2.3 Probabilidade

Diversos aspectos rigorosos da Teoria de Probabilidade necessitam do emprego da Teoria da Medida para serem definidos ou demonstrados. Contudo a Teoria da Medida foge do escopo deste trabalho e não será abordada. O leitor pode consultar [6] para detalhes que não serão apresentados aqui.

Para definir o que é um espaço de probabilidade, precisamos primeiramente definir o que são espaço amostral, σ -álgebra e a medida de probabilidade.

Ao realizar um experimento, podemos considerar um conjunto não vazio que contém todos os possíveis resultados deste experimento.

Definição 2.1 (Espaço amostral) *O conjunto não vazio Ω de todos os resultados possíveis de um experimento aleatório é chamado de Espaço Amostral.*

De acordo com a referência [8], o espaço amostral pode ser classificado:

- finito, se puder ser colocado em correspondência bi-unívoca com um subconjunto finito de \mathbb{N} ;
- infinito enumerável, se puder ser colocado em correspondência bi-unívoca com \mathbb{N} ;
- não enumerável, caso não atenda a nenhuma das duas condições anteriores

Exemplo 2.1 *Considere um experimento aleatório dividido em dois ensaios repetidos idênticamente, sendo que em cada ensaio podemos obter dois resultados distintos b_1 ou b_2 . O espaço amostral desse experimento é dado por*

$$\Omega = \{(b_1, b_1), (b_1, b_2), (b_2, b_1), (b_2, b_2)\}.$$

Definição 2.2 *Uma classe de subconjuntos de Ω , representada por \mathcal{F} , é uma σ -álgebra se e somente se \mathcal{F} possui as seguintes propriedades:*

1. $\Omega \in \mathcal{F}$;
2. Se $A \in \mathcal{F}$ então $A^C \in \mathcal{F}$;
3. Se $A_n \in \mathcal{F}$ para $n = 1, 2, \dots$ então $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

Os elementos da σ -álgebra \mathcal{F} são chamados de eventos e somente a eles atribuímos uma medida positiva chamada probabilidade.

Exemplo 2.2 Uma σ -álgebra \mathcal{F} associada ao espaço amostral Ω do Exemplo 2.1 pode ser dada por

$$\begin{aligned} \mathcal{F} = \{ & \emptyset, \{(b_1, b_1)\}, \{(b_1, b_2)\}, \{(b_2, b_1)\}, \{(b_2, b_2)\}, \{(b_1, b_1), (b_1, b_2)\}, \{(b_1, b_1), (b_2, b_1)\}, \\ & \{(b_1, b_1), (b_2, b_2)\}, \{(b_1, b_2), (b_2, b_1)\}, \{(b_1, b_2), (b_2, b_2)\}, \{(b_2, b_1), (b_2, b_2)\}, \\ & \{(b_1, b_1), (b_1, b_2), (b_2, b_1)\}, \{(b_1, b_1), (b_1, b_2), (b_2, b_2)\}, \\ & \{(b_1, b_1), (b_2, b_1), (b_2, b_2)\}, \{(b_1, b_2), (b_2, b_1), (b_2, b_2)\}, \Omega \}. \end{aligned}$$

Definição 2.3 Uma medida positiva \mathbb{P} definida na σ -álgebra \mathcal{F} é denominada medida probabilidade, se atende aos axiomas:

1. $\mathbb{P}(\Omega) = 1$;
2. Para $A_1, A_2, \dots \in \mathcal{F}$ eventos disjuntos, isto é, $A_i \cap A_j = \emptyset$, se $i \neq j$, então

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

Definição 2.4 Um espaço de probabilidade é uma trinca $(\Omega, \mathcal{F}, \mathbb{P})$, formado pelo espaço amostral Ω , uma σ -álgebra \mathcal{F} em Ω , e a medida de probabilidade \mathbb{P} .

Exemplo 2.3 Com relação aos exemplos 2.1 e 2.2, podemos definir a medida de probabilidade \mathbb{P} , que satisfaz

$$\mathbb{P}(\{(b_i, b_j)\}) = \begin{cases} p^2, & \text{se } i = j = 1, \\ p(1-p), & \text{se } i \neq j, \\ (1-p)^2, & \text{se } i = j = 2, \end{cases}$$

para algum $0 < p < 1$. A trinca $(\Omega, \mathcal{F}, \mathbb{P})$ é um exemplo de espaço de probabilidade.

Definição 2.5 Sejam $(\Omega, \mathcal{F}, \mathbb{P})$ um espaço de probabilidade e $A, B \in \mathcal{F}$ com $\mathbb{P}(B) > 0$. A probabilidade condicional de A dado B é definida por

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

Definição 2.6 *Sejam $(\Omega, \mathcal{F}, \mathbb{P})$ um espaço de probabilidade e $A, B \in \mathcal{F}$. Os eventos A e B são independentes se*

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B).$$

2.4 Variáveis aleatórias e esperança matemática

Frequentemente precisamos indicar uma característica numérica para uma dada amostra $\omega \in \Omega$. Isto é feito por meio de uma função de Ω em \mathbb{R} chamada de variável aleatória.

Definição 2.7 *Seja $(\Omega, \mathcal{F}, \mathbb{P})$ um espaço de probabilidade. Denominamos de variável aleatória qualquer função $X : \Omega \rightarrow \mathbb{R}$ tal que*

$$X^{-1}(I) = \{\omega \in \Omega : X(\omega) \in I\} \in \mathcal{F},$$

para todo intervalo $I \subset \mathbb{R}$.

Pela definição, para uma variável aleatória X podemos dizer, então, que a imagem inversa de intervalos $I \subset \mathbb{R}$ pertencem a σ -álgebra \mathcal{F} . Em linguagem comum, uma variável aleatória é uma função do espaço amostral Ω nos números reais, para a qual é possível calcular a probabilidade de ocorrência de seus valores.

Exemplo 2.4 *Para o espaço de probabilidade do Exemplo 2.3, podemos definir uma variável aleatória que conta as vezes que b_1 foi obtido como resultado do experimento.*

$$X((b_i, b_j)) = \begin{cases} 2, & \text{se } i = j = 1, \\ 1, & \text{se } i \neq j, \\ 0, & \text{se } i = j = 2, \end{cases}$$

Um tipo especial de variável aleatória é a função indicadora de um dado conjunto. Definimos \mathbf{I}_A como a função indicadora do conjunto $A \in \mathcal{F}$, da seguinte forma

$$\mathbf{I}_A(\omega) = \begin{cases} 1, & \text{se } \omega \in A \\ 0, & \text{se } \omega \notin A \end{cases}$$

para $\omega \in \Omega$.

A cada variável aleatória em um dado espaço de probabilidade, podemos associar uma função de \mathbb{R} em $[0, 1]$ denominada Função de Distribuição Acumulada (ou simplesmente Função de Distribuição).

Definição 2.8 *Sejam $(\Omega, \mathcal{F}, \mathbb{P})$ um espaço de probabilidade e X uma variável aleatória. A Função de Distribuição Acumulada F_X é a função $F_X : \mathbb{R} \rightarrow [0, 1]$ dada por*

$$F_X(x) = \mathbb{P}([X \leq x]) = \mathbb{P}(\{\omega \in \Omega : X(\omega) \leq x\}),$$

para cada $x \in \mathbb{R}$.

É comum escrever-se F , omitindo o subscrito X em F_X quando se tem clareza que F é a distribuição de X . Em geral vamos utilizar a notação simplificada $\mathbb{P}(X \leq x)$ para representar $\mathbb{P}([X \leq x])$. O mesmo vale para casos análogos, tais como, $\mathbb{P}(X = x)$, $\mathbb{P}(X < x)$, e assim por diante.

A seguir listamos algumas propriedades que a função de distribuição F obedece (ver [6, 8]):

- $\lim_{x \rightarrow -\infty} F(x) = 0$ e $\lim_{x \rightarrow +\infty} F(x) = 1$;
- F é contínua à direita;
- F é não decrescente.

A variável aleatória é dita *discreta* quando sua imagem for um conjunto finito ou infinito enumerável. Se $\{x_1, x_2, \dots\}$ for o conjunto discreto de valores assumidos por X , podemos definir a função de probabilidade

$$p(x_i) = \mathbb{P}(X = x_i).$$

Se existir uma função positiva f tal que F possa ser escrita como

$$F(x) = \int_{-\infty}^x f(w)dw,$$

para todo $x \in \mathbb{R}$, a variável aleatória X é dita (absolutamente) *contínua* e a função f é denominada *função densidade*.

Definição 2.9 (Esperança Matemática) *Se X for uma variável aleatória discreta assumindo os valores x_1, x_2, x_3, \dots e se $p(\cdot)$ for sua função de probabilidade, a esperança matemática de X é definida pela série*

$$\mathbb{E}[X] = \sum_{i=1}^{\infty} x_i p(x_i).$$

Se X for uma variável aleatória contínua e se f for sua função densidade, a esperança matemática de X é calculada pela integral

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} xf(x)dx$$

A esperança de uma variável aleatória também pode ser chamada de média ou valor esperado.

Exemplo 2.5 A variável aleatória X definida no exemplo 2.4 tem média dada por

$$\mathbb{E}[X] = 2 \times p^2 + 1 \times 2p(1 - p) = 2p.$$

Utilizando o conceito de probabilidade condicional podemos definir a função de distribuição condicional e a esperança condicional associadas a uma variável aleatória.

Definição 2.10 Seja X uma variável aleatória sobre um espaço de probabilidade $(\Omega, \mathcal{F}, \mathbb{P})$. A distribuição condicional de X dado o evento $A \in \mathcal{F}$ é

$$F(x|A) = \mathbb{P}(X \leq x|A),$$

para $x \in \mathbb{R}$.

Se X for uma variável aleatória discreta, a esperança condicional de X dado A é

$$\mathbb{E}[X|A] = \sum_{i=1}^{\infty} x_i p(x_i|A),$$

onde $p(x_i|A) = \mathbb{P}(X = x_i|A)$.

A *Moda* de uma variável aleatória X informa o valor mais frequente (ou mais provável) dentre os possíveis valores que X pode assumir.

Definição 2.11 Se X for uma variável aleatória discreta assumindo os valores x_1, x_2, \dots e se $p(\cdot)$ for sua função de probabilidade, a *Moda* de X é dada por,

$$\mathbb{M}(X) = \arg \max_{i \geq 1} \{p(x_i)\},$$

ou seja, pelo valor x_i que maximiza a função $p(\cdot)$. Se X for uma variável aleatória contínua com função densidade f ,

$$\mathbb{M}(X) = \arg \max_{x \in \mathbb{R}} \{f(x)\},$$

isto é, o valor x que a variável aleatória X assume que maximiza f .

Note que a *moda* depende apenas da distribuição e não dos valores que a variável aleatória assume.

A *Variância* fornece uma medida para o grau de dispersão (ou concentração) de valores X em torno do valor esperado. Se $\mathbb{E}[X] = \mu$ for o valor esperado da variável X então a variância de X é

$$\text{VAR}(X) = \mathbb{E}[(X - \mu)^2].$$

Desenvolvendo esta expressão podemos escrever também que

$$\text{VAR}(X) = \mathbb{E}[X^2] - (\mu)^2.$$

O desvio padrão σ de uma variável aleatória X com média μ é definido como a raiz quadrada da variância de X

$$\sigma = \sqrt{\text{VAR}(X)}.$$

O k -ésimo momento de uma variável aleatória X é definido por $\mathbb{E}[X^k]$.

A *Covariância* entre duas variáveis aleatórias X e Y mede o quanto estas variáveis se apresentam relacionadas. Se $\mathbb{E}[X] = \mu_X$ e $\mathbb{E}[Y] = \mu_Y$, então a covariância entre X e Y é

$$\text{COVAR}(X, Y) = \mathbb{E}[(X - \mu_X)(Y - \mu_Y)].$$

Desenvolvendo esta expressão, podemos calcular a covariância entre X e Y também por

$$\text{COVAR}(X, Y) = \mathbb{E}[XY] - \mu_Y \mu_X.$$

A demonstração pode ser consultada na referência [11].

Definição 2.12 *Considere duas variáveis aleatórias X e Y assumindo valores em $\mathbb{I} \subset \mathbb{R}$. Dizemos que X e Y são identicamente distribuídas se*

$$\mathbb{P}(X \leq x) = \mathbb{P}(Y \leq x), \forall x \in \mathbb{I}.$$

Definição 2.13 *Sejam X_1, \dots, X_n variáveis aleatórias sobre Ω . Se*

$$\mathbb{P}(X_1 \in A_1, X_2 \in A_2, \dots, X_n \in A_n) = \mathbb{P}(X_1 \in A_1) \cdot \dots \cdot \mathbb{P}(X_n \in A_n),$$

para cada $A_i \in \mathcal{F}$, então dizemos que X_1, \dots, X_n são independentes.

Uma sequência de variáveis aleatórias é independente e identicamente distribuída se cada variável aleatória tiver a mesma distribuição de probabilidade das outras e todas forem mutuamente independentes.

2.5 Distribuição Normal e Teorema Central do Limite

Uma variável aleatória contínua X tem distribuição normal (ou Gaussiana) se sua função de densidade de probabilidade for dada por

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[-\frac{1}{2} \left(\frac{x - \mu}{\sigma} \right)^2 \right], \quad x \in (-\infty, \infty),$$

onde $\mu = \mathbb{E}[X]$ e $\sigma = \sqrt{\text{VAR}(X)}$ são, respectivamente, a média e o desvio padrão de X . Usamos a notação $X \sim N(\mu, \sigma^2)$ para indicar que X tem distribuição normal com média μ e variância σ^2 .

2.5.1 Teorema Central do Limite

Considere uma sequência de variáveis aleatórias independentes X_1, X_2, \dots . Existem dois teoremas fundamentais relativos ao comportamento de tais sequências: o teorema central do limite e a lei dos grandes números. A lei dos grandes números não será abordada nesse trabalho, para mais detalhes consultar a referência [14].

O teorema central do limite (TCL) é um dos resultados mais importante da probabilidade e da estatística. Ele é muito útil na inferência estatística, cujo objetivo é fazer afirmações a respeito de uma população através de uma amostra, pois pode ser interpretado como afirmando que à medida que o tamanho da amostra de uma população vai aumentando, a distribuição amostral da sua média vai se aproximando de uma distribuição normal, independentemente da forma da distribuição da população. Dessa forma é usado, por exemplo, para estimar a média populacional através da média amostral ou também o desvio padrão da média populacional através do desvio padrão da média amostral dessa população [3].

No presente texto, o teorema será útil para tecer considerações sobre o comportamento de sequências de variáveis binomiais na seção 3.

Teorema 2.2 [Teorema central do limite] *Sejam X_1, X_2, \dots uma sequência de variáveis aleatórias independentes e identicamente distribuídas, onde cada variável possui média μ e variância σ^2 , em que $0 < \sigma^2 < \infty$. Seja $S_n = \sum_{i=1}^n X_i$, então*

$$\frac{S_n - n\mu}{\sigma\sqrt{n}}$$

converge em distribuição para uma variável normal com média 0 e variância 1 quando $n \rightarrow \infty$.

Demonstrações para esta e outras versões do TCL podem ser vistas em [1], [6] ou [14].

2.5.2 Passeios aleatórios

Para dar continuidade ao tema do trabalho é importante relacionarmos passeios aleatórios (processos aditivos) e o teorema central do limite. Caso o leitor deseje aprofundar-se no assunto, poderá consultar as referências [1], [13] e [14].

A seguir apresentamos uma definição geral de passeio aleatório baseada na de Kallenberg [7].

Definição 2.14 *Sejam X_1, X_2, \dots variáveis aleatórias independentes e identicamente distribuídas. Seja $S_0 = C$, onde C é uma constante e*

$$S_n = S_0 + \sum_{i=1}^n X_i,$$

para cada $n \geq 1$. O processo $\{S_n, n \geq 0\}$ é chamado passeio aleatório.

Notamos, então, que se num passeio aleatório as variáveis X_i possuírem média e variância finitas, podemos aplicar o TCL e inferir as propriedades assintóticas de S_n .

Exemplo 2.6 *Vamos ilustrar um passeio aleatório unidimensional de uma partícula sobre a reta numérica, considerando as seguintes condições: a posição inicial da partícula será o ponto zero; o deslocamento será constante e igual a 1; o deslocamento pode ocorrer para a direita (+1) ou para a esquerda (-1); para definir o sentido do deslocamento da partícula será realizado o lançamento de uma moeda honesta (no caso de cara se deslocará para a direita e no caso de coroa se deslocará para a esquerda).*

A posição da partícula depois de n lançamentos da moeda será soma dos resultados obtidos nos n deslocamentos.

- Após o 1º lançamento, a partícula poderá estar localizada no ponto $+1$ ou -1 , havendo apenas uma chance de resultar em $+1$ e uma chance de resultar em -1 .
- Após o 2º lançamento, a partícula poderá estar localizada nos pontos $+2$, 0 ou -2 , havendo uma chance de resultar em $+2$, uma chance de resultar em -2 e duas chances de resultar em 0 .
- Após o 3º lançamento, a partícula poderá estar localizada nos pontos $+3$, $+1$, -1 ou -3 , havendo uma chance de resultar em $+3$, uma chance de resultar em -3 , três chances de resultar em $+1$ e três chances de resultar em -1 . E assim por diante.

A Tabela 1 mostra as possibilidades de deslocamentos da partícula até o 6º lançamento. Como podemos notar, as possibilidades de posicionamento são números binomiais e formam um triângulo de Pascal. A posição Z da partícula será descrita no Exemplo 3.1, após a apresentação da Distribuição Binomial.

Tabela 1: Possibilidades de deslocamentos da partícula até o 6º lançamento.

| | Posição na reta numérica | | | | | | | | | | | | |
|-------------|--------------------------|----|----|----|----|----|----|----|----|----|----|----|----|
| Lançamento | -6 | -5 | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 | +5 | +6 |
| 0 (inicial) | | | | | | | 1 | | | | | | |
| 1º | | | | | | 1 | | 1 | | | | | |
| 2º | | | | | 1 | | 2 | | 1 | | | | |
| 3º | | | | 1 | | 3 | | 3 | | 1 | | | |
| 4º | | | 1 | | 4 | | 6 | | 4 | | 1 | | |
| 5º | | 1 | | 5 | | 10 | | 10 | | 5 | | 1 | |
| 6º | 1 | | 6 | | 15 | | 20 | | 15 | | 6 | | 1 |

Como acabamos de ver, o passeio aleatório do exemplo exibe características relacionadas a uma distribuição binomial. A seguir, portanto, veremos a formalização do modelo de distribuição binomial e de sua generalização, o modelo multinomial.

3 Distribuições Binomial e Multinomial

3.1 Distribuição Binomial

A distribuição binomial pode modelar um experimento que é dividido em n ensaios repetidos identicamente, em cada um dos quais podemos obter sucesso (com probabilidade p) ou fracasso (com probabilidade $1 - p$). O objetivo é contabilizar a quantidade de sucessos no conjunto de ensaios.

Vamos definir o espaço de probabilidade. Fixe um número $n \in \mathbb{N}$, $n \neq 0$, um conjunto binário $B = \{b_1, b_2\}$, e um número real p , $0 < p < 1$. A ideia subjacente é ter um modelo em que b_1 ocorra com probabilidade p , e b_2 com probabilidade $1 - p$.

O espaço amostral será o conjunto Ω de todas as n -uplas $(\omega_1, \dots, \omega_n)$ em que $\omega_i \in B$. Note, portanto, que $\Omega = B^n$ é discreto com cardinalidade 2^n .

O espaço de eventos \mathcal{F} será dado pela coleção de todos os subconjuntos de Ω , isto é, \mathcal{F} é o conjunto das partes de Ω , também simbolizado por $\mathcal{F} = 2^\Omega$.

A medida de probabilidade \mathbb{P} sobre \mathcal{F} será a função

$$\mathbb{P} : \mathcal{F} \rightarrow [0, 1],$$

dada da seguinte forma. Fixe $\omega = (\omega_1, \dots, \omega_n) \in \Omega$ e suponha que b_1 ocorra i vezes em ω (consequentemente b_2 ocorre $n - i$ vezes). Então

$$\mathbb{P}(\{\omega\}) = p^i(1 - p)^{n-i}.$$

Agora vamos definir a variável aleatória $\mathcal{S} : \Omega \rightarrow \mathbb{R}$, que para cada $\omega = (\omega_1, \dots, \omega_n)$ conta o número de vezes que b_1 ocorre em ω , isto é,

$$\mathcal{S}(\omega) = \sum_{i=1}^n \mathbf{I}_{\{b_1\}}(\omega_i),$$

lembrando que \mathbf{I}_A é a função indicadora do conjunto A .

O evento $[\mathcal{S} = i]$ significa o conjunto de todas as n -uplas em que b_1 ocorre i vezes, logo contém $\binom{n}{i}$ elementos. A distribuição de \mathcal{S} é dada por

$$\mathbb{P}(\mathcal{S} = i) = \binom{n}{i} p^i (1 - p)^{n-i},$$

para cada $i \in \{0, 1, \dots, n\}$.

A variável aleatória \mathcal{S} portanto tem distribuição binomial.

O valor esperado de \mathcal{S} é dado por

$$\mathbb{E}[\mathcal{S}] = np,$$

e a variância, por

$$\text{VAR}(\mathcal{S}) = np(1 - p),$$

(as demonstrações podem ser vistas na referência [11]).

Precisamos saber como encontrar a moda de \mathcal{S} , para isso, temos a seguinte proposição.

Proposição 3.1 *A moda de \mathcal{S} é dada pelo menor inteiro i^* tal que $i^* \geq np - q$, onde $q = 1 - p$.*

Demonstração: De início, note que \mathcal{S} tem função de probabilidade dada por

$$p(i) = \mathbb{P}(\mathcal{S} = i).$$

Vamos supor que o maior valor da função $p(\cdot)$ não ocorra nos extremos. Para $0 \leq i \leq n - 1$ defina

$$r_{i,i+1} = \frac{p(i+1)}{p(i)}.$$

Observe que $r_{0,1}$ será maior do que 1 e $r_{n-1,n}$ será menor do que 1. Assim queremos o menor i que faz com que $r_{i,i+1}$ se torne menor do que ou igual a 1. Temos

$$r_{i,i+1} = \frac{\binom{n}{i+1} p^{i+1} (1-p)^{n-i-1}}{\binom{n}{i} p^i (1-p)^{n-i}} = \frac{i!(n-i)!p}{(i+1)!(n-i-1)!(1-p)} = \frac{n-i}{i+1} \frac{p}{1-p}.$$

Para que $r_{i,i+1}$ seja menor do que ou igual a 1, devemos ter

$$\frac{n-i}{i+1} \frac{p}{1-p} \leq 1$$

que implica em

$$i \geq np - (1 - p).$$

□

Podemos observar que

$$\mathbb{E}[\mathcal{S}] - \mathbb{M}(\mathcal{S}) = np - i^* \leq q.$$

É útil podermos definir uma sequência de variáveis aleatórias em Ω da seguinte forma. Para cada $m \in \{1, \dots, n\}$, seja $\mathcal{S}_m : \Omega \rightarrow \mathbb{R}$ a função que conta o número de vezes que b_1 ocorre nos primeiros m ensaios, isto é,

$$\mathcal{S}_m(\omega) = \sum_{i=1}^m \mathbf{I}_{\{b_1\}}(\omega_i).$$

O evento $[\mathcal{S}_m = i]$ significa o conjunto de todas as n -uplas em que b_1 ocorre i vezes nas primeiras m coordenadas. A distribuição de \mathcal{S}_m é dada por

$$\mathbb{P}(\mathcal{S}_m = i) = \begin{cases} \binom{m}{i} p^i (1-p)^{m-i}, & \text{se } i \leq m \\ 0, & \text{se } i > m \end{cases}$$

para cada $i \in \{0, 1, \dots, n\}$. Note que $\mathcal{S}_n = \mathcal{S}$.

Uma análise importante que podemos fazer a partir da definição da sequência $(\mathcal{S}_1, \dots, \mathcal{S}_n)$ é sobre probabilidades condicionais da forma

$$\mathbb{P}(\mathcal{S}_n = i | \mathcal{S}_m = j),$$

isto é, a probabilidade de que b_1 ocorra i vezes em n ensaios dado que já ocorreu j vezes nos primeiros m ensaios. Se $i < j$, esta probabilidade é zero. No caso em que $i = j$, então do ensaio $m+1$ até o ensaio n ocorre somente b_2 , logo a probabilidade é $(1-p)^{n-m}$. Finalmente, se $i > j$, a probabilidade será igual à probabilidade de b_1 ocorrer $i-j$ vezes nos $n-m$ ensaios restantes. Vamos resumir estas informações na proposição seguinte.

Proposição 3.2 *Para $m \in \{1, \dots, n\}$ e $i, j \in \{0, \dots, n\}$, temos que*

$$\mathbb{P}(\mathcal{S}_n = i | \mathcal{S}_m = j) = \begin{cases} 0, & \text{se } i < j \\ \binom{n-m}{i-j} p^{i-j} (1-p)^{n-m-i+j}, & \text{se } i \geq j \end{cases}$$

Um outro conjunto de variáveis aleatórias que podem ser definidas sobre Ω é a sequência $(\mathcal{X}_1, \dots, \mathcal{X}_n)$, em que cada variável \mathcal{X}_m indica se b_1 ocorreu no m -ésimo ensaio, isto é

$$\mathcal{X}_m(\omega) = \mathbf{I}_{\{b_1\}}(\omega_m)$$

para cada $\omega = (\omega_1, \dots, \omega_n)$.

Decorre que, por serem os ensaios independentes, as variáveis $\mathcal{X}_1, \dots, \mathcal{X}_n$ são independentes. Temos por exemplo,

$$\begin{aligned}\mathbb{P}(\mathcal{X}_i = a, \mathcal{X}_j = b) &= \mathbb{P}(\mathbf{I}_{\{b_1\}}(\omega_i) = a, \mathbf{I}_{\{b_1\}}(\omega_j) = b) \\ &= \mathbb{P}(\mathbf{I}_{\{b_1\}}(\omega_i) = a) \mathbb{P}(\mathbf{I}_{\{b_1\}}(\omega_j) = b) \\ &= \mathbb{P}(\mathcal{X}_i = a) \mathbb{P}(\mathcal{X}_j = b),\end{aligned}$$

para cada $a, b \in \mathbb{R}$.

Além disso, para todo $a \in \mathbb{R}$,

$$\begin{aligned}\mathbb{P}(\mathcal{X}_i = a) &= \mathbb{P}(\mathbf{I}_{\{b_1\}}(\omega_i) = a) \\ &= \mathbb{P}(\mathbf{I}_{\{b_1\}}(\omega_j) = a) \\ &= \mathbb{P}(\mathcal{X}_j = a).\end{aligned}$$

Ou seja, \mathcal{X}_i e \mathcal{X}_j têm a mesma distribuição.

De forma geral, temos a seguinte proposição.

Proposição 3.3 *As variáveis $\mathcal{X}_1, \dots, \mathcal{X}_n$ são independentes e identicamente distribuídas.*

Note que as sequências $\{\mathcal{S}_m\}$ e $\{\mathcal{X}_m\}$ se relacionam:

$$\mathcal{S}_m = \mathcal{S}_{m-1} + \mathcal{X}_m,$$

ou ainda,

$$\mathcal{S}_m = \sum_{i=1}^m \mathcal{X}_i.$$

Isso mostra que $\{\mathcal{S}_m\}$ é um passeio aleatório.

Vemos, além disso, que a média e a variância de \mathcal{X}_i são dadas, respectivamente, por

$$\begin{aligned}\mathbb{E}[\mathcal{X}_i] &= p \\ \text{VAR}(\mathcal{X}_i) &= p(1-p),\end{aligned}$$

para cada $i \geq 1$. Logo podemos aplicar o Teorema Central do Limite³ (Teorema 2.2) para obter o padrão assintótico da sequência $\{\mathcal{S}_m\}$, pois a sequência $\{\mathcal{X}_i\}$ é independente e identicamente distribuída pela Proposição 3.3. Conseqüentemente,

$$\frac{\mathcal{S}_n - np}{\sqrt{np(1-p)}}$$

tem distribuição normal com média 0 e variância 1, quando $n \rightarrow \infty$.

Em particular o comportamento de \mathcal{S}_n para n grande é normal com média np e variância $np(1-p)$. Lembrando que, numa variável normal, a média e a moda são iguais, então, para n grande a moda de \mathcal{S}_n é np .

Exemplo 3.1 *A posição da partícula no Exemplo 2.6 após o n -ésimo lançamento da moeda pode ser indicada por uma variável aleatória \mathcal{Z} cuja distribuição depende de \mathcal{S} . De fato, defina*

$$\begin{aligned}\mathcal{Y}_i &= 2\mathcal{X}_i - 1, \text{ para } i = 1, 2, \dots \\ \mathcal{Z} &= \sum_{i=1}^n \mathcal{Y}_i.\end{aligned}$$

Vemos que os valores de \mathcal{Y}_i são $+1$ ou -1 . Vamos supor que a probabilidade da moeda dar cara seja p , então a partícula anda para a direita com probabilidade

$$p = \mathbb{P}(\mathcal{Y}_i = +1)$$

e após n lançamentos a posição da partícula será dada por \mathcal{Z} . Note que

$$\mathcal{Z} = \sum_{i=1}^n \mathcal{Y}_i = \sum_{i=1}^n [2\mathcal{X}_i - 1] = 2\mathcal{S} - n.$$

Além disso, a distribuição de \mathcal{Z} é

$$\begin{aligned}\mathbb{P}(\mathcal{Z} = m) &= \mathbb{P}(2\mathcal{S} = m + n) \\ &= \mathbb{P}\left(\mathcal{S} = \frac{m + n}{2}\right),\end{aligned}$$

para cada $m \in \mathbb{Z}$.

³Formalmente seria necessário considerar o espaço amostral $\Omega = B^\infty$ formado por seqüências infinitas, porém o detalhamento do espaço de probabilidade decorrente foge ao escopo deste texto.

Em particular, \mathcal{Z} é um passeio aleatório, e

$$\mathbb{E}[\mathcal{Z}] = 2np - n = n(p - q),$$

$$\text{VAR}(\mathcal{Z}) = 4np(1 - p).$$

Note também que a moda de \mathcal{Z} é $n(p - q)$.

3.2 Distribuição Multinomial

A distribuição multinomial generaliza o desenvolvimento da binomial, podendo modelar um experimento também dividido em n ensaios repetidos identicamente, mas neste caso em cada um dos ensaios podemos obter k resultados distintos (cada um com probabilidade p_i de acontecer, $1 \leq i \leq k$).

Para definir o espaço de probabilidade, fixe um conjunto M com k elementos, $M = \{m_1, m_2, \dots, m_k\}$, bem como fixe k números reais p_j , $0 < p_j < 1$, tais que $\sum_{j=1}^k p_j = 1$. O espaço amostral será o conjunto Ω de todas as n -uplas $(\omega_1, \dots, \omega_n)$ em que $\omega_i \in M$. Note, portanto, que $\Omega = M^n$ é discreto com cardinalidade k^n . O espaço de eventos \mathcal{F} será dado pela coleção de todos os subconjuntos de Ω , isto é, $\mathcal{F} = 2^\Omega$.

A medida de probabilidade \mathbb{P} sobre \mathcal{F} será dada da seguinte forma. Fixe $\omega = (\omega_1, \dots, \omega_n) \in \Omega$ e suponha que m_1 ocorra i_1 vezes, m_2 ocorra i_2 vezes, e assim sucessivamente. Então

$$\mathbb{P}(\{\omega\}) = p_1^{i_1} p_2^{i_2} \cdots p_k^{i_k}.$$

Para cada j tal que $1 \leq j \leq k$, definimos a variável aleatória $\tilde{\mathcal{S}}_j : \Omega \rightarrow \mathbb{R}$, que para cada $\omega = (\omega_1, \dots, \omega_n)$ conta o número de vezes que m_j ocorre em ω , isto é,

$$\tilde{\mathcal{S}}_j(\omega) = \sum_{i=1}^n \mathbf{I}_{\{m_j\}}(\omega_i).$$

O evento $[\tilde{\mathcal{S}}_1 = i_1, \dots, \tilde{\mathcal{S}}_k = i_k]$ significa o conjunto de todas as n -uplas em que m_1 ocorre i_1 vezes, m_2 ocorre i_2 vezes, e assim por diante. A distribuição de $(\tilde{\mathcal{S}}_1, \dots, \tilde{\mathcal{S}}_k)$ é dada por

$$\mathbb{P}(\tilde{\mathcal{S}}_1 = i_1, \dots, \tilde{\mathcal{S}}_k = i_k) = \frac{n!}{i_1! i_2! \cdots i_k!} p_1^{i_1} p_2^{i_2} \cdots p_k^{i_k}.$$

O vetor aleatório $(\tilde{\mathcal{S}}_1, \dots, \tilde{\mathcal{S}}_k)$ tem distribuição multinomial. O valor esperado de $\tilde{\mathcal{S}}_j$ é dado por

$$\mathbb{E}[\tilde{\mathcal{S}}_j] = np_j,$$

a variância, por

$$\text{VAR}(\tilde{\mathcal{S}}_j) = np_j(1 - p_j),$$

e a covariância entre $\tilde{\mathcal{S}}_u$ e $\tilde{\mathcal{S}}_v$, por

$$\text{COVAR}(\tilde{\mathcal{S}}_u, \tilde{\mathcal{S}}_v) = -np_u p_v,$$

(ver [11]). Note que cada $\tilde{\mathcal{S}}_j$ tem distribuição binomial.

4 Processos Multiplicativos Aleatórios

Nesta seção vamos apresentar algumas propriedades estatísticas do produto de n variáveis aleatórias. Esse é um dos temas abordados no artigo [12].

Como vimos na subseção 2.5.1 o TCL, dentro de determinadas condições, pode ser utilizado para se apreender o padrão assintótico da soma de variáveis aleatórias. Decorre que os processos aditivos estão cercados de vários estudos. Por outro lado, os processos multiplicativos, apesar de importantes, têm sido pouco estudados, pois não possuem uma versão do TCL que permita estudar seu padrão assintótico e fazer inferências.

Uma forma de tentar entender o comportamento assintótico da distribuição do produto \mathcal{P}_n de n variáveis é analisando o logaritmo do produto, $\ln \mathcal{P}_n$. Este processo transforma o produto de variáveis aleatórias numa soma dos logaritmos destas variáveis. Porém, mesmo que esta soma atenda às condições de alguma versão do TCL, há uma perda de informações sobre os momentos mais extremos dos produtos e o comportamento do processo frente a eventos raros fica ofuscado.

Muitos modelos matemáticos são dados em termos de equações de recorrência a tempo discreto [5]. Quando os coeficientes são aleatórios, podemos ver que a solução é um processo multiplicativo. Considere, por exemplo, a equação a tempo discreto

$$y_{i+1} = a_i y_i + b_i,$$

onde y_i é a sequência a ser determinada, a_i, b_i são em geral sequências aleatórias de números reais, para $i = 0, 1, \dots$. A solução da equação no tempo $n + 1$ pode ser escrita como

$$y_{n+1} = \left[\prod_{i=0}^n a_i \right] y_0 + \sum_{j=1}^n \left[\prod_{i=j}^n a_i \right] b_{j-1},$$

em que se observa o produto de variáveis aleatórias.

Um exemplo específico de aplicação deste modelo é o fracionamento de rochas [12]. Neste caso, ponha $b_i = 0$ e considere que y_i é o tamanho de uma rocha no i -ésimo estágio de sua fragmentação. Então a_i é o fator de redução da rocha entre os estágios i e $i + 1$ e o tamanho da rocha na n -ésima iteração fica dado por

$$y_n = \left[\prod_{i=0}^{n-1} a_i \right] y_0.$$

Uma outra questão que pode surgir é reiniciar o processo multiplicativo para gerar distribuições estacionárias [9]. Como já abordado, o processo $y_{i+1} = a_i y_i$ permite, dentro de determinadas condições, que se conclua sobre a convergência de $\ln y_i$ para uma distribuição log-normal⁴, porém dependente do tempo, ou seja, não estacionária. Contudo, considerando que o produto acumulado y_{i+1} possa ser reiniciado, isto é, a cada etapa do processo, $y_{i+1} = y_0$ com probabilidade q ou $y_{i+1} = a_i y_i$ com probabilidade $1 - q$, então o sistema pode desenvolver uma distribuição estacionária. Esse modelo de processo multiplicativo, com possibilidade de reinício, será tema para pesquisas em trabalhos futuros.

A seguir iremos focar o estudo nos processos multiplicativos binomiais e em seguida generalizar para os processos multiplicativos multinomiais.

4.1 Processo multiplicativo binomial

Fixe um número $n \in \mathbb{N}$, dois números reais distintos $b_1, b_2 > 0$ e um número real p , $0 < p < 1$. Defina também $q = 1 - p$. Estabelecemos o espaço de probabilidade $(\Omega, \mathcal{F}, \mathbb{P})$, onde Ω é o conjunto de todas a n -uplas $(\omega_1, \dots, \omega_n)$ em que $\omega_i \in \{b_1, b_2\}$ e \mathcal{F} é a coleção de todos os subconjuntos de Ω . A medida de probabilidade \mathbb{P} sobre \mathcal{F} é dada da seguinte forma. Fixe $\omega = (\omega_1, \dots, \omega_n) \in \Omega$ e suponha que b_1 ocorra i vezes em ω (consequentemente b_2 ocorre $n - i$ vezes). Então definimos

$$\mathbb{P}(\{\omega\}) = p^i (1 - p)^{n-i}.$$

Agora vamos definir a variável aleatória $\mathcal{P} : \Omega \rightarrow \mathbb{R}$, que para cada $\omega = (\omega_1, \dots, \omega_n)$

⁴Uma variável aleatória X tem distribuição log-normal se $Y = \log(X)$ tem distribuição normal.

calcula o produto $\omega_1\omega_2\dots\omega_n$, isto é,

$$\mathcal{P}(\omega) = \prod_{i=1}^n \omega_i.$$

Vamos calcular a distribuição de \mathcal{P} . Note que o evento $[\mathcal{P} = b_1^i b_2^{n-i}]$ representa as amostras $\omega \in \Omega$ em que b_1 ocorre i vezes. A quantidade dessas amostras é, portanto, $\binom{n}{i}$. Logo

$$\mathbb{P}(\mathcal{P} = b_1^i b_2^{n-i}) = \binom{n}{i} p^i (1-p)^{n-i}.$$

Assim como na Seção 3.1, podemos definir a variável aleatória \mathcal{S} que conta a quantidade de vezes que b_1 ocorre. Sendo assim, a distribuição de \mathcal{S} é dada por

$$\mathbb{P}(\mathcal{S} = i) = \binom{n}{i} p^i (1-p)^{n-i}.$$

Vemos, portanto, que \mathcal{P} e \mathcal{S} têm a mesma distribuição e as duas variáveis se relacionam pela equação

$$\mathcal{P} = b_1^{\mathcal{S}} b_2^{n-\mathcal{S}} = \left(\frac{b_1}{b_2}\right)^{\mathcal{S}} b_2^n. \quad (1)$$

No entanto, claramente seus valores esperados são diferentes. De fato, vamos calcular o valor médio do produto \mathcal{P} . Temos, por definição,

$$\begin{aligned} \mathbb{E}[\mathcal{P}] &= \sum_{i=0}^n b_1^i b_2^{n-i} \mathbb{P}(\mathcal{P} = b_1^i b_2^{n-i}) \\ &= \sum_{i=0}^n b_1^i b_2^{n-i} \binom{n}{i} p^i q^{n-i} \\ &= \sum_{i=0}^n \binom{n}{i} (pb_1)^i (qb_2)^{n-i} \end{aligned}$$

Pelo Teorema do Binômio de Newton temos que

$$\mathbb{E}[\mathcal{P}] = (pb_1 + qb_2)^n,$$

que difere de $\mathbb{E}[\mathcal{S}] = np$.

É fácil de ver que o k -ésimo momento de \mathcal{P} decorre de uma simples generalização do procedimento anterior. Temos

$$\begin{aligned} \mathbb{E}[\mathcal{P}^k] &= \sum_{i=0}^n (b_1^i b_2^{n-i})^k \mathbb{P}(\mathcal{P} = b_1^i b_2^{n-i}) \\ &= (pb_1^k + qb_2^k)^n. \end{aligned}$$

Em particular o segundo momento é dado por

$$\mathbb{E}[\mathcal{P}^2] = (pb_1^2 + qb_2^2)^n,$$

e a variância,

$$\text{VAR}(\mathcal{P}) = (pb_1^2 + qb_2^2)^n - (pb_1 + qb_2)^{2n}.$$

A Moda de \mathcal{P} é obtida observando-se a distribuição dos valores de \mathcal{P} e identificando o valor que tem maior probabilidade de ocorrer, assim como acontece com a variável \mathcal{S} . Como $\mathbb{M}(\mathcal{S}) = i^*$ então, pela equação (1) temos

$$\mathbb{M}(\mathcal{P}) = b_1^{i^*} b_2^{n-i^*}.$$

Lembrando, agora, que no final da Seção 3.1 utilizamos o TCL para aproximar a moda de \mathcal{S} por np , então podemos dizer que, para n grande,

$$\mathbb{M}(\mathcal{P}) = b_1^{np} b_2^{nq}.$$

Vamos resumir estas propriedades em uma proposição.

Proposição 4.1 *A média, a variância, o k -ésimo momento e a moda de \mathcal{P} são dados, respectivamente, por*

$$\begin{aligned} \mathbb{E}[\mathcal{P}] &= (pb_1 + qb_2)^n \\ \text{VAR}(\mathcal{P}) &= (pb_1^2 + qb_2^2)^n - (pb_1 + qb_2)^{2n} \\ \mathbb{E}[\mathcal{P}^k] &= (pb_1^k + qb_2^k)^n \\ \mathbb{M}(\mathcal{P}) &= b_1^{np} b_2^{nq}. \end{aligned}$$

Observação 4.1 *É interessante notar que*

$$\begin{aligned} \mathbb{E}[\mathcal{P}] &= \mathbb{E}[e^{\ln \mathcal{P}}] \\ \mathbb{M}(\mathcal{P}) &= e^{\mathbb{E}[\ln \mathcal{P}]} \end{aligned}$$

A primeira igualdade é óbvia. Já a segunda surge porque $\ln \mathcal{P}$ representa a soma de n variáveis, cada uma com média $p \ln b_1 + q \ln b_2$. Ou ainda visto de outra forma, pela equação (1) temos que

$$\ln \mathcal{P} = \mathcal{S} \ln \left(\frac{b_1}{b_2} \right) + \ln(b_2^n). \quad (2)$$

Calculando a média temos

$$\mathbb{E}[\ln \mathcal{P}] = \mathbb{E}[\mathcal{S}] \ln \left(\frac{b_1}{b_2} \right) + \ln(b_2^n),$$

o que acarreta

$$\mathbb{E}[\ln \mathcal{P}] = n(p \ln b_1 + q \ln b_2). \quad (3)$$

Desta expressão calculamos a exponencial e obtemos

$$\begin{aligned} e^{\mathbb{E}[\ln \mathcal{P}]} &= e^{n(p \ln b_1 + q \ln b_2)} \\ &= e^{\ln b_1^{np} + \ln b_2^{nq}} \\ &= b_1^{np} b_2^{nq}, \end{aligned}$$

o que corresponde à moda de \mathcal{P} .

Exemplo 4.1 Considere dois jogadores disputando uma sequência de n partidas do jogo **Par ou ímpar**. Existem apenas dois resultados possíveis: ganhar ou perder. A cada partida, o ganhador multiplica sua pontuação por 2 e o perdedor multiplica sua pontuação por $\frac{1}{2}$. A pontuação inicial de ambos é 1. Vamos colocar, então, $b_1 = 2$ e $b_2 = \frac{1}{2}$ e, supondo que a probabilidade de sair par ou sair ímpar é a mesma, então teremos $p = q = \frac{1}{2}$.

A pontuação do primeiro jogador depois de n partidas será o produto \mathcal{P} referente aos seus ganhos ou perdas nas n partidas.

- Após a 1ª partida, seu ganho será 2 ou $\frac{1}{2}$, havendo apenas uma chance de resultar em 2 e uma chance de resultar em $\frac{1}{2}$.
- Após a 2ª partida, seu ganho será 4, 1 ou $\frac{1}{4}$, havendo uma chance de resultar em 4, uma chance de resultar em $\frac{1}{4}$ e duas chances de resultar em 1.
- Após a 3ª partida, seu ganho será 8, 2, $\frac{1}{2}$ ou $\frac{1}{8}$, havendo uma chance de resultar em 8, uma chance de resultar em $\frac{1}{8}$, três chances de resultar em 2 e três chances de resultar em $\frac{1}{2}$. E assim por diante.

A Tabela 2 mostra as possibilidades de atingir os ganhos para o primeiro jogador até a 6ª partida. Analogamente ao Exemplo 2.6, essas possibilidades estão relacionadas a números binomiais.

Tabela 2: Número de possibilidades de atingir certos ganhos para o primeiro jogador.

| Partida | Ganho | | | | | | | | | | | | |
|----------------|----------------|----------------|----------------|---------------|---------------|---------------|----|----|----|---|----|----|----|
| | $\frac{1}{64}$ | $\frac{1}{32}$ | $\frac{1}{16}$ | $\frac{1}{8}$ | $\frac{1}{4}$ | $\frac{1}{2}$ | 1 | 2 | 4 | 8 | 16 | 32 | 64 |
| 0 (inicial) | | | | | | | 1 | | | | | | |
| 1 ^a | | | | | | 1 | | 1 | | | | | |
| 2 ^a | | | | | 1 | | 2 | | 1 | | | | |
| 3 ^a | | | | 1 | | 3 | | 3 | | 1 | | | |
| 4 ^a | | | 1 | | 4 | | 6 | | 4 | | 1 | | |
| 5 ^a | | 1 | | 5 | | 10 | | 10 | | 5 | | 1 | |
| 6 ^a | 1 | | 6 | | 15 | | 20 | | 15 | | 6 | | 1 |

O valor médio do produto fica

$$\begin{aligned}
 \mathbb{E}[\mathcal{P}] &= (p \cdot b_1 + q \cdot b_2)^n \\
 &= \left(\frac{1}{2} \cdot 2 + \frac{1}{2} \cdot \frac{1}{2} \right)^n \\
 &= \left(\frac{5}{4} \right)^n = e^{n \cdot \ln(\frac{5}{4})}.
 \end{aligned}$$

Já a moda (ou valor mais provável) é dado por

$$\begin{aligned}
 \mathbb{M}(\mathcal{P}) &= (b_1^p b_2^q)^n \\
 &= \left[2^{\frac{1}{2}} \left(\frac{1}{2} \right)^{\frac{1}{2}} \right]^n \\
 &= \left(1^{\frac{1}{2}} \right)^n \\
 &= 1.
 \end{aligned}$$

Em geral, o ganho \mathcal{P} é dado por

$$\mathcal{P} = 2^{2S-n} = 2^Z,$$

para Z como no Exemplo 3.1.

Observação 4.2 $\mathbb{E}[\mathcal{P}]$ cresce geometricamente com n sempre que $pb_1 + (1-p)b_2 > 1$, ou

seja, quando

$$p > \frac{1 - b_2}{b_1 - b_2}$$

Exemplo 4.2 Note que no caso em que $b_1 = 2$ e $b_2 = \frac{1}{2}$, como no Exemplo 4.1, basta que p seja maior do que $\frac{1}{3}$ para termos $\mathbb{E}[\mathcal{P}]$ com crescimento geométrico com n .

Quando observamos a Moda, no caso em que $b_1 = 2$ e $b_2 = \frac{1}{2}$, temos

$$\mathbb{M}(\mathcal{P}) = 2^{np} \left(\frac{1}{2}\right)^{n-p} = 2^{n(2p-1)}.$$

Assim a Moda será menor do que 1 no caso em que $p < \frac{1}{2}$.

Em vista do valor de p podemos classificar o comportamento do experimento em função de n crescente. Veja a Tabela 3.

Tabela 3: Comportamento do experimento em função de n grande.

| | $\mathbb{E}[\mathcal{P}]$ | $\mathbb{M}(\mathcal{P})$ |
|---------------------------------|---------------------------|---------------------------|
| $0 < p < \frac{1}{3}$ | decresc. geom. | decresc. geom. |
| $p = \frac{1}{3}$ | constante | decresc. geom. |
| $\frac{1}{3} < p < \frac{1}{2}$ | cresc. geom. | decresc. geom. |
| $p = \frac{1}{2}$ | cresc. geom. | constante |
| $\frac{1}{2} < p < 1$ | cresc. geom. | cresc. geom. |

Observe a região de valores para p entre $\frac{1}{3}$ e $\frac{1}{2}$ onde o comportamento de $\mathbb{E}[\mathcal{P}]$ e $\mathbb{M}(\mathcal{P})$ são díspares. Como $\mathbb{M}(\mathcal{P})$ representa o valor mais provável, uma pequena amostra do experimento possui grande chance de conter as sequências que decrescem geometricamente, porém à medida que a amostra cresce o efeito das sequências pouco prováveis mas com valores extremos se acentua e começa a interferir em $\mathbb{E}[\mathcal{P}]$.

4.1.1 Esperança condicional

Agora vamos definir uma sequência de variáveis aleatórias em Ω que representam os produtos parciais dos ω_i nos m primeiros ensaios. Para cada $m \in \{1, \dots, n\}$, seja $\mathcal{P}_m : \Omega \rightarrow \mathbb{R}$ a função dada por

$$\mathcal{P}_m(\omega) = \prod_{i=1}^m \omega_i,$$

para cada $\omega = (\omega_1, \dots, \omega_n)$. Se em $\omega_1 \dots \omega_m$ o valor de b_1 ocorrer i vezes e o valor b_2 ocorrer $m - i$ vezes, então

$$\mathcal{P}_m(\omega) = b_1^i b_2^{m-i}.$$

O evento $[\mathcal{P}_m = b_1^i b_2^{m-i}]$ significa o conjunto de todas as n -uplas em que b_1 ocorre i vezes nas primeiras m coordenadas. A distribuição de \mathcal{P}_m é dada por

$$\mathbb{P}(\mathcal{P}_m = b_1^i b_2^{m-i}) = \begin{cases} \binom{m}{i} p^i (1-p)^{m-i}, & \text{se } i \leq m \\ 0, & \text{se } i > m \end{cases}$$

para cada $i \in \{0, 1, \dots, n\}$. Note, portanto, que \mathcal{P}_m e \mathcal{S}_m (definido na Seção 3.1) têm a mesma distribuição. Além disso observe que $\mathcal{P}_n = \mathcal{P}$.

Estamos interessados no comportamento de \mathcal{P}_n condicionado ao que sabemos sobre \mathcal{P}_m , para $m < n$. Para isso vamos calcular

$$\mathbb{E}[\mathcal{P}_n | \mathcal{P}_m = b_1^j b_2^{m-j}] = \sum_{i=0}^n b_1^i b_2^{n-i} \mathbb{P}(\mathcal{P}_n = b_1^i b_2^{n-i} | \mathcal{P}_m = b_1^j b_2^{m-j})$$

Como \mathcal{P}_m e \mathcal{S}_m têm a mesma distribuição, então os eventos $[\mathcal{P}_m = b_1^j b_2^{m-j}]$ e $[\mathcal{S}_m = j]$ são iguais. Assim

$$\mathbb{E}[\mathcal{P}_n | \mathcal{P}_m = b_1^j b_2^{m-j}] = \sum_{i=0}^n b_1^i b_2^{n-i} \mathbb{P}(\mathcal{S}_n = i | \mathcal{S}_m = j)$$

Aplicando a Proposição 3.2 temos

$$\begin{aligned} \mathbb{E}[\mathcal{P}_n | \mathcal{P}_m = b_1^j b_2^{m-j}] &= \sum_{i=j}^{n-m+j} b_1^i b_2^{n-i} \binom{n-m}{i-j} p^{i-j} (1-p)^{n-m-i+j} \\ &= \sum_{i=0}^{n-m} b_1^{i+j} b_2^{n-i-j} \binom{n-m}{i} p^i (1-p)^{n-m-i} \\ &= b_1^j b_2^{m-j} \sum_{i=0}^{n-m} \binom{n-m}{i} (p b_1)^i [(1-p) b_2]^{n-m-i} \\ &= b_1^j b_2^{m-j} [p b_1 + (1-p) b_2]^{n-m}. \end{aligned}$$

Daí vemos que $\mathbb{E}[\mathcal{P}_n | \mathcal{P}_m = b_1^j b_2^{m-j}]$ sofre maior influência de $p b_1 + (1-p) b_2$, quando n cresce.

Para ilustrar a conclusão deste resultado, vamos fazer um exemplo numérico.

Exemplo 4.3 Fixe $b_1 = 2$, $b_2 = \frac{1}{2}$ e $p = \frac{1}{2}$.

$$\begin{aligned} \mathbb{E} \left[\mathcal{P}_n \mid \mathcal{P}_m = 2^j \left(\frac{1}{2} \right)^{m-j} \right] &= \mathbb{E}[\mathcal{P}_n \mid \mathcal{P}_m = 2^{2j-m}] \\ &= 2^{2j-m} \left(\frac{5}{4} \right)^{n-m}. \end{aligned}$$

Ocorre que, não importa o valor inicial ao qual \mathcal{P}_n esteja condicionado, o valor esperado cresce exponencialmente com n .

Com isso ilustramos o fato de sequências pouco prováveis influenciarem sobremaneira a média, e como o processo multiplicativo em si pode ocasionar valores extremos.

Observação 4.3 Note que a variável $\ln \mathcal{P}$ apresentada na Observação 4.1 e calculada a partir da equação (2) tem distribuição normal com média dada pela equação (3) e variância dada por

$$\text{VAR}(\ln \mathcal{P}) = np(1-p) \left[\ln \left(\frac{b_1}{b_2} \right) \right]^2.$$

Além disso, como $\ln \mathcal{P}$ resulta de um passeio aleatório, então podemos utilizar o TCL para inferir sobre o comportamento assintótico de $\ln \mathcal{P}$ para n grande. Porém, a informação obtida para $\ln \mathcal{P}$ a partir do TCL não contém os efeitos que sequências pouco prováveis com valores extremos têm sobre o produto \mathcal{P} . Para mais detalhes sobre o efeito dos eventos raros sobre os momentos de \mathcal{P} , ver [12].

4.2 Processos multiplicativos multinomiais

Finalizando esta seção, vamos traçar agora as linhas gerais dos processos multiplicativos multinomiais, comentando sobre o comportamento do produto \mathcal{P} frente sua média e moda.

4.2.1 Processo multiplicativo trinomial

No caso de um processo multiplicativo trinomial, fixamos três números distintos $m_1, m_2, m_3 > 0$, bem como 3 números reais $0 < p_1, p_2, p_3 < 1$, tais que $p_1 + p_2 + p_3 = 1$. No modelo, m_i ocorre com probabilidade p_i , para $i = 1, 2, 3$.

Fixe $n \in \mathbb{N}$. O espaço amostral será o conjunto Ω de todas a n -uplas $(\omega_1, \dots, \omega_n)$ em que $\omega_i \in \{m_1, m_2, m_3\}$. O espaço de eventos \mathcal{F} será dado pela coleção de todos os subconjuntos de Ω , isto é, $\mathcal{F} = 2^\Omega$. A medida de probabilidade \mathbb{P} sobre \mathcal{F} será dada da seguinte forma. Fixe $\omega = (\omega_1, \dots, \omega_n) \in \Omega$ e suponha que m_1 ocorra i_1 vezes, m_2 , i_2 vezes e m_3 , i_3 vezes. Então

$$\mathbb{P}(\{\omega\}) = p_1^{i_1} p_2^{i_2} p_3^{i_3},$$

onde $i_1 + i_2 + i_3 = n$.

De forma análoga ao processo multiplicativo binomial, vamos definir a variável \mathcal{P} que retorna o produto aleatório de m_1 , m_2 e m_3 . Se $\omega = (\omega_1, \dots, \omega_n)$ então

$$\mathcal{P}(\omega) = \omega_1 \omega_2 \cdots \omega_n.$$

O evento $[\mathcal{P} = m_1^{i_1} m_2^{i_2} m_3^{i_3}]$ significa o conjunto de todas as n -uplas em que m_1 ocorre i_1 vezes, m_2 , i_2 vezes e m_3 , i_3 vezes. No total há $\frac{n!}{i_1! i_2! i_3!}$ n -uplas nesse evento. Portanto

$$\mathbb{P}(\mathcal{P} = m_1^{i_1} m_2^{i_2} m_3^{i_3}) = \frac{n!}{i_1! i_2! i_3!} p_1^{i_1} p_2^{i_2} p_3^{i_3}.$$

Para calcular o valor médio do produto, $\mathbb{E}[\mathcal{P}]$, calculamos a média sobre todos resultados possíveis:

$$\begin{aligned} \mathbb{E}[\mathcal{P}] &= \sum_{i_1=0}^n \sum_{i_2=0}^{n-i_1} m_1^{i_1} m_2^{i_2} m_3^{n-i_1-i_2} \mathbb{P}(\mathcal{P} = m_1^{i_1} m_2^{i_2} m_3^{n-i_1-i_2}) \\ &= \sum_{i_1=0}^n \sum_{i_2=0}^{n-i_1} \frac{n!}{i_1! i_2! (n-i_1-i_2)!} p_1^{i_1} p_2^{i_2} p_3^{n-i_1-i_2} m_1^{i_1} m_2^{i_2} m_3^{n-i_1-i_2} \\ &= \sum_{i_1=0}^n \sum_{i_2=0}^{n-i_1} \frac{n!}{i_1! i_2! (n-i_1-i_2)!} (p_1 m_1)^{i_1} (p_2 m_2)^{i_2} (p_3 m_3)^{n-i_1-i_2}, \end{aligned}$$

e pelo Teorema Multinomial (Teorema 2.1) obtemos

$$\mathbb{E}[\mathcal{P}] = (p_1 m_1 + p_2 m_2 + p_3 m_3)^n.$$

Para calcular a moda de \mathcal{P} (ou o valor mais provável do produto de n fatores), observamos que \mathcal{P} pode ser escrito como

$$\mathcal{P} = m_1^{\tilde{\mathcal{S}}_1} m_2^{\tilde{\mathcal{S}}_2} m_3^{\tilde{\mathcal{S}}_3}.$$

Sendo o valor mais frequente de $\tilde{\mathcal{S}}_j$ igual a np_j , então

$$\mathbb{M}(\mathcal{P}) = m_1^{np_1} m_2^{np_2} m_3^{np_3} = (m_1^{p_1} m_2^{p_2} m_3^{p_3})^n.$$

Exemplo 4.4 *Vamos considerar a seguinte situação: dois jogadores estão disputando uma sequência de partidas do jogo **Pedra, Papel e Tesoura**. Existem 3 resultados possíveis: ganhar, empatar ou perder. O jogador multiplica sua pontuação por 2 quando ganha, multiplica por $\frac{1}{2}$ quando perde e, em caso de empate, multiplica sua pontuação por 1. Considere $m_1 = 2$, $m_2 = 1$ e $m_3 = \frac{1}{2}$. Supondo que as probabilidades de ganhar, perder ou empatar são iguais, então $p_1 = p_2 = p_3 = \frac{1}{3}$.*

O valor médio do produto será

$$\begin{aligned}\mathbb{E}(\mathcal{P}) &= (p_1 m_1 + p_2 m_2 + p_3 m_3)^n \\ &= \left(\frac{1}{3}2 + \frac{1}{3}1 + \frac{1}{3}\frac{1}{2}\right)^n \\ &= \left(\frac{7}{6}\right)^n = e^{n \ln(\frac{7}{6})},\end{aligned}$$

que cresce geometricamente com n . O valor mais provável será

$$\begin{aligned}\mathbb{M}(\mathcal{P}) &= (m_1^{p_1} m_2^{p_2} m_3^{p_3})^n \\ &= \left[2^{\frac{1}{3}} 1^{\frac{1}{3}} \left(\frac{1}{2}\right)^{\frac{1}{3}}\right]^n \\ &= 1.\end{aligned}$$

Vemos, assim, que os eventos pouco prováveis afetam muito a média.

4.2.2 Processos multiplicativos multinomiais

Neste caso generalizado, fixe k números distintos $m_1, m_2, \dots, m_k > 0$, bem como k números reais $0 < p_1, \dots, p_k < 1$, tais que $\sum_{j=1}^k p_j = 1$. Neste modelo, m_1 ocorre com probabilidade p_1 , m_2 , com probabilidade p_2 e assim por diante. O espaço amostral será o conjunto Ω de todas a n -uplas $(\omega_1, \dots, \omega_n)$ em que $\omega_i \in \{m_1, m_2, \dots, m_k\}$. A medida de probabilidade \mathbb{P} será dada da seguinte forma. Fixe $\omega = (\omega_1, \dots, \omega_n) \in \Omega$ e suponha que m_1 ocorra i_1 vezes, m_2 ocorra i_2 vezes, e assim sucessivamente. Então

$$\mathbb{P}(\{\omega\}) = p_1^{i_1} p_2^{i_2} \cdots p_k^{i_k}.$$

Definimos a variável aleatória \mathcal{P} por

$$\mathcal{P}(\omega) = \omega_1 \dots \omega_n,$$

sendo que o evento $[\mathcal{P} = m_1^{i_1} \dots m_k^{i_k}]$, que reúne todas as n -uplas em que m_1 ocorre i_1 vezes, m_2 ocorre i_2 vezes, e assim por diante, tem probabilidade

$$\mathbb{P}(\mathcal{P} = m_1^{i_1} \dots m_k^{i_k}) = \frac{n!}{i_1! i_2! \dots i_k!} p_1^{i_1} p_2^{i_2} \dots p_k^{i_k}.$$

O valor médio do produto \mathcal{P} é

$$\mathbb{E}(\mathcal{P}) = (p_1 m_1 + p_2 m_2 + \dots + p_k m_k)^n$$

e o valor mais provável é

$$\mathbb{M}(\mathcal{P}) = m_1^{np_1} m_2^{np_2} \dots m_k^{np_k} = (m_1^{p_1} m_2^{p_2} \dots m_k^{p_k})^n.$$

5 Proposta de Aplicação em Sala de Aula

Nessa seção apresentamos uma proposta de atividade para ser desenvolvida em sala de aula a partir do 7º ano do Ensino Fundamental, pois os alunos deverão ter uma prévia noção de probabilidade. A atividade terá como base o jogo do **Par ou ímpar** apresentado na referência [1]. O sistema de pontuação terá um processo aditivo e um processo multiplicativo para que os alunos possam observar principalmente o comportamento extremo que o processo multiplicativo pode ter, atingindo valores extremamente altos (ou extremamente baixos).

5.1 Jogo do par ou ímpar

Essa atividade será dividida em algumas etapas e em cada etapa faremos sugestões de questionários que poderão ser aplicados aos alunos complementando a atividade.

5.1.1 1ª etapa (processo aditivo):

Os alunos serão informados sobre as regras do jogo e em seguida responderão ao questionário inicial (opcional).

Regras do jogo:

- A turma deverá ser dividida em duplas para a realização das atividades.

| Jogo do Par ou Ímpar | | | | | | | | | | | | | |
|----------------------|----------|---|---|---|---|---|---|---|---|---|----|----|----|
| Jogadores | Partidas | | | | | | | | | | | | |
| | Início | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| | 10 | | | | | | | | | | | | |
| | 10 | | | | | | | | | | | | |

| Jogo do Par ou Ímpar | | | | | | | | | | | | | |
|----------------------|----------|----|----|----|----|----|----|----|----|----|----|----|----|
| Jogadores | Partidas | | | | | | | | | | | | |
| | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 |
| | | | | | | | | | | | | | |
| | | | | | | | | | | | | | |

Figura 1: Tabela para anotações da pontuação de cada jogador

- Os jogadores deverão escolher entre par ou ímpar e a escolha não poderá ser alterada durante a realização das partidas do jogo.
- O jogo terá um total de 25 partidas (ou mais). Em cada partida, os jogadores deverão indicar, com os dedos, os números de 1 a 5. Se a soma dos números indicados pelos jogadores for par, ganhará a partida o jogador que escolheu par, caso contrário ganhará o jogador que escolheu ímpar.
- Cada jogador terá uma pontuação inicial, por exemplo igual a 10. Em caso de vitória, o jogador somará 1 ponto à sua pontuação. Em caso de derrota, o jogador perderá 1 ponto de sua pontuação. Utilizar as tabelas da Figura 1 para anotar a pontuação no decorrer do jogo.
- O vencedor do jogo será o jogador que obtiver a maior pontuação ao final das 25 partidas.

Sugestões de perguntas para o 1º questionário:

1. Alguma vez você já jogou par ou ímpar onde não era permitido indicar o número 0?
2. Você acha que o fato de ter excluído o “zero dedos” das opções muda a probabilidade de sair par ou sair ímpar?

3. Se você acha que as probabilidades de sair par ou sair ímpar são diferentes, então qual das opções tem maior probabilidade?

5.1.2 2ª etapa (processo multiplicativo):

Para a segunda etapa, haverá apenas uma alteração nas regras do jogo (sobre a pontuação).

Regra específica da 2ª etapa:

- Cada jogador terá uma pontuação inicial, por exemplo igual a 10. Em caso de vitória, o jogador multiplicará sua pontuação por 2. Em caso de derrota, o jogador multiplicará sua pontuação por um meio.

Sugestões de perguntas para o 2º questionário

1. Você acha que esse sistema de pontuação apresentado (dobrar ou dividir por 2) é bom para os jogadores?
2. Se você pudesse escolher entre o sistema de pontuação apresentado na 1ª etapa ou o sistema apresentado na 2ª etapa, qual você escolheria?

5.1.3 Etapa final

Depois que os alunos concluírem as duas etapas do jogo do par ou ímpar será realizado o questionário final, caso o professor tenha optado pela realização dos questionários. Em seguida o professor poderá realizar os cálculos das probabilidades de sair par ou de sair ímpar juntamente com os alunos.

Sugestões de perguntas para o questionário final:

1. O vencedor do jogo escolheu par ou ímpar?

2. Após jogar as 25 partidas, você acha que o fato de ter excluído o “zero dedos” das opções mudou a probabilidade de sair par ou sair ímpar? Você conseguiria calcular essas probabilidades?
3. Depois da experiência, você acha que esse sistema de pontuação apresentado (dobrar ou dividir por dois) é bom para os jogadores? Por quê?

5.1.4 Cálculo da probabilidade do Jogo par ou ímpar excluindo “zero dedos”

Resultados possíveis: $\mathcal{C} = \{(1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (2, 1), (2, 2), (2, 3), (2, 4), (2, 5), (3, 1), (3, 2), (3, 3), (3, 4), (3, 5), (4, 1), (4, 2), (4, 3), (4, 4), (4, 5), (5, 1), (5, 2), (5, 3), (5, 4), (5, 5)\}$.

Podemos verificar que dos 25 resultados possíveis 13 são pares e 12 são ímpares. Ou seja, a probabilidade de sair par é de $\frac{13}{25}$ ou 0,52 e a probabilidade de sair ímpar é de $\frac{12}{25}$ ou 0,48.

Suponha que estejamos analisando o ganho do jogador que escolheu *par* durante todo o jogo. Após n partidas, na etapa 1, seu ganho será $\mathcal{Z} + 10$, sendo o valor médio igual à moda e ambos iguais a

$$\mathbb{E}[\mathcal{Z} + 10] = n(0,52 - 0,48) + 10 = 0,04n + 10.$$

Na etapa 2, o ganho será $10\mathcal{P}$. O valor médio será

$$\begin{aligned} \mathbb{E}[10\mathcal{P}] &= 10 \left(0,52 \times 2 + 0,48 \times \frac{1}{2} \right)^n \\ &= 10 (1,28)^n \end{aligned}$$

e a moda,

$$\begin{aligned} \mathbb{M}(10\mathcal{P}) &= 10 \left[2^{0,52} \left(\frac{1}{2} \right)^{0,48} \right]^n \\ &= 10 \cdot 2^{0,04n}. \end{aligned}$$

O sistema de pontuação na 2^a etapa é muito mais vantajoso para o jogador que escolheu par, pois os ganhos tendem a ser muito altos em média, apesar do valor mais frequente ser decrescente com n .

6 Conclusão

Nesse trabalho discutimos o comportamento assintótico dos processos aditivos aleatórios e dos processos multiplicativos aleatórios. Sobre os processos aditivos mostramos a relação entre passeio aleatório e o Teorema Central do Limite. Sobre os processos multiplicativos, especificamente os binomiais e multinomiais, desenvolvemos algumas propriedades estatísticas importantes como a média, a variância, o k -ésimo momento e a moda, enfatizando a influência de eventos extremos.

Elaboramos, também, uma atividade para ser trabalhada em sala de aula, que contém um processo aditivo e um processo multiplicativo. A atividade foi pensada em ser aplicada a partir do 7º ano do ensino fundamental e exige apenas conhecimentos básicos como adição de números inteiros, multiplicação de números racionais e cálculo de probabilidade. A atividade tem como base o jogo do par ou ímpar, mas com dois diferenciais: (1) os jogadores não poderiam jogar com o “zero dedos” a fim de que as probabilidades de vitória fossem diferentes de 0,5; (2) haviam dois tipos de pontuações, um envolvendo um processo aditivo e outro envolvendo um processo multiplicativo. A atividade, que deveria ser realizada de forma presencial, não pôde ser executada devido à ocorrência da pandemia de Covid-19, obrigando as aulas do ano letivo de 2020 serem remotas.

7 Agradecimentos

Agradeço a Deus, pela vida, saúde e ter me abençoado para chegar nesse momento. Agradeço minha família pelo incentivo, em especial minha mãe Helia Maria pelo amor, educação e apoio incondicional. Agradeço ao meu orientador, Professor Telles Timóteo da Silva, a quem admiro pela extrema competência. Muito obrigada pela atenção, paciência e dedicação na execução desse trabalho. Agradeço aos professores da UFSJ - CAP, pelos valiosos ensinamentos. Agradeço aos meus colegas de turma, sem eles essa caminhada seria muito mais difícil.

Referências

- [1] BORGES, Pablo dos Santos. **Jogo do par ou ímpar**. 2014. 88f. Trabalho de Conclusão de Curso (Mestrado Profissional em Matemática em Rede Nacional) - Universidade Federal de Goiás - UFG, Goiania, 2014.
- [2] BUSSAB, Wilton de O, e MORETTIN, Pedro A., *Estatística Básica*. 6^a edição. São Paulo: Saraiva, 2009.
- [3] COLUCCINI, Cláudia Mara Cardoso Reis. **A Matemática das Pesquisas por Amostragem: um olhar sobre as pesquisas de intenção de votos**. 2017. 41f. Trabalho de Conclusão de Curso (Mestrado Profissional em Matemática em Rede Nacional) - Universidade Federal de São João del Rei - UFSJ / Campus Alto do Paraopeba, Ouro Branco, 2017.
- [4] DANTE, Luiz Roberto. *Matemática Contexto e Aplicações*. São Paulo: Editora Ática, 2012.
- [5] DE JESUS, Eliane Alves. **Sistemas Dinâmicos Discretos**. 2016. 78f. Trabalho de Conclusão de Curso (Mestrado Profissional em Matemática em Rede Nacional) - Universidade Federal de São João del Rei - UFSJ / Campus Alto do Paraopeba, Ouro Branco, 2016.
- [6] JAMES, Barry R. *Probabilidade: um curso em nível intermediário*. 2a.ed. Rio de Janeiro: Instituto de Matemática Pura e Aplicada, 1996.
- [7] KALLENBERG, Olav. *Foundations of Modern Probability*. 2nd.ed. New York: Springer, 2002.
- [8] MAGALHÃES, Marcos Nascimento. *Probabilidade e Variáveis Aleatórias*. 2^a edição. São Paulo: Editora da Universidade de São Paulo, 2006.
- [9] MANRUBIA, Susanna C. e ZANETTE, Damian H. Stochastic multiplicative processes with reset events. **Physical Review E** **59** (5) 1999.

- [10] MORETTIN, Luiz Gonzaga, *Estatística Básica, Volume 1, Probabilidade*. 7ª edição. São Paulo: Pearson Makron Books, 1999.
- [11] OLIVEIRA, Jaqueline Patrícia Duarte. **Abordando as distribuições binomial e multinomial: propriedades e um exemplo de processo estocástico**. 2015. 43f. Trabalho de Conclusão de Curso (Mestrado Profissional em Matemática em Rede Nacional) - Universidade Federal de São João del Rei - UFSJ / Campus Alto do Paraopeba, Ouro Branco, 2015.
- [12] REDNER, S. Random multiplicative processes: an elementary tutorial. **Am. J. Phys** **58** (3) 1990.
- [13] TEODORO, Marcos Paulo. **Passeios aleatórios combinados à sequência de Gibonacci**. 2017. 42f. Trabalho de Conclusão de Curso (Mestrado Profissional em Matemática em Rede Nacional) - Universidade Federal de São João del Rei - UFSJ / Campus Alto do Paraopeba, Ouro Branco, 2017.
- [14] TOME, Tânia, e OLIVEIRA, Mário José. *Dinâmica Estocástica e Irreversibilidade*. São Paulo: Editora da Universidade de São Paulo, 2001.