2017: DISSERTAÇÃO DE MESTRADO
Mestrado Profissional em Matemática - PROFMAT, Universidade Federal de São João Del-Rei - UFSJ
Sociedade Brasileira de Matemática - SBM

Probabilidades no futebol

Cézar Augusto Garófalo Carneiro [1]
Humberto Cesar Fernandes Lemos[2]

Resumo: Este trabalho apresenta a Teoria das Probabilidades aplicada ao futebol através da modelagem matemática. Inicialmente, estudaremos os conceitos fundamentais de probabilidade, as suas definições (clássica e frequentista) e diferenças, a linguagem padrão utilizada, os axiomas importantes e suas propriedades necessárias para a aplicação em diversas situações. Em seguida, retrataremos a ideia geral da Lei dos Grandes Números e os tipos de convergências que discriminam suas particularidades. Enfim, detalharemos dois modelos matemáticos para o cálculo de diversas probabilidades que envolvem os times em um campeonato de futebol e apresentaremos algumas propostas de atividades em sala de aula envolvendo a Lei dos Grandes Números e a probabilidade no futebol visando contribuir com os professores na composição do aprendizado dos seus alunos.

Palavras-chave: Probabilidade, Teoria das Probabilidades, Lei dos Grandes Números, Probabilidade no Futebol.

Abstract: This work introduces the Theory of Probability applied to soccer through mathematical modeling. Initially, we are going to study the fundamental concepts of probability, its classical and frequentist definitions and differences, the standard language utilized, relevant axioms and their necessary features to be applied to different situations. Following to that, we are going to depict the general idea of the Law of Large Numbers as well as the various types of convergences which discriminate their particularity. Finally, we are going to detail two mathematical models to calculate the diverse probabilities which involve the teams participating in a soccer championship. With the aim of contributing to teachers composing their students' learning, we are also going to introduce a few in-class activity proposals containing the Law of Large Numbers and the probabilities in soccer.

Keywords: Probability, Theory of Probability, Law of Large Numbers, Probability in Soccer.

[1]Aluno de Mestrado do PROFMAT, Turma 2014, Universidade Federal de São João Del-Rei - UFSJ, cezargarofalo@gmail.com

[2]Professor orientador, Departamento de Física e Matemática - DEFIM, UFSJ - CAP, humbertolemos@ufsj.edu.br

1 Introdução

Matemática e futebol, uma combinação perfeita para quem é apaixonado por números e um bom jogo no fim de semana. Algumas pessoas acham que não há uma relação entre os dois, mas se o futebol é tão imprevisível como dizem ser, como podemos usar a Matemática para mensurar e analisar improváveis resultados?

A Matemática oferece a Probabilidade e a Estatística como ferramentas importantes no futebol. Quando assistimos uma partida, várias informações como número de faltas, posse de bola, chutes a gol, escanteios e outros são fornecidos, estatisticamente, retratando as características do jogo. Nos programas esportivos da TV e jornais, as probabilidades dos times de serem campeões ou serem rebaixados são calculadas e informadas gerando algum alívio ou indignação nos torcedores, e sendo um motivo a mais para uma discussão sobre o campeonato. Na internet, sites esportivos proporcionam aos internautas ganhos financeiros com apostas em possíveis resultados dos jogos que ainda serão disputados e demonstram a probabilidade para cada resultado estimando o valor das apostas e dos prêmios. Eis alguns exemplos de situações que comprovam o quanto a Matemática está presente no futebol.

Este trabalho procura explicar o uso da Matemática no futebol através da Teoria das Probabilidades a fim de determinar as chances de um ou mais times vencerem ou se livrarem de um rebaixamento durante o campeonato. Os procedimentos para o cálculo destas probabilidades são bem interessantes, despertando ideias para que o leitor possa criar o seu próprio modelo ou, simplesmente, satisfazer a sua curiosidade.

Na seção 2, apresentaremos um estudo sobre a Teoria das Probabilidades, destacando seus conceitos básicos e definindo a sua linguagem padrão, trataremos das definições das probabilidades clássica e frequentista distinguindo suas diferenças e aplicações, destacaremos as condições da sigma-álgebra, as principais e importantes propriedades, a probabilidade condicional e a independência de eventos, exemplificando todos eles para um melhor entendimento do leitor. Na seção 3, apresentaremos a Lei dos Grandes Números, o conceito geral detalhando a ideia em que elas se baseiam, os tipos de convergências para definirmos a Lei Forte dos Grandes Números de Tchebychev e a lei Fraca dos Grandes Números de Kolmogorov. Na seção 4, detalharemos dois modelos matemáticos para o cálculo de probabilidades no futebol. O primeiro modelo, muito utilizado nos programas esportivos em Minas Gerais, foi criado por um grupo do Departamento de Matemática da UFMG que se preocupa em divulgar este modelo como sugestão para a criação de novos modelos para o futebol ou outros esportes. O segundo modelo, pela simplicidade, poderá ser utilizado como atividade em sala de aula visando tabalhar com modelagem matemática e probabilidade com os alunos. Na seção 5, apresentaremos três propostas de atividades em sala de aula, uma envolvendo o cálculo de probabilidades no futebol e modelagem matemática e outras duas trabalhando o conceito de probabilidade e a ideia geral da Lei dos Grandes Números.

2 Noções de probabilidade

2.1 Experimentos aleatórios

Existem experimentos que, apesar de serem repetidos muitas vezes e sob as mesmas condições, não apresentam resultados idênticos, por exemplo, no lançamento de uma moeda perfeita. Este experimento é aplicado em uma das regras do Futebol [1] (regra nº8) para garantir que as duas equipes tenham a mesma chance na escolha do lado do campo para início da partida, pois só é possível obter um dos dois valores: coroa ou cara. Esse tipo de experimento chamamos de experimento aleatório.

Definição 2.1 *Experimentos aleatórios são experimentos que ao serem repetidos várias vezes, sob condições semelhantes, não apresentam o mesmo resultado (resultados imprevisíveis).*

Exemplo 2.1 *Podemos observar alguns experimentos abaixo que mesmo repetidos várias vezes não há como prever o resultado obtido em um determinado momento.*

- *Lançamento de um dado.*
- *Retirar uma carta do baralho.*
- *Sortear um número no jogo de roleta.*

A área da Matemática que apura a chance de ocorrer um determinado resultado em um experimento aleatório é denominada Teoria das Probabilidades. Alguns registros sobre essa teoria aparecem na obra sobre jogos de azar do italiano Girolamo Cardano (1501-1576). Aproximadamente cem anos depois, Blaise Pascal (1623-1662) desenvolveu a Teoria das Probabilidades baseada numa discussão de problemas relacionados a jogos (cartas) com Pierre Fermat (1601-1665) e alguns tópicos relacionados à probabilidade foram encontrados na sua obra sobre o triângulo aritmético. Em 1713, surge o primeiro artigo completo sobre a Teoria das Probabilidades na obra *Ars Conjectandi* (Arte de conjecturar)([12]) de Jakob Bernoulli (1654-1705). A obra *The Doctrine of Chaucesme* lançada, em 1718, pelo Abraham de Moivre (1667-1754) apresenta o primeiro conceito do Teorema Central do Limite. Em 1812, Pierre Simon Laplace (1749-1827) enunciou a definição clássica de probabilidade na majestosa obra *Theorie Analytique des Probabilités*. A partir de então, outros matemáticos contribuíram no desenvolvimento da Teoria das Probabilidades sendo aplicada nas áreas da economia, saúde, biologia, política, etc.

Destaque para os matemáticos russos Pafnuti Tchebychev (1821-1894), Andrei Andreyevich Markov (1856-1922) e, principalmente, Andrei Nikolaevich Kolmogorov (1903-1987) que, em 1933, desenvolveu o conceito axiomático de probabilidade na sua publicação *Foundations of The Theory of Probability*.

2.2 A linguagem das probabilidades

A aplicação da Teoria das Probabilidades em várias áreas do conhecimento contribuiu para o desenvolvimento de modelos matemáticos visando estudar diversos experimentos aleatórios. Estes modelos variam de acordo com o grau de complexidade de cada experimento, mas possuem conceitos básicos comuns. Destacaremos primeiramente a linguagem das probabilidades, que é bem peculiar.

Em nosso modelo matemático, consideraremos cada possível resultado de um experimento aleatório qualquer como ponto amostral. O conjunto de todos os pontos amostrais será denominado espaço amostral.

Definição 2.2 *Espaço amostral é o conjunto de todos os possíveis resultados de um experimento aleatório e será representado por* Ω*.*

Os elementos do espaço amostral serão denominados eventos simples ou pontos amostrais.

Exemplo 2.2 *No lançamento de um dado de 6 faces, podemos obter os seguintes resultados* 1*,* 2*,* 3*,* 4*,* 5 *e* 6*, logo, o espaço amostral deste experimento é* $\Omega = \{1, 2, 3, 4, 5, 6\}$.

Exemplo 2.3 *No lançamento de uma moeda, sucessivamente, até que o resultado seja coroa pela primeira vez, todos os resultados possíveis podem ser relacionados biunivocamente com o conjunto dos números naturais. Note que o espaço amostral é infinito e enumerável, ou seja, seus elementos podem ser dipostos em uma sequência.*

No exemplo 2.2 o espaço amostral é finito e no exemplo 2.3 o espaço amostral é infinito, neste caso, enumerável. Há também espaços amostrais infinitos e não enumeráveis.

Definição 2.3 *Denominaremos de evento todo subconjunto de elementos pertencente à* σ*-álgebra* $\mathbb{A}$ *de um espaço amostral* Ω *de um experimento aleatório.*

Observe que o conceito de σ-álgebra será tratado posteriormente na seção 2.4.

Exemplo 2.4 *Considere o espaço amostral* $\Omega = \{1, 2, 3, 4, 5, 6\}$ *no lançamento de um dado de 6 faces, citado no exemplo* 2.2*, podemos considerar alguns possíveis eventos:*

- $A = \{\textit{números pares}\} = \{2, 4, 6\}$.
- $B = \{\textit{números primos}\} = \{2, 3, 5\}$.
- $C = \{\textit{divisores de } 60\} = \{1, 2, 3, 4, 5, 6\}$.
- $D = \{\textit{múltiplos de } 8\} = \emptyset$.

Todo evento que contém apenas um ponto amostral (conjunto unitário) será denominado evento simples e um determinado evento A ocorre se pelo menos um evento simples pertencente a A ocorrer no final de um experimento aleatório.

No exemplo 2.4 consideremos o evento $A = \{\text{números pares}\}$. Tem-se: $A = \{2, 4, 6\}$. O evento A ocorrerá se um dos três eventos simples 2, 4 ou 6 ocorrer. Da mesma forma, o evento $B = \{2, 3, 5\}$ ocorrerá se um dos eventos simples 2, 3 ou 5 ocorrer. Observe que todos os elementos do evento C pertencem ao espaço amostral Ω e será chamado evento certo. Já o evento D que não possui elementos será chamado evento impossível. Estes são os eventos triviais na Teoria das Probabilidades.

Por se tratar de conjuntos, citaremos a Álgebra dos conjuntos e algumas de suas propriedades que serão essenciais para o aprofundamento da Teoria das Probabilidades.

Considere dois eventos A e B de um espaço amostral Ω.

A união dos eventos A e B é o evento que ocorre se pelo menos um deles ocorre e, denotaremos por $A \cup B$.

A interseção dos eventos A e B é o evento que ocorre quando ambos ocorrem e denotaremos por $A \cap B$.

O complementar do evento A é o evento que ocorre quando A não ocorre e denotaremos por A^c.

Podemos definir as operações de união e interseção para um grupo enumerável de eventos.

A união dos eventos A_i, para $i = 1, 2, ...$, é o evento que ocorre se pelo menos um dos eventos A_i ocorre e denotaremos por $\bigcup_{i=1}^{\infty} A_i$.

A interseção dos eventos A_i, para $i = 1, 2, ...$, é o evento que ocorre se todos os eventos A_i ocorrerem e denotaremos por $\bigcap_{i=1}^{\infty} A_i$.

Exemplo 2.5 *Uma urna contém quinze bolas numeradas de 1 a 15. Uma bola é retirada ao acaso. Considere os seguintes eventos:*
$A = \{\textit{retirar uma bola de número par}\} = \{2, 4, 6, 8, 10, 12, 14\}$ *e*
$B = \{\textit{retirar uma bola de número múltiplo de } 3\} = \{3, 6, 9, 12, 15\}$.
Aplicando as operações da união, interseção e complementar temos:
$A \cup B = \{\textit{retirar uma bola de número par ou múltiplo de } 3\} = \{2, 3, 4, 6, 8, 9, 10, 12, 14, 15\}$,
$A \cap B = \{\textit{retirar uma bola de número par e múltiplo de } 3\} = \{6, 12\}$ *e*
$A^c = \{\textit{retirar uma bola de número não par}\} = \{1, 3, 5, 7, 9, 11, 13, 15\}$.

Quando todos os elementos do evento A pertencem ao evento B, ou seja, $A \subset B$, podemos afirmar que os eventos A e B ocorrem simultaneamente se o evento A ocorrer. Quando $A \subset B$ e $B \subset A$, os dois eventos ocorrem simultaneamente se ocorrer o evento A ou o evento B. Neste caso, consideramos que os eventos são iguais. Quando dois eventos A e B apresentam uma interseção vazia, ou seja, $A \cap B = \emptyset$, podemos afirmar que os eventos A e B não podem ocorrer simultaneamente. E, nesta situação, dizemos que os eventos são **mutuamente exclusivos**.

Existem algumas propriedades importantes envolvendo as operaçãos apresentadas, são elas:

p.1) $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$;
p.2) $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$;
p.3) $(A \cap B)^c = A^c \cup B^c$;
p.4) $(A \cup B)^c = A^c \cap B^c$.

2.3 Definições de probabilidade

Os jogos que envolviam dados já existiam há milhares de anos antes de Cristo, mas estudos sobre as chances de ocorrerem tais resultados não foram encontrados neste período, e sim, na Idade Média. Galileu Galilei definiu o conceito de eventos igualmente possíveis. Este conceito será utilizado na definição clássica de Probabilidade.

Definição 2.4 (Eventos equiprováveis a piori) *Dado um espaço amostral Ω com $n(\Omega)$ eventos simples e um evento A de Ω contendo $n(A)$ eventos simples, e considerando que todos os eventos simples de Ω são igualmente possíveis, a definição clássica de probabilidade deste evento A ocorrer é dada por:*

$$P(A) = \frac{n(A)}{n(\Omega)}.$$

Exemplo 2.6 *Um dado ideal é lançado sobre uma mesa. O espaço amostral Ω é composto por 6 eventos simples $\Omega = \{1, 2, 3, 4, 5, 6\}$ e cada um deles possui a probabilidade de $1/6$, todos igualmente possíveis. Seja um evento $A = \{obter\ um\ número\ ímpar\}$ composto por 3 eventos simples $A = \{1, 3, 5\}$, a probabilidade do evento evento A ocorrer é*

$$P(A) = \frac{n(A)}{n(\Omega)} = \frac{3}{6} = \frac{1}{2}.$$

A definição 2.4 é aplicada quando atribuímos a mesma chance de ocorrência a todos os eventos simples de um determinado experimento. Mas quando o número de eventos simples que compõem o espaço amostral for infinito? Nesta situação, uma maneira de determinar a probabilidade da ocorrência de um determinado evento A deve-se repetir n vezes o experimento aleatório registrando o número de vezes que o evento A ocorreu. Quanto maior n, a probabilidade deste evento A tende a uma constante p. A probabilidade frequentista será melhor abordada na seção 3 mas segue abaixo a sua definição.

Definição 2.5 (Interpretação frequentista de probabilidade) *Considere $f(A)$ a frequência obtida na ocorrência do evento A em n repetições de um experimento aleatório. A probabilidade frequentista do evento A é a razão dada por:*

$$P(A) = \lim_{n \to \infty} \frac{f(A)}{n}.$$

Exemplo 2.7 *Considere os dados da tabela 1. O número total de nascimentos durante este período é 88723 que representa o espaço amostral. O evento $A = \{nascimento\ de\ mulheres\}$ ocorreu 42591 vezes, ou seja, este número representa a frequência do evento A.*

A probabilidade do evento A , em um ano, foi $P(A) = \dfrac{42591}{88273} \approx 0,4825.$

A partir destas definições, apresentaremos os axiomas e principais propriedades (básicas e intuitivas) que serão necessárias para aprofundarmos um pouco mais na Teoria das Probabilidades.

Tabela 1: Número de nascimentos em uma cidade no período de um ano.

Meses	Jan.	Fev.	Mar.	Abr.	Mai.	Jun.	Jul.	Ago.	Set.	Out.	Nov.	Dez.
Homens	3964	3797	3712	3512	3392	3761	3743	3550	4017	4173	4117	3944
Mulheres	3621	3596	3491	3391	3160	3371	3537	3407	3866	3711	3775	3665
Total	7585	7393	7203	6903	6552	7132	7280	6957	7883	7884	7892	7609

2.4 Propriedades da probabilidade

Duas importantes definições de probabilidade foram apresentadas e podem ser aplicadas quando Ω é finito ou infinito em diversas situações. Porém, quando Ω apresentar algum subconjunto não mensurável, ao mesmo não poderá ser atribuído uma probabilidade e não vamos abordar sobre este assunto, pois utilizaria o conceito geral de medida de um conjunto que está fora dos propósitos deste trabalho.

A fim de formalizar certas propriedades da Teoria das Probabilidades, vamos indicar por $\mathbb{A}$ a classe dos eventos alatórios, ou seja, a coleção de eventos de Ω aos quais podem ser atribuídos probabilidade (subconjuntos de Ω). Vejamos o exemplo 2.8 para melhor compreensão.

Exemplo 2.8 *Considere* $\Omega = \{1, 2, 3\}$*, então*

$$\mathbb{A} = \mathcal{P}(\Omega) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \{1,2,3\}\}.$$

Note que se a classe $\mathbb{A}$ *for o conjunto das partes de* Ω *finito, sempre terá* 2^n *eventos aleatórios para este experimento.*

Definiremos $\mathbb{A}$ para satisfazer igualmente o tratamento dos casos distintos de Ω (finito e infinito, ambos os casos enumeráveis).

Definição 2.6 *Considere uma classe* $\mathbb{A}$ *de subconjuntos de* Ω *não vazio, denominaremos álgebra de subconjuntos de* Ω*, caso ela satisfaça as seguintes condições:*
1) $\Omega \in \mathbb{A}$*;*
2) *Se* $A \in \mathbb{A}$*, então* $A^c \in \mathbb{A}$*;*
3) *Se* $A_1 \in \mathbb{A}$ *e* $A_2 \in \mathbb{A}$*, então* $A_1 \cup A_2 \in \mathbb{A}$*;*

Podemos afirmar também que são válidas, a partir de 1, 2 e 3, as condições abaixo:
4) $\emptyset \in \mathbb{A}$;
Demonstração: Se $\Omega \in \mathbb{A}$ então $\Omega^c \in \mathbb{A}$. Como $\Omega^c = \emptyset$, concluímos que $\emptyset \in \mathbb{A}$. $\square$

5) Se $A_1, A_2, \ldots, A_n \in \mathbb{A}$, então $\bigcup_{k=1}^{n} A_k \in \mathbb{A}$;
Demonstração: $A_1 \cup A_2 \in \mathbb{A} \Rightarrow (A_1 \cup A_2) \cup A_3 \in \mathbb{A} \Rightarrow \cdots \Rightarrow \bigcup_{k=1}^{n} A_k \in \mathbb{A}$. $\square$

6) Se $A_1, A_2, \ldots, A_n \in \mathbb{A}$, então $\bigcap_{k=1}^{n} A_k \in \mathbb{A}$;

Demonstração: Se $A_1, A_2, \ldots, A_n \in \mathbb{A}$, então $A_1^c, A_2^c, \ldots, A_n^c \in \mathbb{A}$, logo $\bigcup_{k=1}^{n} A_k^c \in \mathbb{A}$. Como $\left(\bigcup_{k=1}^{n} A_k^c\right)^c = \bigcap_{k=1}^{n} A_k$, concluímos que $\bigcap_{k=1}^{n} A_k \in \mathbb{A}$. $\square$

Para atender à classe de eventos aleatórios $\mathbb{A}$ infinita e enumerável, definiremos a σ-álgebra (sigma-álgebra).

Definição 2.7 *Considere uma classe $\mathbb{A}$ de subconjuntos de Ω não vazio, denominaremos σ-álgebra (sigma-álgebra) de subconjuntos de Ω, caso ela satisfaça as seguintes condições:*
1) $\Omega \in \mathbb{A}$;
2) *Se* $A \in \mathbb{A}$, *então* $A^c \in \mathbb{A}$;
$3'$) *Se* $A_1, A_2, A_3, \ldots \in \mathbb{A}$, *então* $\bigcup_{k=1}^{\infty} A_k \in \mathbb{A}$.

É perceptível que as condições 1 e 2 são as mesmas apresentadas na definição 2.6 e que $3'$ implica a condição 3 (definição 2.6), pois $A_1 \cup A_2 = A_1 \cup A_2 \cup A_2 \cup A_2 \ldots \in \mathbb{A}$. Concluímos que σ-álgebra é uma álgebra mas não podemos garantir a reciprocidade. Atribuímos também à σ-álgebra as seguintes condições:

$$\emptyset \in \mathbb{A} \quad \text{e} \quad \bigcap_{k=1}^{\infty} A_k \in \mathbb{A} \text{ com } A_1, A_2, A_3, \ldots \in \mathbb{A}.$$

Consideraremos que qualquer evento aleatório A de uma certa σ-álgebra $\mathbb{A}$ de um espaço amostral Ω admita uma probabilidade P que é qualquer função $P : \mathbb{A} \to \mathbb{R}$ tal que:

i) $P(\Omega) = 1$;

ii) $P(A) \geq 0$, $\forall A \in \sigma-$álgebra;

iii) Se $(A_k)_{k \geq 1}$ é uma sequência de eventos de Ω tais que os A_k são dois a dois mutuamente exclusivos, então

$$P\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} P(A_k);$$

iv) Se $(A_k)_{k \geq 1}$ é uma sequência de eventos de Ω e $A_k \downarrow \emptyset$ (decresce para o vazio), então

$$P(A_k) \to 0 \quad \text{quando} \quad k \to \infty.$$

E consequentemente seguem as seguintes propriedades da probabilidade:

Propriedade 2.4.1 *Se A^c é o complementar do evento A, então: $P(A^c) = 1 - P(A)$.*

Demonstração: Se A e A^c são eventos mutuamente exclusivos e $A \cup A^c = \Omega$. Podemos afirmar por i e iii que $P(\Omega) = P(A \cup A^c) = P(A) + P(A^c) = 1$ e subtraindo $P(A)$ em ambos os membros temos $P(A^c) = 1 - P(A)$. $\square$

Propriedade 2.4.2 *Se $A \subset B$ e $B \subset \Omega$, tem-se $P(A) \leq P(B)$.*

Demonstração: Se $A \subset B$ então $B = A \cup (A^c \cap B)$, aplicando iii pois A e $(A^c \cap B)$ são dois eventos mutuamente exclusivos, temos $P(B) = P(A) + P(A^c \cap B)$. Como $P(A^c \cap B) \geq 0$, por i, conclui-se que $P(A) \leq P(B)$. □

Propriedade 2.4.3 *Para quaisquer dois eventos $A, B \subset \Omega$, temos*

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

Demonstração: Podemos escrever $A \cup B$ como a união de dois eventos mutuamente exclusivos A e $(A^c \cap B)$ e por iii temos

$$P(A \cup B) = P(A) + P(A^c \cap B). \tag{1}$$

Analogamente, $B = (A^c \cap B) \cup (A \cap B)$ e são mutuamente exclusivos, então por iii temos

$$P(B) = P(A^c \cap B) + P(A \cap B). \tag{2}$$

Isolando $P(A^c \cap B)$ em (2) e substituindo em (1), temos

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

□

Propriedade 2.4.4 *Considere $A_1, A_2, A_3, \ldots$ eventos do espaço amostral Ω, temos:*

$$P\left(\bigcup_{k=1}^{\infty} A_k\right) = 1 - P\left(\bigcap_{k=1}^{\infty} A_k^c\right).$$

Propriedade 2.4.5 *Considere $A_1, A_2, A_3, \ldots$ eventos mutuamente exclusivos do espaço amostral Ω, temos:*

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

Propriedade 2.4.6 *Considere $(A_k)_{k \geq 1}$ é um sequência de eventos de Ω, se A_k cresce para o A ($A_k \uparrow A$) então $P(A_k) \uparrow P(A)$, consequentemente, se $A_k \downarrow A$ então $P(A_k) \downarrow P(A)$.*

2.5 Probabilidade condicional e independência

Nesta seção, ao contrário das anteriores, estaremos interessados na probabilidade associada a mais de um evento aleatório e estudaremos as relações do mesmo quando a ocorrência de um ou mais interferirem ou não na probabilidade de ocorrência de um determinado evento.

2.5.1 Probabilidade condicional

Pela definição 2.3, todo evento é um subconjunto pertencente à σ-álgebra $\mathbb{A}$ de um espaço amostral Ω e quando determinamos $P(A)$ encontramos uma relação do subconjunto A com um espaço amostral Ω. Utilizando desta análise, poderíamos fazer uso da notação $P(A \mid \Omega)$ (lê-se: probabilidade de A dado Ω) sabendo que $P(A)$ está condicionada ao espaço amostral Ω. Claro que quando definimos probabilidade a um único evento de um experimento aleatório, não há necessidade do uso desta notação pois já está subentendida esta condição, mas há casos em que a probabilidade de um evento A pode ser influenciada após a ocorrência de um outro evento B, sabendo que $A, B \subset \Omega$. Para esta situação adotaremos a notação $P(A \mid B)$ que definiremos após o exemplo 2.9.

Exemplo 2.9 *No lançamento consecutivo de três moedas e atribuindo C para "cara"e K para "coroa", obtemos o seguinte espaço amostral:*

$$\Omega = \{CCC, CCK, CKC, CKK, KCC, KCK, KKC, KKK\}.$$

Considere o evento $A = \{$obter exatamente dois resultados cara$\}$ e $B = \{$obter coroa no lançamento da primeira moeda$\}$, então:

$$A = \{CCK, CKC, KCC\} \quad e \quad B = \{KCC, KCK, KKC, KKK\}.$$

Então podemos concluir que $P(A) = 3/8$.

No entanto, qual seria a probabilidade de obter exatamente dois resultados cara sabendo que o resultado da primeira moeda foi coroa? Isto é, qual é a probabilidade de A ocorrer sabendo que B ocorreu?

Neste caso, como já sabemos o resultado da primeira moeda, o espaço amostral é reduzido a 4 resultados possíveis representado por $n(B)$, e podemos afirmar que B é o nosso novo espaço amostral. Desta forma a probabilidade de ocorrer A dado B é igual a $1/4$, pois dos quatro resultados possíveis de B apenas $KCC \in A$.

Observe que para determinarmos a probabilidade de A dado B obtemos a razão de $n(A \cap B)$ por $n(B)$ e uma nova expressão pode ser obtida se dividirmos o numerador e o denominador por $n(\Omega)$, então:

$$P(A \mid B) = \frac{n(A \cap B)}{n(B)} \Rightarrow P(A \mid B) = \frac{\frac{n(A \cap B)}{n(\Omega)}}{\frac{n(B)}{n(\Omega)}} \Rightarrow P(A \mid B) = \frac{P(A \cap B)}{P(B)}.$$

Definição 2.8 *Considere dois eventos A e B de um espaço amostral Ω, com $P(B) > 0$. A probabilidade condicional do evento A dado B pode ser definida por*

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

Até o momento, nos preocupamos em definir e exemplificar a probabilidade condicional, mas quando há necessidade de obter a probabilidade da ocorrência simultânea dos eventos A e B podemos obter uma nova expressão a partir da definição 2.8. Essa nova expressão, importantíssima, é conhecida como Teorema da Multiplicação.

Teorema 2.1 (Teorema da Multiplicação) *Dado dois eventos A e B de um espaço amostral Ω, a probabilidade simultânea de A e B é definida por:*

$$P(A \cap B) = P(A) \cdot P(B \mid A) \quad \text{ou} \quad P(A \cap B) = P(B) \cdot P(A \mid B).$$

Vejamos a aplicação do Teorema da Multiplicação no exemplo 2.10.

Exemplo 2.10 *Considere duas caixas com garrafas de vinho dispostas da seguinte forma:*

- *Caixa* 1 *contém* 8 *garrafas de vinho, sendo* 3 *garrafas de vinho nacional e* 5 *de vinho importado;*

- *Caixa* 2 *contém* 6 *garrafas de vinho, sendo* 2 *garrafas de vinho nacional e* 4 *de vinho importado.*

Uma caixa é escolhida aleatoriamente e desta é retirada uma garrafa ao acaso. Qual é a probabilidade desta garrafa ser de vinho importado?

Vamos considerar que o evento A seja retirar uma garrafa de vinho importado e o evento B escolher uma das duas caixas com garrafas de vinho. Como existem duas caixas, a probabilidade de B é $1/2$. Já o evento A depende da caixa que foi escolhida (probabilidade condicional) pois se escolhermos a caixa 1, temos $P(A \mid B) = 5/8$ e se escolhermos a caixa 2, temos $P(A \mid B) = 4/6 = 2/3$. Para calcularmos a probabilidade deste problema devemos obter a probabilidade simultânea de cada situação aplicando o Teorema da Multiplicação. Note que temos duas situações:

- *(1^a situação) Retirar uma garrafa de vinho importado após escolher a caixa* 1.
  $P_1(A \cap B) = P_1(B) \cdot P_1(A \mid B) = \dfrac{1}{2} \cdot \dfrac{5}{8} = \dfrac{5}{16}$.

- *(2^a situação) Retirar uma garrafa de vinho importado após escolher a caixa* 2.
  $P_2(A \cap B) = P_2(B) \cdot P_2(A \mid B) = \dfrac{1}{2} \cdot \dfrac{2}{3} = \dfrac{1}{3}$.

As duas situações são distintas logo, a probabilidade do problema é a soma das probabilidades das possíveis situações:

$$P = P_1(A \cap B) + P_2(A \cap B) = \frac{5}{16} + \frac{1}{3} = \frac{31}{48}.$$

Logo, a probabilidade de escolher uma garrafa de vinho importado é $31/48$.

2.5.2 Independência

Considere o seguinte problema:

Exemplo 2.11 *Uma caixa contêm* 12 *canetas vermelhas e* 8 *canetas pretas. Uma pessoa retira, ao acaso, sucessivamente e com reposição, duas dessas canetas. Qual é a probabilidade de que as duas canetas retiradas sejam vermelhas?*

Chamaremos a primeira retirada de evento A e a segunda retirada de evento B, logo $P(A) = 12/20 = 3/5$ *e* $P(B) = 12/20 = 3/5$. *Observe que a ocorrência de um evento não afeta a ocorrência (ou não ocorrência) do outro evento, podemos afirmar então que* $P(A \mid B) = P(A)$ *e* $P(B \mid A) = P(B)$. *Estes eventos são chamados independentes.*

Aplicando o Teorema da Multiplicação, chegaremos ao resultado do problema.

$$P(A \cap B) = P(A) \cdot P(B) = \frac{3}{5} \cdot \frac{3}{5} = \frac{9}{25}.$$

Logo, a probabilidade de retirar duas canetas vermelhas é $9/25$.

No exemplo 2.11, foi perceptível a independência dos eventos A e B, mas nem sempre será fácil avaliar se a ocorrência de um evento interfere ou não a ocorrência de outro.

Uma forma mais confiável de avaliar se dois eventos são independentes é verificar a veracidade da condição $P(A \cap B) = P(A) \cdot P(B)$. Podemos definir a independência de eventos da seguinte forma:

Definição 2.9 *Dois eventos, A e B, são independentes se, e somente se,*

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

3 Lei dos grandes números

3.1 Conceito geral

O estudo sobre probabilidades abordada até o momento se aplica à probabilidade clássica e frequentista. Vimos que na probabilidade clássica, os eventos são equiprováveis e suas probabilidades podem ser pré-determinadas sem a necessidade da ocorrência do experimento aleatório. Na prática, esta definição não é satisfatória pois existem poucos experimentos que facilmente podemos identificar os eventos simples e equiprováveis. Já a probabilidade frequentista obtida através da frequência relativa da ocorrência de um evento qualquer, pode se dizer, que geralmente é a mais adequada pois existem inúmeras situações em que os eventos não são equiprováveis ou não há como determinar a probabilidade de um evento precisamente sem a ocorrência de inúmeras (número bem grande) repetições do experimento aleatório. Podemos citar algumas destas situações: a probabilidade de um peça estar defeituosa em uma linha de produção de uma empresa, a probabilidade de um time ser campeão, a probabilidade de um candidato vencer as eleições e etc.

Na subseção 2.3, apresentamos um exemplo para cada modelo (exemplos: 2.6 e 2.7), mas para compreendermos um pouco mais sobre a probabilidade frequentista, consideraremos um experimento simples e bem conhecido para que possamos, em breve, definir a importantíssima Lei dos Grandes Números.

Vamos considerar o lançamento de uma moeda não viciada. Sabemos que a probabilidade clássica de obter cara (evento A) é $P(A) = 1/2 = 0,5 = 50\%$, ou seja, no total de dois possíveis resultados há apenas uma única maneira de se obter cara, lembre-se que não necessitamos da ocorrência do experimento para tal cálculo. Esse resultado obtido também nos dá uma ideia que a cada dois lançamentos consecutivos da moeda um é cara e o outro coroa, mas na prática isto nem sempre acontece. Não podemos garantir que se caso o resultado do primeiro lançamento for coroa, o segundo será com certeza cara, ou que, em dez lançamentos consecutivos, 5 resultados serão cara e 5 coroa independente da ordem. Suponhamos que em 10 lançamentos consecutivos desta mesma moeda, obtemos seis resultados cara, pela probabilidade frequentista temos $P(A) = 6/10 = 0,6 = 60\%$. Podemos perceber que a probabilidade frequentista não equivale à clássica, então o modelo frequentista não é confiável? Não! É confiável com certeza, mas quando o número de experimentos é relativamente grande.

Tabela 2: Cálculo da probabilidade frequentista em relação ao número total de lançamentos de uma moeda.

Total de lançamentos	**Número de caras**	**Probabilidade frequentista**	**Diferença em relação à probabilidade clássica**
1	1	100%	50%
10	6	60%	10%
100	56	56%	6%
1000	486	48,6%	1,4%
10000	4956	49,56%	0,44%
100000	50060	50,06%	0,06%

A tabela 2 apresenta alguns resultados obtidos de uma simulação que realizei, através do Excel, do experimento "lançamento de uma moeda não viciada". É importante lembrar que quando nos referimos à expressão "vários lançamentos" consideraremos que foram lançadas várias moedas ou que foram realizados, consecutivamente, vários lançamentos da mesma moeda. Após a realização destes experimentos independentes, podemos perceber que há uma diferença da probabilidade frequentista em relação à clássica, mas à medida que o número de lançamentos aumenta a probabilidade frequentista se aproxima cada vez mais da clássica, ou seja, aumenta o seu nível de confiabilidade em relação ao valor esperado. Definimos **valor esperado** de uma variável aleatória como a soma do produto da probabilidade de cada possibilidade de saída da experiência pelo seu respectivo valor.

Podemos verificar através do gráfico da figura 1, de maneira mais ampla, a aproximação da probabilidade frequentista em relação à clássica, ou seja, no valor esperado (50%).

A partir da verificação do comportamento do gráfico da figura 1, podemos definir de maneira generalizada a Lei dos Grandes Números.

Definição 3.1 *Considere X uma variável aleatória e n o número de ensaios de um experimento, temos $X_1, X_2, X_3, \cdots, X_{n-1}, X_n$ uma sequência de variáveis aleatórias independentes e identicamente distribuídas para $n \geq 2$. A média aritmética de n ensaios converge para o valor esperado EX_n quando $n \to \infty$, obedecendo à Lei dos Grandes Números.*

A Lei dos Grandes Números foi provada pela primeira vez pelo matemático Jakob Bernoulli, em 1692, e publicada em sua obra [12], em 1713. Nesta obra ele afirma que:

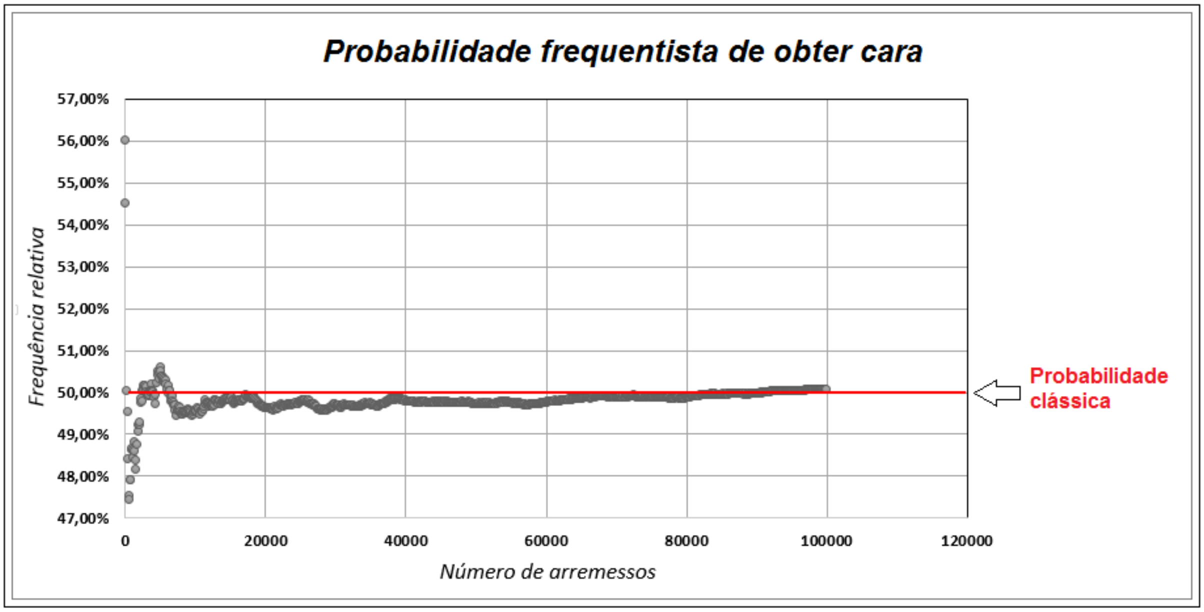

Figura 1: Gráfico do comportamento da probabilidade frequentista em relação ao valor esperado de acordo com o número de arremessos.

> Se um evento de probabilidade p é observado repetidamente em ocasiões independentes, a proporção da frequência observada do mesmo evento em relação ao número total de repetições convergem em direção a p à medida que o número de repetições se torna arbitrariamente grande.

O surgimento desta lei se deve à necessidade de estimar um evento casual, por exemplo, prever o sexo de um bebê antes do seu nascimento. Bernoulli pensando nisso, mostrou que é possível e abriu as portas para o estudo de outros matemáticos (Siméon-Denis Poisson, Pafnuti Tchebychev, Émile Borel, Andrei Markov, Francesco Paolo Cantelli, Andrei Kolmogorov e Aleksandr Khinchin) que contribuiram no desenvolvimento desta lei dando origem a duas formas: a Lei Fraca dos Grandes Números e a Lei Forte dos Grandes Números. Essas duas formas possuem o mesmo conceito mas diferem no tipo de convergência que veremos a seguir.

3.2 Tipos de convergência

Apresentaremos nesta seção dois tipos de convergências para uma sequência de variáveis aleatórias, os quais são necessários para definirmos as duas formas da Lei dos Grandes Números. Destacaremos apenas os conceitos, pois as demonstrações dos mesmos estão além da finalidade do trabalho.

Considere um espaço de probabilidade $(\Omega, \mathbb{A}, P)$ e uma sequência de variáveis aleatórias $X_1, X_2, X_3, \cdots, X_n$ definida no mesmo.

Definição 3.2 (Convergência em probabilidade) *A sequência de variáveis aleatórias X_n converge para a variável aleatória X em probabilidade se para todo $\varepsilon > 0$,*

$P\left(|(X_n - X)| \geq \varepsilon\right) \to 0$ *quando* $n \to \infty$. *Escreve-se* $X_n \xrightarrow{P} X$.

A convergência em probabilidade garante que quando n está suficientemente grande, a probabilidade de X_n não pertencer ao intervalo $[X - \varepsilon, X + \varepsilon]$ é praticamente nula.

Definição 3.3 (Convergência quase certa) *A sequência de variáveis aleatórias* X_n *converge para a variável aleatória* X *quase certamente se* $P\left(X_n \to X\right) = 1$ *quando* $n \to \infty$. *Escreve-se* $X_n \xrightarrow{q.c} X$.

A convergência quase certa é mais precisa, ou seja, é pontual, pois quando n for suficientemente grande, X_n estará convergindo para X com probabilidade 1.

Podemos dizer que a convergência quase certa é mais forte que a convergência em probabilidade pois:

$$\begin{array}{rcl} \text{convergência quase certa} & \Rightarrow & \text{convergência em probabilidade,} \\ & \text{mas} & \\ \text{convergência em probabilidade} & \not\Rightarrow & \text{convergência quase certa.} \end{array}$$

Uma boa prova para a implicação e um exemplo para a não implicação apresentadas acima podem ser encontrados em James ([15], p.210-212).

3.3 As duas formas da Lei dos Grandes Números

Depois de apresentadas as convergências (em probabilidade e quase certa) podemos agora definir a Lei Fraca e a Lei Forte dos Grandes Números.

Definição 3.4 (Lei Fraca dos Grandes Números de Tchebychev) *Seja uma sequência de variáveis aleatórias* $X_1, X_2, X_3, \cdots$, *independentes e identicamente distribuídas com média* μ *e variância* σ^2 *e* $S_n = \sum_{n}^{k=1} X_k$. *Tem-se* $\dfrac{S_n}{n} \xrightarrow{p} \mu$.

O matemático russo Tchebychev desenvolveu uma desigualdade para demonstrar trivialmente a Lei Fraca dos Grandes Números, já a prova de Bernoulli é muito mais complexa. Essas provas podem ser consultadas pelo leitor em James ([15], p.213). A Lei Fraca dos Grandes Números estabelece basicamente que quanto maior o número de observações (um número suficientemente grande), a sua média estará nas proximidades do valor esperado com uma probabilidade altíssima.

Definição 3.5 (Lei Forte dos Grandes Números de Kolmogorov) *Seja uma sequência de variáveis aleatórias* $X_1, X_2, X_3, \cdots$, *independentes, identicamente distribuídas e integráveis, com média* μ *e* $S_n = \sum_{n}^{k=1} X_k$. *Tem-se* $\dfrac{S_n}{n} \xrightarrow{q.c} \mu$.

A Lei Forte dos Grandes Números estabelece que a média das observações converge para o valor esperado com probabilidade 1 à medida que o número de observações se torna cada vez maior (suficientemente grande).

Como podemos observar, a grande diferença entre às leis apresentadas é que uma converge em probabilidade (Lei Fraca) e a outra converge quase certamente (Lei Forte), mas não apresentam conceitos diferentes da Lei dos Grandes Números.

4 Probabilidades e o futebol

O futebol, hoje, é considerado o esporte mais popular do mundo pois para a sua realização não são necessários muitos equipamentos ou recursos. No Brasil, conhecido como a "Terra do Futebol", várias pessoas são apaixonadas pelo esporte e, geralmente, adotam um time para torcer fervorosamente.

Durante o campeonato brasileiro, os programas esportivos utilizam da estatística para analisarem os dados de cada partida, observam qual time se portou melhor e apresentam os dados mais importantes que determinam a classificação de cada equipe até o momento. Já a probabilidade é usada para estimar as chances de cada equipe ser campeão, ser rebaixado, estar classificado para algum torneio e etc. Estes números geram muita discussão entre os torcedores pois alguns consideram inúteis, já outros consideram úteis mas não sabem como são obtidos.

Muitos jovens do Ensino Médio, quando estão aprendendo probabilidade, têm a curiosidade de saber como podemos aplicá-la em diversas áreas e grande parte deles perguntam: como podemos calcular as chances de um time de futebol ser campeão, se todos têm a mesma chance de vencer uma partida? Esta indagação não pertence somente a estes alunos, e sim, a todos apaixonados pelo futebol, e por isso escolhemos este esporte para mostrarmos a aplicação da Teoria das Probabilidades. Mas lembre-se, o que vamos apresentar nesta seção poderá ser utilizado em outros esportes ou até mesmo em outras áreas.

A Teoria das Probabilidades dentro do futebol pode ser utilizada em diversas situações, vejamos algumas delas:

- avaliar uma possível contusão de uma atleta;
- estimar o tempo necessário para que o atleta esteja no auge do seu preparo físico;
- presumir o público pagante para a próxima partida;
- calcular as chances de um time vencer uma determinada partida;
- determinar o futuro financeiro caso contrate um grande jogador e etc.

Vamos nos preocupar em aplicar a Teoria das Probabilidades para definir as chances de cada time no resultado de uma ou mais partidas e as chances das possíveis posições ao final do campeonato.

4.1 Um modelo matemático para probabilidade no futebol

"Caixinha de surpresas"[3], este é o termo conhecido pelo futebol por apresentar resultados imprevisíveis, por exemplo, ninguém imagina que o líder do campeonato, jogando em casa, perderá para o último colocado, e isso pode acontecer como já vimos algumas vezes em alguns campeonatos pelo mundo, mas se considerarmos todos os campeonatos

[3] A expressão que virou clichê foi inventada pelo comentarista Benjamim Wright, pai do ex-árbitro José Roberto Wright. Essa é apenas uma das famosas expressões do jargão futebolístico.

já ocorridos por todo o planeta, pouquíssimas vezes nos deparamos com tal fato. E por se tratar de um experimento aleatório, a probabilidade tem seu papel fundamental, estimando matematicamente a chance deste evento ocorrer. Qualquer matemático sabe que quando falamos sobre probabilidade, estamos nos referindo ao tratamento das incertezas, possibilidades, chances de um acontecimento ocorrer ou não, e o futebol apresenta tudo isso. Há uma ligação entre futebol e probabilidade. Quando tentamos mensurar a chance de um time ser campeão ou ser rebaixado não estamos fazendo diferente do que foi feito pelos matemáticos, no século XVI, com os jogos de azar (cartas, dados, moedas, etc), tentando medir o número de ocorrências de um evento durante um longo tempo.

Existe uma grande diferença de obtermos a probabilidade para o futebol, como queremos, da qual é obtida nos jogos de azar. No lançamento de uma moeda honesta para cada possível resultado temos a probabilidade de $1/2$, no lançamento de um dado não viciado temos para cada número uma probabilidade de $1/6$, na mega-sena temos

$$P = \frac{1}{C_6^{60}} = \frac{6!54!}{60!} = \frac{1}{50063860}$$

para cada combinação de seis números (um jogo simples), podemos perceber que a probabilidade é obtida *a priori* pois tratam-se de experimentos aleatórios com resultados equiprováveis, todos têm a mesma probabilidade que é comum nos jogos de azar. Isso não ocorre no futebol! Não é conveniente atribuir a mesma probabilidade de vencer uma partida para um time que está disparado na liderança e para um time que está entre os últimos colocados do campeonato. Vários fatores podem interferir: escalações diferentes, árbitros e bandeirinhas, mando de campo, local do jogo, torcida, clima, contratações recentes, estado emocional, desempenho do ataque, desempenho da defesa, convocações de jogadores para as suas respectivas seleções e etc. São inúmeros fatores que podem interferir no rendimento de um time no campeonato, alguns com mais intensidade do que outros, e isto é um grande problema para estimar uma probabilidade para os possíveis resultados em uma partida de futebol.

Como as equipes não apresentam as mesmas chances no decorrer do campeonato e ainda existem vários fatores que afetam os possíveis resultados das partidas, o ideal é que seja feito um modelo de simulação de resultados e, a partir deste, fazer uso de inúmeras simulações aplicando a Lei dos Grandes Números na qual coletaremos os dados para obtermos as probabilidades desejadas. É possível que alguém esteja pensando porque não usar a razão do número de casos favoráveis das combinações dos jogos restantes que leve um determinado time a ser campeão pelo número total de combinações possíveis? É impotante salientar que usaremos um modelo criado para o campeonato brasileiro de futebol. Um campeonato de pontos corridos (ida e volta) disputado por 20 clubes, temos 38 rodadas com 10 partidas cada totalizando 380 jogos e cada jogo há 3 possibilidades: vitória do mandante, empate ou vitória do visitante. Então o número total dos possíveis resultados é 3^{380}, um número extraordinariamente grande a ser computado gerando todos eles. Em [18], os autores mostram que:

Em 15 rodadas este número sobe para 3^{150}. Bem, em tese, 3^{150} (aproximadamente 3×10^{71}) é até tratável em um computador. De fato, para sistemas criptográficos atuais são usados códigos de tamanho até 2^{1024}, como é o caso da nova codificação do principal banco estatal brasileiro. Agora, o que é absolutamente impraticável é gerar em "tempo finito" todos os universos possíveis. Assim, por exemplo, se cada caso fosse gerado em 1 trilhonésimo de segundo (processadores na faixa de Terahertz), o computador gastaria $1,18 \times 10^{50}$ séculos para gerar todos eles. A título de comparação, as estimativas atuais para a idade do Universo apontam para $1,4 \times 10^{8}$ séculos.

Desta forma, também, estaríamos desconsiderando o poder de ataque ou defesa das equipes pois o saldo de gols é um importante critério para desempate. A vitória de um time mandante, através de combinações, descarta se ele venceu por um ou mais gols e quando houver empate em pontos em uma das combinações, qual time levaria a melhor?

O modelo que será apresentado foi desenvolvido por um grupo do Departamento de Matemática da UFMG, em [18] os autores sugerem que a base deste sirva para criação de outros modelos tanto para o futebol quanto para outros esportes, por exemplo, a fórmula 1. No modelo desenvolvido por eles, alguns fatores importantes que, possivelmente, interferem no resultado de uma partida serão peças importantes para toda a metodologia que veremos a seguir, é válido lembrar que estamos falando de probabilidades e os resultados que poderão ser obtidos não afirmam que ocorrerão ou não, serão apenas as chances de cada resultado ocorrer mantendo o comportamento dos times nos jogos realizados até o momento.

4.1.1 Vetores força de cada time

Logo no início do campeonato, eles agregam dois vetores tridimensionais $\vec{P}_M$ e $\vec{P}_V$ para cada time e estes vetores permanecerão com o mesmo até o fim do campeonato sofrendo algumas alterações de acordo com seu retrospecto ou durante as simulações. O vetor $\vec{P}_M$ representará a força como mandante na partida e o vetor $\vec{P}_V$ a força como visitante. Segundo os autores, essa necessidade de agregar dois vetores para cada time se deve pelo fator "mando de campo" pois um time se comporta de uma maneira diferente quando joga em casa com o apoio de sua torcida do que quando joga no campo do adversário.

Os vetores apresentam as seguintes coordenadas:

$$\vec{P}_M = (PMV, PME, PMD) \quad \text{e} \quad \vec{P}_V = (PVD, PVE, PVV).$$

Em que:

- PMV é a probabilidade de vitória como mandante;
- PME é a probabilidade de empate como mandante;
- PMD é a probabilidade de derrota como mandante;
- PVD é a probabilidade de derrota como visitante;

- PVE é a probabilidade de empate como visitante;

- PVV é a probabilidade de vitória como visitante.

Cada coordenada é representada por um número real entre 0 e 1 e a soma das coordenadas de cada vetor resulta em 1, ou seja, 100%. Como eles não levam em consideração o retrospecto da cada time nos campeonatos anteriores e nem a tradição adquirida, então todos os times deverão iniciar com as mesmas probabilidades. O vetor $\vec{P}_M$ será definido da seguinte forma $\vec{P}_M = (0,4; 0,3; 0,3)$, ou seja, 40% de vitória, 30% de empate e 30% de derrota em casa. Eles disseram que poderiam atribuir 1/3 para cada coordenada, mas isso é uma decisão pessoal. Geralmente, o time mandante possui mais fatores que lhe favorecem e por isso eles decidiram que a probabilidade de vitória seja maior do que empate e derrota. A mesma ideia eles consideraram para o time quando joga como visitante, ele costuma ter muitas dificuldades, então o vetor $\vec{P}_V$ será definido da seguinte forma $\vec{P}_V = (0,4; 0,3; 0,3)$.

Esses vetores representam a força que o time tem quando joga dentro ou fora de casa. A seguir veremos como as coordenadas destes vetores de cada time irão se alterar conforme seu desempenho ou durantes às simulações.

4.1.2 Vetor probabilidade de cada partida

Na seção anterior, descrevemos como foram definidos os vetores iniciais $\vec{P}_M$ e $\vec{P}_V$ de cada time, mas com o desenvolver do campeonato ou das simulações de jogos, a cada rodada um dos vetores sofrerá mudanças de acordo com o resultado obtido na partida (resultado real ou simulado) e o nível de dificuldade do adversário.

Vamos considerar que já estamos com algumas rodadas realizadas no campeonato e dois times irão se enfrentar na próxima rodada, por exemplo, o duelo São Paulo $\times$ Grêmio. O São Paulo apresenta os vetores $\vec{P}_{M_{SAO}} = (0,25; 0,45; 0,30)$ e $\vec{P}_{V_{SAO}} = (0,42; 0,29; 0,29)$ e o Grêmio, $\vec{P}_{M_{GRE}} = (0,32; 0,37; 0,31)$ e $\vec{P}_{V_{GRE}} = (0,33; 0,57; 0,10)$. O vetor probabilidade da partida será obtido através da fórmula:

$$P_{A\times B} = \left(\frac{PMV_A + PVD_B}{2}, \frac{PME_A + PVE_B}{2}, \frac{PMD_A + PVV_B}{2}\right).$$

Como podemos observar, os criadores do modelo definiram que o vetor probabilidade da partida é a média aritmética simples dos vetores força $\vec{P}_M$ e $\vec{P}_V$, sendo o primeiro pertencente ao time mandante e o segundo do time visitante. Segundo eles, a primeira coordenada dos vetores citados representa a probabilidade de vitória do mandante e a probabilidade de derrota do time visitante, por isso a média entre estas coordenadas, pois a chance de vitória do mandante instiga a derrota do visitante e pensamento análogo às demais coordenadas. Então o vetor probabilidade da partida São Paulo $\times$ Grêmio será:

$$
\begin{aligned}
P_{SAO\times GRE} &= \left(\frac{PMV_{SAO}+PVD_{GRE}}{2}, \frac{PME_{SAO}+PVE_{GRE}}{2}, \frac{PMD_{SAO}+PVV_{GRE}}{2}\right)\\
&= \left(\frac{0,25+0,33}{2}, \frac{0,45+0,57}{2}, \frac{0,30+0,10}{2}\right)\\
&= (0,29; 0,51; 0,20)\,.
\end{aligned}
$$

As probabilidades são: 29% de chances para a vitória do São Paulo, 51% de chances de resultar em empate e 20% de chances para a vitória do Grêmio. Essas probabilidades são utilizadas para a simulação desta partida, veremos isso na seção 4.1.4. Vejamos as chances para a partida Grêmio $\times$ São Paulo:

$$
\begin{aligned}
P_{GRE\times SAO} &= \left(\frac{PMV_{GRE}+PVD_{SAO}}{2}, \frac{PME_{GRE}+PVE_{SAO}}{2}, \frac{PMD_{GRE}+PVV_{SAO}}{2}\right)\\
&= \left(\frac{0,32+0,42}{2}, \frac{0,37+0,29}{2}, \frac{0,31+0,29}{2}\right)\\
&= (0,37; 0,33; 0,30)\,.
\end{aligned}
$$

Vitória do Grêmio 37%, empate 33% e vitória do São Paulo 30%.

Em [18], os autores mencionam a existência de alguns casos particulares para determinar as probabilidades de uma partida, um deles são os clássicos regionais. Para esta situação os vetores $\vec{P}_V$ de cada time não são utlizados, vejamos a fórmula abaixo:

$$
P_{A\times B} = \left(\frac{PMV_A+PMD_B}{2}, \frac{PME_A+PME_B}{2}, \frac{PMD_A+PMV_B}{2}\right).
$$

Os autores justificam que essa variação na forma de calcular as probabilidades de uma partida se deve ao fato de que o time mandante, mesmo jogando no seu estádio, não irá afetar o seu adversário que já está ambientado com o local e não tem que se desgastar com viagem, e é claro que, não podemos esquecer da grande rivalidade. “Clássico é clássico!” O Grenal, clássico entre Grêmio e Internacional, é um exemplo de um jogo muito equilibrado considerando a história entres estes dois times e para não deixarmos de lado o retrospecto no campeonato, utilizam os vetores $\vec{P}_{V_{GRE}}$ e $\vec{P}_{V_{INT}}$. Observe que as probabilidades são as mesmas independente quem seja o mandante pois serão utilizados os mesmos vetores.

Outras situações podem levar ao mesmo procedimento, por exemplo, o Flamengo jogando em Brasília contra o Corinthians ou o Palmeiras jogando no seu estádio com portões fechados sem a presença da sua torcida. Segundo eles, são situações em que o matemático tem que reconhecer e avaliar se são fatores relevantes ou não para o seu modelo. Isso é uma decisão pessoal, mas deve-se sempre procurar fazer com que os resultados sejam adequados aos observados na realidade.

4.1.3 Fim de jogo e suas consequências

As probabilidades da partida foram definidas e quando o resultado é obtido, seja real ou simulado, este atualizará as coordenadas dos vetores força envolvidos de forma negativa ou positiva. Se consideramos o time mandante, uma vitória em casa aumenta suas chances de vencer novamente como mandante e consequentemente diminui as chances de empate e derrota pois somando as probabilidades, o resultado deverá ser igual a 1. Já o time visitante perdendo aumenta suas chances de uma nova derrota fora de casa e diminuirá as chances de empate e vitória. Esta alteração das probabilidades depende do resultado obtido, do rendimento que o adversário tem no campeonato até o momento e um peso α.

O resultado definirá qual a fórmula a ser utilizada pois existem 3 possíveis resultados: vitória do mandante, empate e vitória do visitante. E para cada possível resultado, existirá uma fórmula que atualiza os vetores força $\vec{P}_M$ e $\vec{P}_V$ dos times, mandante e visitante, respectivamente.

O rendimento do adversário "*Rd*"é levado em consideração neste modelo pois como foi dito antes, vencer um líder e vencer o último colocado são situações bem adversas e o moral do time, com certeza, cresce significativamente derrotando o líder do que o lanterna. Segundo os autores, optar pelo rendimento do adversário do que a posição na tabela ou pontos adquiridos, se deve ao fato que em algumas vezes no campeonato os times não apresentam a mesma quantidade de jogos realizados, e pode acontecer de um time, momentaneamente, estar em uma colocação melhor mas com um rendimento inferior a outras equipes. O rendimento Rd é obtido da razão dos pontos conquistados pelos pontos já disputados.

O peso α representa a intensidade em que o resultado obtido poderá afetar os vetores força dos times envolvidos na partida, quanto menor α, maior a sensibilidade, por exemplo, se um time mandante vencer e for utilizado $\alpha = 0$, o aumento das chances de vitória será bem significativa dando muita força ao time para vencer a próxima partida em casa, já quando α é um valor muito alto, a vitória conquistada não interfere quase nada, ou seja, praticamente é desconsiderado o fator "moral". O grupo de matemáticos da UFMG, autores do modelo, geralmente utilizam o peso $\alpha = 4$.

Considere a vitória do time mandante, os vetores força iniciais $\vec{P}_{iM}$ e $\vec{P}_{iV}$, respectivamente, do time mandante e visitante, serão atualizados e transformados nos vetores força finais $\vec{P}_{f_M}$ e $\vec{P}_{f_V}$, conforme as seguintes fórmulas:

$$\vec{P}_{f_M} = \frac{\alpha \cdot \vec{P}_{iM} + Rd \cdot (1,0,0)}{\alpha + Rd} \quad \text{e} \quad \vec{P}_{f_V} = \frac{\alpha \cdot \vec{P}_{iV} + (1 - Rd) \cdot (1,0,0)}{\alpha + (1 - Rd)}.$$

Observação: Rd é o rendimento do time **adversário**.

Os vetores força finais, $\vec{P}_{f_M}$ e $\vec{P}_{f_V}$, mantêm a condição de que a soma das suas três coordenadas resulta na constante 1, conforme a demonstração abaixo. Considere o raciocínio análogo para as demais fórmulas de atualização de vetores força.

Demonstração: Considere o vetor $\vec{P}_{iM} = (PMV, PME, PMD)$, sabemos que $PMV + PME + PMD = 1$. Queremos provar que a soma das coordenadas do vetor $\vec{P}_{f_M}$ também resulta em 1.

Vamos, inicialmente, obter as coordenadas do vetor $\vec{P}_{f_M}$:

$$\begin{aligned}\vec{P}_{f_M} &= \frac{\alpha \cdot \vec{P}_{iM} + Rd \cdot (1,0,0)}{\alpha + Rd} \\ &= \frac{\alpha \cdot (PMV, PME, PMD) + Rd \cdot (1,0,0)}{\alpha + Rd} \\ &= \left(\frac{\alpha PMV + Rd}{\alpha + Rd}, \frac{\alpha PME}{\alpha + Rd}, \frac{\alpha PMD}{\alpha + Rd} \right).\end{aligned}$$

Somando as coordenadas do vetor $\vec{P}_{f_M}$ temos:

$$\begin{aligned}\frac{\alpha PMV + Rd}{\alpha + Rd} + \frac{\alpha PME}{\alpha + Rd} + \frac{\alpha PMD}{\alpha + Rd} &= \frac{\alpha PMV + Rd + \alpha PME + \alpha PMD}{\alpha + Rd} = \\ = \frac{\alpha(PMV + PME + PMD) + Rd}{\alpha + Rd} &= \frac{\alpha + Rd}{\alpha + Rd} = 1.\end{aligned}$$

$\square$

Considere a derrota do time mandante, os vetores força iniciais $\vec{P}_{iM}$ e $\vec{P}_{iV}$, respectivamente, do time mandante e visitante, serão atualizados e transformados nos vetores força $\vec{P}_{f_M}$ e $\vec{P}_{f_V}$, conforme as seguintes fórmulas:

$$\vec{P}_{f_M} = \frac{\alpha \cdot \vec{P}_{iM} + (1 - Rd) \cdot (0,0,1)}{\alpha + (1 - Rd)} \quad \text{e} \quad \vec{P}_{f_V} = \frac{\alpha \cdot \vec{P}_{iV} + Rd \cdot (0,0,1)}{\alpha + Rd}.$$

Conforme o modelo do grupo de matemáticos do Departamento de Matemática da UFMG, em caso de empate, a mesma ideia não pode ser aplicada pois se um time mandante empata não significa que no próximo jogo em casa ele terá maior chance de empatar e as chances para vencer ou perder serão menores. Segundo eles, o empate pode ser considerado quase uma derrota ou quase uma vitória depende do adversário. Se o adversário é um time que está no topo da tabela, suponhamos que seja o líder do campeonato, o time mandante empatar com este líder, mesmo jogando em casa, o moral irá subir, claro que não seria o mesmo se vencesse o líder, mas dá moral para o time com certeza, então a única chance que diminuiria para o próximo jogo em casa seria uma derrota, as chances de empatar e vencer aumentariam. Se o adversário é um time que está mal na tabela, empatar em casa com ele altera o moral do time negativamente, pois seria quase uma derrota, claro que ser derrotado seria mais desastroso, então as chances de empatar e perder aumentam, enquanto a chance de vencer diminui.

Eles adotaram duas possibilidades de atualização do vetor força quando o resultado for um empate:

- quando o adversário tem um rendimento maior que 4/9;

- quando o adversário tem um rendimento de 0 até 4/9.

Considere o empate do time mandante com um adversário que possui $Rd > 4/9$, o vetor força $\vec{P}_{iM}$ será atualizado e transformado no vetor força $\vec{P}_{fM}$, conforme a fórmula seguinte:

$$\vec{P}_{fM} = \frac{\alpha \cdot \vec{P}_{iM} + (2Rd - 1) \cdot \left(\frac{1}{2}, \frac{1}{2}, 0\right) + 2\,(1 - Rd) \cdot (0, 1, 0)}{\alpha + 1}.$$

Considere o empate do time mandante com um adversário que possui $0 \leq Rd \leq 4/9$, o vetor força $\vec{P}_{iM}$ será atualizado e transformado no vetor força $\vec{P}_{fM}$, conforme a fórmula seguinte:

$$\vec{P}_{fM} = \frac{\alpha \cdot \vec{P}_{iM} + (1 - 2Rd) \cdot \left(0, \frac{1}{2}, \frac{1}{2}\right) + 2Rd \cdot (0, 1, 0)}{\alpha + 1}.$$

Essas duas fórmulas apresentadas caso o time mandante empate são válidas também para o time visitante, considerando os mesmos critérios.

4.1.4 Simulação de cada partida

Após as atualizações dos vetores força e definidas as probabilidades para a próxima rodada, um programa deverá simular cada partida.

Para um melhor entendimento, vamos usar as probabilidades da partida São Paulo × Grêmio citadas na seção 4.1.2. O vetor probabilidade $P_{SAO \times GRE} = (0,29; 0,51; 0,20)$ informa que o São Paulo tem chance de 29% de vitória, empate 51% e vitória do Grêmio 20%. De acordo com o modelo da UFMG, o intervalo $[0, 1]$ será dividido em três subintervalos: $[0; 0,29]$ (vitória do mandante), $(0,29; 0,8)$ (empate) e $[0,8; 1]$ (vitória do visitante). O computador sorteia um número real x aleatoriamente de 0 a 1 e este determina o resultado. Suponhamos o número sorteado seja 0,8125, este valor pertence ao subintervalo que dá a vitória ao Grêmio.

Generalizando, considere P_v a probabilidade de vitória do mandante, P_e a probabilidade de empate e P_d a probabilidade de vitória do visitante. O intervalo [0,1] deverá ser dividido da seguinte forma:

- $[0, P_v]$ é o subintervalo que representa a vitória do mandante;
- $(P_v, P_v + P_e)$ é o subintervalo que representa empate na partida;
- $[P_v + P_e, 1]$ é o subintervalo que representa a vitória do visitante.

Um número real x, de 0 a 1, é sorteado aleatoriamente e simulará o resultado da partida usando os seguintes critérios:

- se $x \leq P_v$, então o mandante vence a partida;
- se $P_v < x < P_v + P_e$, então a partida termina empatada;

- se $P_v + P_e \leq x \leq 1$, então a vitória é dada ao visitante.

O placar da partida, também é determinado a partir de um sorteio. Quando há um vencedor na partida, um número natural n de 1 até 6 é sorteado para determinar o número de gols feitos pelo vencedor e em seguinda, sorteia-se um número natural de 0 até $n-1$ para determinar o número de gols feitos pelo perdedor. No caso em que a partida termina empatada, é sorteado apenas um número natural de 0 até 3 para determinar o número de gols feitos pelos dois times na partida (mandante e visitante).

Vamos supor a partida São Paulo $\times$ Grêmio e que no sorteio para determinar o resultado da partida, o São Paulo seja o vencedor. Então será sorteado um número natural de 1 até 6, caso o 4 seja o número sorteado, um segundo número de 0 até 3 também será sorteado. Se este segundo número natural for o 2, o resultado da partida será São Paulo 4$\times$2 Grêmio. Se o resultado fosse empate e no sorteio de um número natural de 0 até 3 obtermos 1, o resultado seria São Paulo 1$\times$1 Grêmio. É importante a realização desse sorteio pois os critérios de desempate no campeaonato brasileiro quando dois ou mais times estão empatados em pontos são:

- maior número de vitórias;
- maior saldo de gols;
- maior números de gols feitos.

4.1.5 Probabilidades pós simulações

Vimos nas seções anteriores como são definidos os vetores força de um time e as probabilidades de cada partida, como os vetores força de cada time são atualizados após os resultados, sejam eles reais ou simulados, e como simular os resultados das partidas. Todos esses procedimentos estão definidos em um programa, criado pelo grupo de matemáticos da UFMG, que fará toda a simulação do restante do campeonato e registrará os resultados obtidos. E para que eles obtenham as probabilidades dos times ao final do campeonato é necessário que seja repetida esta simulação inúmeras vezes (geralmente cerca de 2 milhões de simulações), visando uma alta confiabilidade nos dados obtidos.

A probabilidade de um time ser campeão é obtida da frequência relativa dos resultados das simulações realizadas em que este time foi campeão.

Exemplo 4.1 *Se em* 2 *milhões de simulações o Palmeiras fosse campeão* 1 *milhão e* 600 *mil vezes a sua probabilidade de ser campeão é* 80%*, pois* $1600000/2000000 = 0,80$.

Note que o número obtido representa a frequência que certos resultados acontecem quando o experimento é repetido inúmeras vezes e não o número de combinações geradas para que o time seja campeão. Essa diferença também favorece para que a simulação computacional transcorra em um tempo bem curto.

As probabilidades de um time ser rebaixado, de classificar para a Copa Sul Americana e para a Copa Libertadores são obtidas após as simulações conforme o exemplo 4.1.

Os autores também determinam as probabilidades da pontuação mínima necessária para um time ser campeão, mas os procedimentos são diferentes. As probabilidades neste caso são obtidas do número de pontos do segundo colocado acrescido de 1 unidade, pois um time para ser campeão precisa superar o vice em pontos, ou seja, se um determinado time está em segundo com 68 pontos, o primeiro colocado deverá fazer no mínimo 69 pontos para ser campeão, claro desconsiderando outros critérios de desempate. Então a probabilidade de cada pontuação para ser campeão é a soma da frequência relativa da pontuação desejada com as frequências relativas das pontuações inferiores, isso porque se um time precisa fazer no mínimo 69 pontos para ser campeão e fizer 70 pontos ele terá uma chance ainda maior de ser campeão que será a frequência relativa de fazer 70 pontos mais a frequência relativa de fazer 69 pontos. Analogamente, podemos determinar a pontuação mínima para classificar para a Copa Libertadores e Copa Sul Americana. Já a pontuação máxima para o rebaixamento, a referência é a pontuação do 16º colocado diminuído 1 unidade e a soma deverá ser feita com a frequência relativa da pontuação desejada com as frequências relativas das pontuações superiores.

4.2 Um outro modelo para o futebol

Apresentaremos, nesta seção, um outro modelo para determinar as probabilidades de uma partida de futebol. No modelo anterior, cada time possuia dois vetores (vetor força mandante e vetor força visitante) e são usados em combinação com os vetores do adversário para obter as probabilidades da partida. Neste modelo não há vetores força para cada time e sim um número que representará sua força e esta será calculada levando em consideração três aspectos relevantes que podem interferir no resultado de uma partida: o local da partida, o atual momento técnico-emocional e o desempenho geral no campeonato.

4.2.1 Fator campo

O autor, em [11], acredita que o local onde o jogo será realizado pode interferir no resultado da partida e então representa o fator campo pela letra C, e este será composto por dois valores:

- C_m é o fator campo mandante;
- C_v é o fator campo visitante.

O fator campo mandante, C_m, é a média aritmética simples dos pontos obtidos nas partidas em que o time jogou como mandante. O fator campo visitante, C_v, é a média aritmética simples dos pontos obtidos nas partidas em que o time jogou como visitante. Cada time possuirá os dois fatores campo, C_m e C_v, mas apenas um será utilizado em cada rodada para o cálculo da força conforme o local onde será realizada a partida.

Exemplo 4.2 *Suponhamos que um time tenha feito* 14 *pontos em* 10 *partidas disputadas no seu mando de campo, então seu fator campo mandante é* $C_m = 14/10 = 1,4$.

4.2.2 Fator momento técnico-emocional

Segundo o autor, em [11], os resultados mais recentes podem influenciar no resultado da próxima partida, pois se um time vencer as últimas seis partidas chegará muito motivado e com grandes chances de vencer a próxima partida, enquanto que se perder as últimas seis partidas chegará com o moral muito baixo esperando ser derrotado mais uma vez.

O fator momento técnico-emocional, representado por M, é a média ponderada dos pontos obtidos nas seis últimas partidas de forma que a partida mais recente tem peso 6, a segunda mais recente peso 5 e analogamente até peso 1.

Esta média é calculada pelas seis últimas partidas pois este é o número médio de partidas realizadas por uma equipe no período de um mês e os pesos decrescem pois o resultado mais recente representa muito mais o momento atual do time.

Exemplo 4.3 *Considere que um time teve a seguinte sequência das últimas seis partidas em que disputou:*

1 *rodada anterior*	$\rightarrow$	*Empate;*
2 *rodadas anteriores*	$\rightarrow$	*Vitória;*
3 *rodadas anteriores*	$\rightarrow$	*Vitória;*
4 *rodadas anteriores*	$\rightarrow$	*Vitória;*
5 *rodadas anteriores*	$\rightarrow$	*Vitória;*
6 *rodadas anteriores*	$\rightarrow$	*Derrota.*

Então o seu fator momento técnico-emocional M é calculado da seguinte forma:

$$M = \frac{6 \cdot 1 + 5 \cdot 3 + 4 \cdot 3 + 3 \cdot 3 + 2 \cdot 3 + 1 \cdot 0}{6 + 5 + 4 + 3 + 2 + 1} = \frac{6 + 15 + 12 + 9 + 6 + 0}{21} = \frac{48}{21} \approx 2,3.$$

É importante ressaltar que o time recebe 3 *pontos por vitória e* 1 *ponto por empate. Caso contrário não recebe pontos.*

4.2.3 Fator desempenho global

O rendimento e qualidade técnica do time durante o campeonato também é um aspecto relevante para determinar as probabilidades do mesmo na próxima rodada, conforme o autor em [11]. Ele denomina este aspecto de fator desempenho global e será representado por G e é obtido da média de pontos conquistados por partida durante o campeonato até o momento atual.

Exemplo 4.4 *Suponhamos que um time conquistou* 22 *pontos até a* 20^a *rodada do campeonato, então seu fator desempenho global G será:*

$$G = \frac{22}{20} = 1,1.$$

4.2.4 Força

A força F que o time apresenta para a próxima partida é calculada, pelo autor, através da média ponderada dos fatores C, M e G, com os respectivos pesos 5, 4 e 1 conforme a seguinte fórmula:

$$F = \frac{5C + 4M + 1G}{5 + 4 + 1}.$$

Os pesos diferentes são critérios definidos pelo autor em [11] e como os fatores variam de 0 a 3, a força também é um número real de 0 a 3.

Exemplo 4.5 *Considere os dados abaixo de um time conforme os exemplos* 4.2*,* 4.3 *e* 4.4.

$$C_m = 1,4, \quad M = 2,3 \quad e \quad G = 1,1.$$

A força F deste time será calculada da seguinte forma:

$$F = \frac{5C + 4M + 1G}{5 + 4 + 1} = \frac{5 \cdot 1,4 + 4 \cdot 2,3 + 1 \cdot 1,1}{10} = \frac{7 + 9,2 + 1,1}{10} = \frac{17,3}{10} = 1,73.$$

4.2.5 Probabilidades da partida

As probabilidades dos três possíveis resultados (vitória do mandante, empate e vitória do visitante) são obtidas a partir das forças dos times que se enfrentarão.

O primeiro passo é determinar a probabilidade da partida terminar empatada. Claro que quando as forças dos times que se enfrentarão estão bem próximas, a chance de resultar empate é grande, mas nunca ultrapassando 50%, por mais equilibrado que seja, de acordo com o autor em [11].

A probabilidade de empate P_e é definida da metade da razão entre a força de menor valor pela força de maior valor, pode ser escrita da seguinte forma:

$$P_e = \frac{F_{menor}}{F_{maior}} \cdot 0,5.$$

As probabilidades dos três possíveis resultados juntas totalizam em 100%, e como já definimos a probabilidade de empate P_e, aplicando a propriedade 2.4.1, temos a probabilidade $1 - P_e$ que deverá ser distribuída proporcionalmente às forças dos times envolvidos. E assim teremos a probabilidade de vitória do mandante, de empate e de vitória do visitante.

Exemplo 4.6 *Suponhamos que dois times A e B possuam, respectivamente, os seguintes fatores força:* $F_A = 2,5$ *e* $F_B = 1,5$*. Vamos obter as probabilidades dos resultados possíveis da partida Time A×Time B. Primeiramente, devemos obter a probabilidade de empate* P_e *que será a metade da razão do fator força menor pelo fator força maior, então:*

$$P_e = \frac{1,5}{2,5} \cdot 0,5 = 0,6 \cdot 0,5 = 0,30 = 30\%.$$

Já sabemos que um possível empate na partida tem a probabilidade de 30%*, então a probabilidade de não ocorrer empate é de* 70%*. Devemos distribuir* 70% *proporcionalmente* *aos fatores força $F_A = 2,5$ e $F_B = 1,5$ dos times envolvidos. Logo, temos a seguinte relação:*

$$\frac{1,5+2,5}{0,7} = \frac{1,5}{P_v} = \frac{2,5}{P_m}.$$

Resolvendo as proporções, obtemos a probabilidade de vitória do mandante $P_m = 0,4375 = 43,75\%$ e a probabilidade de vitória do visitante $P_v = 0,2625 = 26,25\%$. Então as probabilidades dos possíveis resultados são: $43,75\%$ de vitória do Time A, 30% de terminar empatado e $26,25\%$ de vitória do Time B.

Este modelo encontrado em [11] explica uma forma de calcular as probabilidades das partidas que ainda serão disputadas e não para todo o campeonato, mas pode ser adaptado ao primeiro modelo que foi explicado na seção 4.1 em que um computador faz inúmeras simulações e as probabilidades do campeonato são obtidas a partir da frequência relativa do que se quer calcular.

5 Propostas de Atividades em Sala de Aula

Nesta seção iremos propor três atividades que podem ser utilizadas em sala de aula sobre a Lei dos Grandes Números e a probabilidade no futebol a fim de que os estudantes do Ensino Médio possam se habituar com o conceito fundamental da probabilidade, sejam estimulados a desenvolver o raciocínio lógico através da criação de modelos matemáticos para o futebol e que constatem a importância da aplicação da probabilidade em diversas áreas.

5.1 Atividade: Modelos matemáticos para o futebol

A atividade visa destacar a importância do uso da probabilidade no futebol apresentando um dos modelos deste trabalho de uma maneira bem dinâmica para que os alunos possam discutir novas ideias para a construção de um novo modelo.

Primeiramente, o professor deverá detalhar um dos modelos aos seus alunos explicando todos os processos de como obter as probabilidades das partidas, preferencialmente, da próxima rodada que irá acontecer em breve. Por isso sugerimos o segundo modelo pois ele não necessita dos resultados de todas as partidas anteriores, e sim, apenas dos resultados das seis últimas rodadas e os dados da classificação do campeonato naquele momento.

A próxima etapa o professor deverá solicitar que os alunos formem uns quatro ou cinco grupos e a partir do modelo apresentado criar outras formas de obter a força de cada time e consequentemenete, determinar as probabilidades dos times nas próximas partidas. Após esta criação, eles deverão apresentar o modelo criado aos outros grupos e o professor poderá aproveitar este momento para conversar sobre os conceitos de probabilidades e as Leis dos Grandes Números, comentando sobre as simulações que são realizadas através de um programa e que determinam as probabilidades dos times para serem campeões ou para outras situações.

A escolha da próxima rodada é ideal para que, ao final dos modelos construídos e apresentados, assim que as partidas ocorrerem na realidade, os alunos possam comparar os resultados obtidos e estes servindo como pauta de discussão em sala de aula criando oportunidade para o professor esclarecer a diferença entre “chance de ocorrer” e a “ocorrência real”.

5.2 Atividade: Estimando eventos

A atividade tem como objetivo envolver os alunos com os cálculos de probabilidade e estimativa de uma maneira lúdica. O professor deverá preparar o material previamente, providenciando urnas contendo 10 bolas de isopor, sendo 3 vermelhas e 7 pretas, de mesmo peso e dimensão.

No início da atividade, o professor deverá separar os alunos em grupos contendo 3 ou 4 alunos. Em seguida informará aos alunos que cada urna contêm 10 bolas, algumas vermelhas e outras brancas, mas sem informar essa quantidade, pois eles deverão estimar qual a quantidade de bolas vermelhas e brancas após realizadas várias repetições do experimento.

O professor deverá explicar também que para realizar o experimento, um aluno deve retirar uma bola da urna, registrar qual a cor obtida numa folha ou no caderno e recolocar a bola na urna. Este experimento deverá ser realizado pelo menos 30 vezes.

Os alunos, após os 30 sorteios, deverão observar os resultados obtidos e estimar o número de bolas vermelhas e brancas que estão dentro da urna. Neste momento, o professor deverá observar quantos acertaram ou erraram e aproveitar para discutir sobre os resultados obtidos pelos alunos e comparar a atividade à pesquisa eleitoral.

O professor, caso queira, pode realizar a atividade novamente aumentando ou alterando o número de bolas ou até mesmo incluir bolas de outras cores.

5.3 Atividade: Repetindo experimentos

Esta atividade procura contribuir para um melhor entendimento dos alunos sobre a ideia principal da Lei dos Grandes Números e os cálculos de probabilidades. Um experimento simples deve ser escolhido pelo professor para ser utilizado nesta atividade. Os experimentos simples sugeridos são: lançamento de uma moeda, lançamento de um dado ou retirada de uma carta entre outras. Explicaremos a atividade com a utilização do lançamento de uma moeda.

Várias duplas deverão ser formadas e cada uma receberá uma moeda. Cada dupla deverá lançar a moeda 50 vezes e anotar o resultado obtido em cada lançamento.

Após estes 50 lançamentos, as duplas deverão obter a razão do número de coroas em relação ao total de lançamentos realizados e informar ao professor.

O professor deverá anotar a razão obtida por cada dupla no quadro, analisar estes resultados junto com seus alunos e enfatizar que a maioria dos resultados devem ser diferentes de $1/2$, mas que estão bem próximos deste valor e que em pouquíssimos casos teremos resultados mais afastados. O professor também poderá juntar os dados obtidos pelas duplas, obter uma nova razão e mostrar que ela está mais próxima de $1/2$, enfatizando a ideia principal da Lei dos Grandes Números, que quanto maior for o número de repetições mais próximo estará da probabilidade real.

Considerações Finais

Observamos que a probabilidade pode ser aplicada à diversas áreas mas para que isso seja perceptível aos alunos do Ensino Médio é necessário um aprofundamento no conteúdo, ao contrário do que é realizado atualmente nas escolas.

A escolha pela aplicação da probabilidade no futebol neste trabalho se deve ao grande interesse e questionamento dos alunos em saber como são calculadas as probabilidades de um time ser campeão e, consequentemente, pela dificuldade apresentada pelos professores em explicar ou descrever como são obtidos estes cálculos por nunca terem lido algo a respeito. Acreditamos que esta dificuldade dos professores é justificada pela ausência de materiais disponíveis para este estudo, pois em muitos casos não são revelados os detalhes de todo o processo para obter estas probabilidades, e poucos divulgam estas informações como é o caso do Departamento de Matemática da UFMG.

Esperamos que este trabalho sirva de parâmetro para o desenvolvimento de outros trabalhos e que seja uma referência no tratamento deste tema contribuindo para excelência nas aulas de muitos professores do Ensino Médio.

Em trabalhos futuros, pretende-se desenvolver um novo modelo para o cálculo de probabilidades no futebol utilizando uma programação própria, pois esta foi a principal dificuldade apresentada neste trabalho, e ampliar os estudos sobre a Lei dos Grandes Números.

Agradecimentos

Gostaria de agradecer primeiramente a Deus pois sem ele nada disso seria possível, aos meus pais pela paciência por me aguentar neste período de muito trabalho, pelo incentivo aos estudos desde criança e pelas orações. Aos meus amigos Fred e Roberto pela compreensão da minha ausência nestes anos e pela contribuição neste trabalho.

Quero agradecer aos meus professores que contribuíram para a minha vida acadêmica, aos professores e coordenadores da UFSJ/CAP, ao professor Gilcione da UFMG e principalmente ao orientador Humberto que me encorajou para a conclusão deste trabalho.

Agradeço aos meus colegas do PROFMAT, parceiros de todas as sextas-feiras, que lutaram comigo em busca de um aprimoramento profissional, aos que riram das minhas piadas durante as viagens e aos que participaram do grupo de estudo no hotel se preparando para as provas. Em especial, à Francilene minha parceira dos exercícios e ao grande amigo Anderson que me ajudou demais com seu companheirismo, com a sua humildade e prestatividade.

Agradeço à CAPES pelo apoio financeiro e pela oportunidade da realização deste sonho.

E por fim, um agradecimento mais que especial à minha namorada Renata que vivenciou e me deu forças em todos os momentos de angústia, ansiedade, lamúria e fraqueza, acreditando sempre no meu potencial e orando pelo meu sucesso.

Referências

[1] Regras de Futebol. Disponível em <http://www.cbf.com.br/arbitragem/regras-futebol-e-livros/livro-de-regras-de-futebol-2016-2017-portugues#.WG5gIVMrLIU>. Acesso em 18 Dez. 2016.

[2] Portal Action. Disponível em <http://www.portalaction.com.br/probabilidades>. Acesso em 20 Out. 2016.

[3] Chance de Gol. Disponível em < http://chancedegol.uol.com.br/comofunciona.htm >. Acesso em 18 Set. 2016.

[4] Infoescola. Disponível em < http://www.infoescola.com/matematica/probabilidade-em-estatistica-e-lei-dos-grandes-numeros/ >. Acesso em 17 Ago. 2016.

[5] As Leis dos Grandes Números. Disponível em < http://ferrari.dmat.fct.unl.pt/personal/mle/DocPE/PE0809/Docmts/LGN-0708.pdf >. Acesso em 13 Ago. 2016.

[6] Probabilidade2. Disponível em < http://www.de.ufpb.br/˜tarciana/Probabilidade2/ >. Acesso em 05 Out. 2016.

[7] Introdução aos processos estocásticos. Disponível em < http://www.cpdee.ufmg.br/˜emmendes/teoremas.pdf >. Acesso em 20 Ago. 2016.

[8] Sites de Apostas. Disponível em < http://www.sites-de-apostas.net/estrategia-como-determinar-probabilidades-para-um-jogo-de-futebol.php >. Acesso em 30 Abr. 2016.

[9] Uol apoio escolar. Disponível em < http://clickeaprenda.uol.com.br/portal/mostrarConteudo.php?idPagina=28938 >. Acesso em 22 Out. 2016.

[10] Probabilidades no futebol. Disponível em < http://www.mat.ufmg.br/futebol/#hero >. Acesso em 25 Nov. 2016.

[11] L. M. W. d. ALMEIDA, J. d. L. ARAÚJO, and E. BISOGNIN. *Práticas de modelagem matemática na educação matemática*. SciELO-EDUEL, 2011.

[12] J. BERNOULLI. *Ars conjectandi*. Impensis Thurnisiorum, fratrum, 1713.

[13] C. A. B. DANTAS. *Probabilidade: Um Curso Introdutório*. Edusp, 2008.

[14] J. R. FONSECA. *Estatística matemática Vol.1*. 2001.

[15] B. R. JAMES. *Probabilidade: um curso em nível intermediário*. Number 519.2. IMPA, 2015.

[16] L. K. JUNIOR. Probabilidade geométrica e grandes números. 2013.

[17] B. N. LIMA, L. M. CIOLETTI, M. d. O. T. CUNHA, and G. A. BRAGA. Entropia: introdução à teoria matemática da (des) informação. 2012.

[18] B. N. LIMA, G. N. COSTA, R. NACIFE, R. V. MARTINS, and R. GUIMARÃES. Probabilidades no esporte.

[19] P. L. MEYER. Probabilidade: aplicações à estatística. In *Probabilidade: aplicações à estatística*. Livro Técnico, 1970.

[20] P. RATHIE, Pushpa N. ; ZORNIG. *Teoria da Probabilidade*. Editora UnB, 2012.

[21] M. R. SPIEGEL. *Probabilidade e Estatística-: Coleção Schaum*. Bookman Editora, 1978.

[22] J. H. VUOLO. *Fundamentos da teoria dos erros*. E. Blucher, 1996.