Universidade Federal do Rio Grande do Norte
Centro de Ciências Exatas e da Terra
Programa de Pós-Graduação em Matemática - PROFMAT
Mestrado Profissional em Matemática em Rede Nacional

Roberto Fagner Raquel

Princípios de Probabilidade

Natal, dezembro de 2014.

Roberto Fagner Raquel

Princípios de Probabilidade

Dissertação de Mestrado submetida ao Programa de Pós-Graduação em Matemática em Rede Nacional (PROFMAT) da Universidade Federal do Rio Grande do Norte como parte dos requesitos para a obtenção do grau de Mestre em Matemática.

Área de Concentração: Probabilidade.

Orientadora:
Prof^a. Dr^a. Débora Borges Ferreira

Natal, dezembro de 2014.

Catalogação da Publicação na Fonte. UFRN / SISBI / Biblioteca Setorial
Centro de Ciências Exatas e da Terra – CCET.

Raquel, Roberto Fagner.
Princípios de probabilidade / Roberto Fagner Raquel. - Natal, 2014.
48 f. : il.

Orientadora: Profa. Dra. Débora Borges Ferreira.

Dissertação (Mestrado) – Universidade Federal do Rio Grande do Norte. Centro de Ciências Exatas e da Terra. Programa de Pós-Graduação em Matemática em Rede Nacional.

1. Probabilidade – Dissertação. 2. Aleatório – Dissertação. 3. Equiprovável – Dissertação. I. Ferreira, Débora Borges. II. Título.

RN/UF/BSE-CCET CDU: 519.2

Roberto Fagner Raquel

Princípios de Probabilidade

Dissertação de Mestrado submetida ao Programa de Pós-Graduação em Matemática em Rede Nacional (PROFMAT) da Universidade Federal do Rio Grande do Norte como parte dos requesitos para a obtenção do grau de Mestre em Matemática.

Aprovado em: 08 / 12 / 2014

Banca Examinadora:

Débora Borges Ferreira

Prof^a. Dr^a. Débora Borges Ferreira
Departamento de Matemática - UFRN
Orientadora

Fabiana T. Santana

Prof^a. Dr^a. Fabiana Tristão de Santana
Escola de Ciências e Tecnologia - UFRN
Examinador Interno

W Batista Santos

Prof^o. Dr^o. Walter Batista dos Santos
UFG
Examinador Externo

Dedicatória

Ao meu querido amigo, professor e ídolo, Ântonio Roberto da Silva, um exemplo a ser seguido.

Agradecimentos

À professora Débora Borges Ferreira, por encorajar a realização deste trabalho, sua ajuda foi fundamental.

Aos professores do curso e aos colegas, em especial a Almir, Cláudio, Márcio e Venício, companheiros de estudos e viagens.

Aos meus pais, Antonieta e Erivaldo, e irmãos, Alexandre, Eduardo e Priscila, pelo apoio. Aos compadres Bianca, Juciano, Paulo e Marília, pelas acolhidas.

À minha esposa Josemária e ao meu filho Thiago, por tornarem minha vida completa.

“Os ensinamentos das pessoas sábias são uma fonte de vida.”

Provérbios 13.14

Resumo

Historicamente, as Teorias da Probabilidade começaram com os estudos de jogos de azar, como a roleta e as cartas. Atualmente, a Probabilidade tem aplicações em diversas ciências, tais como Economia, Física e Química. Os conceitos e aplicabilidade da Probabilidade são ensinados cada vez mais no Ensino Médio. Normalmente, é no ensino da Análise Combinatória que começa a se estabelecer os conceitos e as noções básicas para a compreensão do cálculo de probabilidades. Este trabalho apresenta definições básicas e necessárias para o entendimento do assunto, assim como uma exploração da teoria na visão dos autores Barry James e Marcos Nascimento Magalhães. Para finalizar, apresentamos uma sugestão de aula prática do assunto para o ensino médio.

Palavras-chave: Aleatório, Probabilidade, Equiprovável.

Abstract

Historically, Probability theories began with studies of gambling such as roulette and cards. Currently, the Probability has applications in various sciences, such as Economics, Physics and Chemistry. The concepts and applicability of Probability are taught increasingly in high school. It's usually in the teaching of Combinatorics that the concepts and the basics for understanding the calculus of probabilities begin to establish. This paper presents basic and necessary definitions for understanding the subject, as well as an exploration of the theory based on the views of the authors Barry James and Marcos Nascimento Magalhães. Finally, we present a suggestion for a high school practical class on the subject.

Keywords: Random, Probability, Equiprobable.

Sumário

Introdução

Diariamente acompanhamos a previsão do tempo. Seja na televisão, nos jornais ou na internet (computador, celular, tablet, etc.) buscamos orientações de como será o tempo para que possamos melhor conduzir nosso dia. Quais roupas e acessórios usar, qual caminho e o melhor horário para trafegar, etc. Mas, a previsão do tempo, como o próprio nome diz, é uma "previsão", ou seja, algo que se estima que vá acontecer, isso porque as condições climáticas não podem ser estabelecidas com total certeza.

A incerteza está presente em nossas vidas nas mais diversas formas: número de pessoas que iremos encontrar durante o dia, quanto tempo ficaremos parados em semáforos, pesquisas eleitorais, crescimento da inflação, oscilações da bolsa de valores e mais uma grande quantidade de outras situações em que não temos certeza quanto à ocorrência ou não das possíveis alternativas. Essas incertezas nos levam a falar sobre probabilidade.

O interesse humano em estudar as possibilidades de acontecer ou não um fenômeno fez surgir um ramo da matemática aplicada chamado de Probabilidade. A palavra probabilidade deriva do latim *probare* (provar ou testar). A teoria da probabilidade tenta quantificar a noção de provável. Mesmo hoje em dia, antes de realizar um jogo ou uma aposta, pensamos nas chances que temos de ganhar, ou seja, na probabilidade de vencermos. Quem não sonha em ganhar na mega sena? Mesmo o jogador que faz apenas um jogo com seis números tem esperança de vencer, apesar de sua chance ser de apenas 1 em 50 milhões. Essas "chances" estão relacionadas a fenômenos que não estão sob nosso controle, fenômenos chamados de aleatórios.

Historiadores afirmam que a Probabilidade teve origem no século XVII, quando matemáticos começaram a estudar alguns jogos de azar e relacionar as apostas com os possíveis resultados. Nos séculos seguintes foram publicados trabalhos que alicerçaram o assunto, e a Teoria das Probabilidades cresceu muito além das fronteiras dos jogos de azar, buscando estimar as chances de ocorrer um determinado acontecimento, elaborando e pesquisando modelos para estudar experimentos ou fenômenos aleatórios.

Segundo Boyer [3] e Cramér [5], as primeiras considerações matemáticas acerca dos jogos e apostas são atribuídas a Pascal e Fermat (séc. XVII), que alicerçaram a teoria do cálculo das probabilidades e da análise combinatória, relacionando as apostas no jogo de dados com os possíveis resultados.

Berlinghoff [1] afirma que no final do séc. XVIII, o matemático francês Laplace começou a trabalhar com probabilidade, escrevendo uma série de artigos sobre o assunto. Em 1812, publicou a *Teoria Analítica das Probabilidades*, onde reuniu tudo que ele e outros tinham estudado sobre o assunto. Mas, apenas aproximadamente no início do séc. XX, Kolmogorov desenvolveu uma teoria matemática rigorosa, baseada em axiomas, definições e teoremas (ver mais em Garding [7]).

Este trabalho será realizado com a intenção de trazer contribuições para o ensino das teorias da probabilidade no Ensino Médio, direcionado a professores que queiram entender ou relembrar um pouco mais das formalidades desse assunto e aplicá-las em sala de aula. Buscamos trabalhar os conceitos com exemplos simples e do cotidiano, trabalhando com experimentos equiprováveis e não equiprováveis.

O tema proposto foi estudado por autores, com diferentes abordagens e metodologias. O trabalho de César Augusto Vieira Lima [10], intitulado "Probabilidade para o Ensino Médio", aplica os conhecimentos de probabilidades na resolução de problemas do cotidiano. Neste trabalho, que tem enfoque nos casos equiprováveis, ele resolve, de forma clara e de fácil compreensão, problemas envolvendo o conceito de probabilidade, por exemplo: como ganhar na mega-sena, coincidência de aniversários, probabilidade do amigo oculto, etc. Nosso trabalho trata, de forma diferencial, dos casos equiprováveis e também dos não equiprováveis.

Já o estudo de Cileda de Queiroz e Silva Coutinho [4], intitulado "Introdução ao Conceito de Probabilidade por uma Visão Frequentista", estuda as concepções espontâneas ou pré-construídas dos alunos à propósito do acaso e de probabilidades, a partir da observação da estabilização da frequência relativa de um evento após um grande número de repetições da experiência aleatória. Segundo a autora, o objetivo final da escolha frequentista é estender a noção de probabilidades às situações não somente de “casos igualmente prováveis”. Como objetivo didático, busca ligar o ensino às condições de aprendizagem nas quais o aluno está inserido. É válido ressaltar que a visão frequentista é não equiprovável.

Nosso trabalho é de natureza bibliográfica, no sentido de aprofundar conhecimentos, e está dividido em três capítulos. No primeiro capítulo veremos a definição de aleatoriedade, assim como de espaço amostral, eventos e variáveis aleatórias. Também vamos estudar as classes de conjuntos sobre as quais aplicaremos a definição de função de probabilidade chamadas de álgebra e σ-álgebra, baseados nas obras de James [9] e Magalhães [11].

No capítulo seguinte vamos nos aprofundar um pouco na definição da função de probabilidade e suas propriedades, estudando também probabilidade condicional, independência de eventos e função de distribuição. Ainda baseados em James [9] e Magalhães [11], definiremos a probabilidade como uma função que satisfaz os axiomas de Kolmogorov, em uma σ-álgebra. Para explorar o assunto, demonstramos algumas propriedades e exemplos.

Para finalizar, apresentamos uma sugestão de aula prática de probabilidade para o ensino médio, abordando os principais conceitos e a função de probabilidade para espaços amostrais equiprováveis e não equiprováveis.

Capítulo 1

Definições Básicas

1.1 Aleatoriedade

As palavras "incerto", "azar", "sorte", "risco", "talvez", "duvidoso", "acaso", "indeterminado", "provável", "imprevisível" e "aleatório" figuram diariamente em todos os meios de comunicação. Tais palavras estão relacionadas a fenômenos que não podemos determinar com exatidão seus resultados.

Definição 1.1 *Usamos a palavra aleatoriedade para designar algo que depende de um acontecimento sujeito às incertezas do acaso. Fenômenos ou experimentos aleatórios são aqueles que, mesmo que tenham sido observados ou repetidos sob as mesmas condições, produzem resultados que não podem ser previstos com certeza.*

Exemplo 1.1

i) E_1: Lançar uma moeda honesta e observar a face voltada para cima.

ii) E_2: Retirar uma bola ao acaso de uma urna com 10 bolas numeradas de 1 a 10 e observar o número sorteado.

iii) E_3: Lançar um dado e observar a face superior.

iv) E_4: Em uma linha de produção, fabrique peças em série e conte o número de peças defeituosas produzidas em um período de 1 hora.

v) E_5: Em uma linha de produção, escolher uma peça e verificar sua funcionalidade.

vi) E_6: De uma gaveta contendo meias de diversas cores retirar, sem enxergar, um par de meias e verificar se têm a mesma cor.

vii) E_7: Um número natural é escolhido ao acaso.

viii) E_8: Um número é escolhido aleatoriamente no intervalo (0, 1).

Cada um desses experimentos tem um conjunto de possíveis resultados. Esse conjunto é o *espaço amostral*.

1.2 Espaço Amostral

Definição 1.2 *Denominamos espaço amostral associado a um experimento aleatório ao conjunto de todos os resultados possíveis deste experimento. Ele será representado pela letra grega* Ω *(ômega). Seus subconjuntos são denominados eventos e o conjunto vazio (conjunto sem elementos) é denotado por* $\emptyset$.

O espaço amostral pode ser finito, infinito enumerável ou infinito não enumerável, dependendo do experimento aleatório estudado. Cada resultado possível do experimento é denominado *ponto* ou *elemento* de Ω e denotado genericamente por ω.

Vamos considerar os exemplos anteriores (Exemplo 1.1) e descrever o espaço amostral para cada um deles. O espaço Ω_i refere-se ao exemplo E_i.

Exemplo 1.2

i) $\Omega_1 = \{\text{cara, coroa}\}$.

ii) $\Omega_2 = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$.

iii) $\Omega_3 = \{1, 2, 3, 4, 5, 6\}$.

iv) $\Omega_4 = \{0, 1, 2, ..., n\}$, onde n é o número máximo de peças que podem ser produzidas em 1 hora.

v) $\Omega_5 = \{\text{boa, defeituosa}\}$.

vi) $\Omega_6 = \{\text{cores iguais, cores diferentes}\}$.

vii) $\Omega_7 = \{1, 2, 3, 4, ...\}$.

viii) $\Omega_8 = (0, 1)$.

Os espaços amostrais observados em S_1, S_2, S_3, S_4, S_5 e S_6 são finitos, S_7 é infinito enumerável e S_8 é infinito não enumerável, porque é um intervalo da reta real, logo contém infinitos números e como contém os irracionais, dizemos que é não enumerável.

1.3 Eventos

Definição 1.3 *Os eventos são os subconjuntos do espaço amostral. São representados pelas letras latinas maiúsculas do início do alfabeto, como A, B, etc.*

A seguir alguns exemplos de eventos. Novamente, vamos nos referir aos exemplos anteriores (Exemplo 1.1): o evento A_i se referirá ao exemplo E_i.

Exemplo 1.3

i) $A_1 = \{\text{cara}\}$, isto é, ocorre cara.

ii) $A_2 = \{2, 3, 5, 7\}$, isto é, ocorre um número primo.

iii) $A_3 = \{1, 2, 3, 4\}$, isto é, ocorre um número menor que cinco.

iv) $A_4 = \{1\}$, isto é, apenas uma peça foi defeituosa.

v) $A_5 = \{\text{boa}\}$, isto é, a peça escolhida é boa.

vi) $A_6 = \{\text{cores iguais}\}$, isto é, o par de meias escolhido tem a mesma cor.

vii) $A_7 = \{1, 3, 5, 7, ...\}$, isto é, o número escolhido é ímpar.

viii) $A_8 = \left(0, \frac{1}{2}\right)$, isto é, o número escolhido é maior que zero e menor que um meio.

Segundo James [9], um evento A ao qual atribuímos probabilidade será chamado *evento aleatório*. O autor afirma que nem todo evento tem uma probabilidade, mas esse fato não nos causará problemas, pois é possível restringir a atenção a alguns eventos.

1.4 Variáveis Aleatórias

Quando realizamos um experimento aleatório buscamos quantidades, valores que exprimam as probabilidades de ocorrência dos diversos eventos possíveis. Quando conhecemos o espaço amostral fica mais fácil avaliar a probabilidade de cada evento.

Por exemplo, suponha que sejam dispostas em uma mesa quatro folhas de papel, cada uma com um número desconhecido escrito no lado sob a mesa. Quais as chances de desvirarmos uma folha e o número ser par? Se não conhecemos os números que estão nas folhas (espaço amostral), não podemos responder a essa pergunta. Supondo que sabemos que os números escritos são $\{1, 2, 3, 4\}$, então a resposta à pergunta seria 2 chances em 4, ou seja, 1/2.

A grande maioria dos experimentos aleatórios de interesse possuem eventos aleatórios abstratos: chove ou não chove, cara ou coroa, meias iguais ou distintas... Para tornar matematicamente tratável esses estudos, precisamos associar o "abstrato" a números reais. Faremos isso atribuindo funções com domínio no espaço amostral e imagem nos reais e esta função pode até ser a identidade.

Por exemplo, se o experimento é lançar um dado, então podemos associar a face obtida no lançamento ao número que aparece nela. Assim, temos uma função com domínio no espaço amostral (faces do dado) e imagem nos reais $\{1, 2, 3, 4, 5, 6\}$.

Definição 1.4 *Variável aleatória é uma função definida num espaço amostral, que assume valores reais. É um característico numérico do resultado de um experimento.*

Da definição anterior, entende-se que variável aleatória é uma variável quantitativa, cujo resultado depende de fatores aleatórios.

Exemplo 1.4.1 Dois jogadores A e B lançam uma moeda honesta. Se sair cara, o jogador A ganha 10 reais. Se sair coroa, o jogador B ganha 10 reais. Quando um jogador ganha, o outro é quem paga.

Temos $\Omega = \{\text{cara, coroa}\}$ e podemos definir duas variáveis aleatórias relacionadas a esse experimento:

$$\begin{array}{rccc} x_A: & \Omega & \to & \mathbb{R} \\ & \text{cara} & \mapsto & +10 \\ & \text{coroa} & \mapsto & -10 \end{array}$$

e

$$\begin{array}{rccc} x_B: & \Omega & \to & \mathbb{R} \\ & \text{cara} & \mapsto & -10 \\ & \text{coroa} & \mapsto & +10 \end{array}$$

Como observamos nesse exemplo, as variáveis aleatórias podem ser convenientemente escolhidas em cada situação.

Os exemplos a seguir estão relacionados com os anteriores (Exemplo 1.1) e servem como exemplos de possíveis variáveis aleatórias.

Exemplo 1.4.2

i) V_1: $\{$cara, coroa$\} \to \{0, 1\}$, isto é, se sair cara, associa ao 0, se sair coroa, ao 1.

ii) V_2: $\{$10 bolas numeradas de 1 a 10$\} \to \{1, 2, ..., 10\}$, associa cada bola com o número que aparece nela.

iii) V_3: $\{$número obtido na face do dado$\} \to \{1, 2, 3, 4, 5, 6\}$, associa o número obtido na face do dado a ele mesmo.

iv) V_4: Associa o número de peças produzidas em 1 hora a ele mesmo.

v) V_5: $\{$boa, defeituosa$\} \to \{0, 1\}$, retira-se uma peça, associa a ela o número 1 se for defeituosa e 0 se for boa.

vi) V_6: $\{$cores iguais, cores diferentes$\} \to \{0, 1\}$, associa ao 0 se o par for de cores iguais e ao 1 se for de cores diferentes.

vii) V_7: Associa o número escolhido a ele mesmo.

viii) V_8: Idem ao V_7.

Mais uma vez, a função da variável aleatória é a partir do resultado abstrato de um experimento (meias, peças, face da moeda, ...) associar números reais para podermos resolver problemas com uma roupagem matemática.

É importante ressaltar que a Definição 1.4 não é a rigor correta, é apenas intuitiva. A seguir apresentaremos a definição rigorosa de variáveis aleatórias. Antes, é preciso entender os conceitos de classes de eventos.

1.5 Classes de Eventos

1.5.1 Álgebra

Essa seção tem como base as obras de James [9] e Magalhães [11]. Posteriormente, veremos a relação entre variáveis aleatórias e probabilidade, e para definir esta, segundo os autores, é preciso entender as definições que seguem.

Do ponto de vista rigorosamente matemático temos que Ω é um conjunto abstrato de pontos. Vamos estabelecer uma classe de subconjuntos de Ω com algumas propriedades convenientes. Essas propriedades serão essenciais para o posterior desenvolvimento da teoria e do cálculo de probabilidades.

Definição 1.5.1 *Seja $\mathcal{F}$ uma classe não vazia de subconjuntos de Ω que satisfaz às seguintes propriedades*

P_1. $\Omega \in \mathcal{F}$;

P_2. Se $A \in \mathcal{F}$, então $A^C \in \mathcal{F}$;

P_3. Se $A \in \mathcal{F}$ e $B \in \mathcal{F}$, então $(A \bigcup B) \in \mathcal{F}$.

Então $\mathcal{F}$ é denominada de álgebra de subconjuntos de Ω.

Exemplo 1.5.1.1 Dado um espaço amostral Ω, seja $\mathcal{F} = \{\emptyset, \Omega\}$. É fácil observar que $\mathcal{F}$ satisfaz às propriedades P_1, P_2 e P_3, logo $\mathcal{F}$ é uma álgebra de subconjuntos de Ω.

Exemplo 1.5.1.2 Dado um espaço amostral $\Omega = \{0, 1\}$, seja $\mathcal{F} = \{\emptyset, \Omega, \{0\}, \{1\}\}$. P_1 está satisfeita. Como $\emptyset^C = \Omega$, $\Omega^C = \emptyset$, $\{0\}^C = \{1\}$ e $\{1\}^C = \{0\}$, P_2 também está satisfeita. Para verificar P_3 basta observar que $\{0\} \bigcup \{1\} = \Omega \in \mathcal{F}$, pois a união com

$\emptyset$ é irrelevante e a união com Ω dá o próprio Ω, logo P_3 está satisfeita e a classe $\mathcal{F}$ de subconjuntos de Ω é uma álgebra de subconjuntos de Ω.

Propriedade 1.5.1 Toda álgebra satisfaz

P_4. $\emptyset \in \mathcal{F}$;

P_5. Se $A_i \in \mathcal{F}$, $i \geq 1$, então $\bigcup\limits_{i=1}^{n} A_i \in \mathcal{F}$ e $\bigcap\limits_{i=1}^{n} A_i \in \mathcal{F}$.

Demonstração: P_4 é verdade pois, por P_1, $\Omega \in \mathcal{F}$, e por P_2, $\Omega^C = \emptyset \in \mathcal{F}$. Provaremos P_5 por indução. Primeiramente, mostraremos que se $A_i \in \mathcal{F}$, $i \geq 1$, então $\bigcup\limits_{i=1}^{n} A_i \in \mathcal{F}$.

Vejamos que é verdade para $n = 2$, pois sejam $A_1, A_2 \in \mathcal{F}$, então $\bigcup\limits_{i=1}^{2} A_i = \left(A_1 \bigcup A_2\right) \in \mathcal{F}$ (pela propriedade P_3). Portanto, a propriedade é verdadeira para $n = 2$.

Suponha que a propriedade seja verdadeira para n, vamos verificar para $n+1$. Sejam $A_1, A_2, ..., A_n, A_{n+1} \in \mathcal{F}$, então $\bigcup\limits_{i=1}^{n+1} A_i = \left(A_1 \bigcup A_2 \bigcup ... \bigcup A_n \bigcup A_{n+1}\right) = \left(A_1 \bigcup A_2 \bigcup ... \bigcup A_n\right) \bigcup A_{n+1} \in \mathcal{F}$ (pela propriedade P_3, pois estamos considerando que $\left(A_1 \bigcup A_2 \bigcup ... \bigcup A_n\right) \in \mathcal{F}$). Portanto, a propriedade é verdadeira para $n + 1$.

Para finalizar, mostraremos a segunda parte, que é se $A_i \in \mathcal{F}$, $i \geq 1$, então $\bigcap\limits_{i=1}^{n} A_i \in \mathcal{F}$.

Para $n = 2$, sejam $A_1, A_2 \in \mathcal{F}$, então $A_1^C, A_2^C \in \mathcal{F}$ (por P_2), $\bigcup\limits_{i=1}^{2} A_i^C = \left(A_1^C \bigcup A_2^C\right) \in \mathcal{F}$ (por P_3) e $\left(A_1^C \bigcup A_2^C\right)^C \in \mathcal{F}$ (novamente por P_2). Agora, basta observar que $\left(\bigcup\limits_{i=1}^{2} A_i^C\right)^C = \left(A_1^C \bigcup A_2^C\right)^C = A_1 \bigcap A_2 = \bigcap\limits_{i=1}^{2} A_i \in \mathcal{F}$. Portanto, a propriedade é verdadeira para $n = 2$.

Suponha que a propriedade seja verdadeira para n, ou seja, $A_1 \bigcap A_2 \bigcap ... \bigcap A_n \in \mathcal{F}$, vamos verificar para $n+1$. Sejam $A_1, A_2, ..., A_n, A_{n+1} \in \mathcal{F}$, então $A_1^C, A_2^C, ..., A_n^C, A_{n+1}^C \in \mathcal{F}$ (por P_2), $\bigcup_{i=1}^{n+1} A_i^C = A_1^C \bigcup A_2^C \bigcup ... \bigcup A_n^C \bigcup A_{n+1}^C \in \mathcal{F}$ (pela primeira parte de P_5) e $\left(\bigcup_{i=1}^{n+1} A_i^C\right)^C = (A_1^C \bigcup A_2^C \bigcup ... \bigcup A_n^C \bigcup \ A_{n+1}^C)^C \in \mathcal{F}$ (novamente por P_2). Agora, temos que $\left(A_1^C \bigcup A_2^C \bigcup ... \bigcup A_n^C \bigcup A_{n+1}^C\right)^C = A_1 \bigcap A_2 \bigcap ... \bigcap A_n \bigcap A_{n+1} = \bigcap_{i=1}^{n+1} A_i \in \mathcal{F}$. Portanto, a propriedade é verdadeira para $n+1$.

As duas partes da propriedade são verdadeiras para $n = 2$ e provamos que a validez de n implica a validez de $n+1$, então a propriedade é válida para todo número natural $n \geq 2$.

Da propriedade P_5, podemos concluir que uma álgebra é fechada para um número finito de uniões e interseções. Agora, considere o exemplo a seguir:

Exemplo 1.5.1.3 Seja $\Omega = \{1, 3, 5, 7, ...\}$ o conjunto dos números naturais ímpares e $\mathcal{F}$ o conjunto das partes de Ω, ou seja, o conjunto de todos os subconjuntos de Ω.

Em alguns casos, como no exemplo anterior, pode ser insuficiente tomar $\mathcal{F}$ como uma álgebra de subconjuntos de Ω, pois há eventos que se repetem infinitamente e precisamos descrevê-los. Pode ser necessário que $\mathcal{F}$ seja fechada sob um número infinito enumerável de operações da teoria dos conjuntos. Precisamos de uma classe de subconjuntos mais restrita, denominada *σ-álgebra*.

1.5.2 σ-Álgebra

Uma *σ-álgebra* é um caso particular de álgebra, então são válidas as mesmas propriedades, mas com uma modificação em P_3, a união deve ser verificada de forma a que todas as possíveis uniões venham a pertencer ao conjunto.

Definição 1.5.2.1 *Seja $\mathcal{F}$ uma classe não vazia de subconjuntos de Ω que satisfaz as seguintes propriedades*

P_1. $\Omega \in \mathcal{F}$;

P_2. Se $A \in \mathcal{F}$, então $A^C \in \mathcal{F}$;

P_3'. Se $A_n \in \mathcal{F}$ para $n = 1, 2, 3, ...$, então $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$.

Então $\mathcal{F}$ é denominada de σ-álgebra de subconjuntos de Ω.

Propriedade 1.5.2 Seja $\mathcal{F}$ uma σ-álgebra de subconjuntos de Ω. Se $A_n \in \mathcal{F}$ para $n = 1, 2, 3, ...$, então $\bigcap_{n=1}^{\infty} A_n \in \mathcal{F}$.

Demonstração: Pela Definição 1.5.2.1, se $\mathcal{F}$ é uma σ-álgebra e $A_n \in \mathcal{F}$ para $n = 1, 2, 3, ...$, então $A_n^C \in \mathcal{F}$ para $n = 1, 2, 3, ...$, e, consequentemente $\bigcup_{n=1}^{\infty} A_n^C \in \mathcal{F}$. Novamente, pela propriedade do complementar temos que $\left(\bigcup_{n=1}^{\infty} A_n^C\right)^C \in \mathcal{F}$.

Assim, $\bigcap_{n=1}^{\infty} A_n = \left(\bigcup_{n=1}^{\infty} A_n^C\right)^C \in \mathcal{F}$, concluindo a demonstração.

Uma σ-álgebra de subconjuntos de Ω é fechada para um número enumerável de operações de conjuntos. Toda σ-álgebra é uma álgebra, pois P_3 é consequência de P_3', já que $A \bigcup B = A \bigcup B \bigcup B \bigcup B ... \in \mathcal{F}$, se $\mathcal{F}$ é σ-álgebra.

A "menor" σ-álgebra gerada por um subconjunto é chamada de σ-álgebra de Borel. Quando definida em $\mathbb{R}$ é denotada por $\mathcal{B}(\mathbb{R})$ e contém todos os intervalos abertos dos reais. Seus elementos são os **conjuntos de Borel** ou **borelianos** (ver mais em James [9]).

Exemplo 1.5.2.1 Considere $\Omega = \{a, b, c\}$, onde $a, b, c \in \mathbb{Z}$ e as seguintes coleções de subconjuntos:

$\mathcal{F}_1 = \{\emptyset, \Omega, \{a\}, \{b, c\}\}$;

$\mathcal{F}_2 = \{\emptyset, \Omega, \{a\}, \{b\}, \{a, c\}, \{b, c\}\}$;

$\mathcal{F}_3 = \{\emptyset, \Omega, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}\}$.

Quais dessas coleções são σ-álgebras?

Vamos verificar se as propriedades P_1, P_2 e P_3' estão satisfeitas.

P_1. $\Omega \in \mathcal{F}$: Observa-se que $\Omega \in \mathcal{F}_1$, $\mathcal{F}_2$ e $\mathcal{F}_3$, logo, P_1 está satisfeita para as três coleções.

P_2. Se $A \in \mathcal{F}$, então $A^C \in \mathcal{F}$:

Para $\mathcal{F}_1$ temos: $\emptyset^C = \Omega$, $\Omega^C = \emptyset$, $\{a\}^C = \{b, c\}$ e $\{b, c\}^C = \{a\}$. Logo $\mathcal{F}_1$ satisfaz a propriedade P_2. Analogamente, verificamos que o mesmo ocorre com as coleções $\mathcal{F}_2$ e $\mathcal{F}_3$.

P_3'. Se $A_n \in \mathcal{F}$ para $n = 1, 2, 3, ...$, então $\displaystyle\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$:

Para verificar P_3' basta verificar se todas as uniões possíveis com os elementos de cada coleção também pertence à coleção. A união com o vazio é irrelevante e a com o espaço amostral Ω dá o próprio Ω, vamos então verificar as demais uniões possíveis. Para $\mathcal{F}_1$ temos $\{a\} \bigcup \{b, c\} = \Omega \in \mathcal{F}_1$, e P_3' está satisfeita. Portanto, a coleção $\mathcal{F}_1$ é uma σ-álgebra. Tomando as uniões relevantes em $\mathcal{F}_2$ é fácil observar que $\{a\} \bigcup \{b\} = \{a, b\} \notin \mathcal{F}_2$, logo, P_3' não é satisfeita e a coleção de conjuntos $\mathcal{F}_2$ não forma uma σ-álgebra. Verificando as uniões possíveis em $\mathcal{F}_3$ verificamos que P_3' está satisfeita. Dessa forma, a coleção de conjuntos $\mathcal{F}_3$ também é uma σ-álgebra.

Observe que a coleção de conjuntos $\mathcal{F}_2$ também não é uma álgebra, pois não satisfaz a propriedade P_3 da Definição 1.5.1.

Observação 1.5.2

(i) A menor álgebra de eventos é $\mathcal{F} = \{\emptyset, \Omega\}$.

(ii) A maior álgebra de eventos é o conjunto das partes de Ω.

(iii) Se o espaço amostral for finito, toda álgebra é uma σ-álgebra, pois só existem um número finito de eventos diferentes. Se o espaço amostral for infinito, existem álgebras que não são σ-álgebras, como mostra o exemplo seguinte:

Exemplo 1.5.2.2 Será que toda álgebra é também σ-álgebra? O exemplo a seguir mostra que não: Selecionar, ao acaso, um ponto do intervalo [0, 1]. Aqui, $\Omega = [0,1]$ e $\mathcal{F} =$ todos os subconjuntos cujo comprimento esteja bem definido (por comprimento, dizemos uma função tal como $f([a,b]) = b - a$).

Inicialmente, vamos considerar uniões finitas de intervalos e seja $\mathcal{F}_0 = \{A \subset [0,1] : A$ é união finita de intervalos$\}$. $\mathcal{F}_0$ é álgebra? Devemos verificar se são satisfeitas as propriedades P_1, P_2 e P_3:

$\Omega \in \mathcal{F}_0$, logo P_1 é satisfeita. Se $A \in \mathcal{F}_0$ então A^C também é união finita de intervalos e pertence a $\mathcal{F}_0$, logo P_2 é satisfeita. Se A e B são uniões finitas de intervalos e A, $B \in \mathcal{F}_0$ então $A \bigcup B \in \mathcal{F}_0$, satistazendo P_3. O conjunto vazio $\emptyset$ e o evento elementar $\{\omega\}$, onde $\omega \in [0,1]$, serão interpretados como intervalos degenerados de comprimento 0, portando são elementos de $\mathcal{F}_0$. Podemos concluir que $\mathcal{F}_0$ é uma álgebra de eventos.

Mas $\mathcal{F}_0$ não é σ-álgebra, pois não contém toda união enumerável de intervalos, como teria que conter se fosse σ-álgebra. Por exemplo, o evento

$$A = \left(0, \frac{1}{2}\right) \bigcup \left(\frac{1}{2}, \frac{3}{4}\right) \bigcup \left(\frac{3}{4}, \frac{7}{8}\right) \bigcup ... \bigcup \left(1 - \frac{1}{2^n}, 1 - \frac{1}{2^{n+1}}\right) \bigcup ...$$

é união enumerável de intervalos mas não é união finita.

Em James [9] temos a seguinte definição de variável aleatória:

Definição 1.5.2.2 *Uma função* $X : \Omega \to \mathbb{R}$ *é variável aleatória se* $[X \leq x] \in \mathcal{F}\ \forall x \in \mathbb{R}$.

O conjunto $[X \leq x]$ é um subconjunto de Ω, $[X \leq x] = \{\omega \in \Omega : X(\omega) \leq x\}$, isto é, são os eventos que a imagem pela variável aleatória (imagem é um número real) é menor que x.

Exemplo 1.5.2.3 Os exemplos dados anteriormente (Exemplo 1.4.1) satisfazem a Definição 1.5.2.2, como podemos observar:
i) $\Omega = \{$cara, coroa$\}$, $\mathcal{F} = \{\emptyset, \Omega, \{$cara$\}, \{$coroa$\}\}$ é σ-álgebra. Vamos definir a variável aleatória V_1:

$$
\begin{array}{rccc}
V_1 : & \Omega & \to & \mathbb{R} \\
 & \text{cara} & \mapsto & 0 \\
 & \text{coroa} & \mapsto & 1
\end{array}
$$

Dessa forma, temos:

$$
\{\omega \in \Omega : V_1(\omega) \in (-\infty, x]\} = \begin{cases} \emptyset, & \text{se } x < 0; \\ \{\text{cara}\}, & \text{se } 0 \leq x < 1; \\ \Omega, & \text{se } x \geq 1. \end{cases}
$$

A leitura é: Quem está no espaço amostral tal que a imagem pela variável é menor que 0? Ninguém! Quem tem imagem menor que 1/2? Apenas a cara tem imagem menor que 1/2, pois sua imagem é 0. Quem tem imagem menor ou igual a 2?

Veja que a Definição 1.5.2.2 de variável aleatória é distinta da Definição 1.4, mas não nos preocuparemos com ela, pois é muito difícil encontrar um exemplo que não satisfaça a Definição 1.4.

Capítulo 2

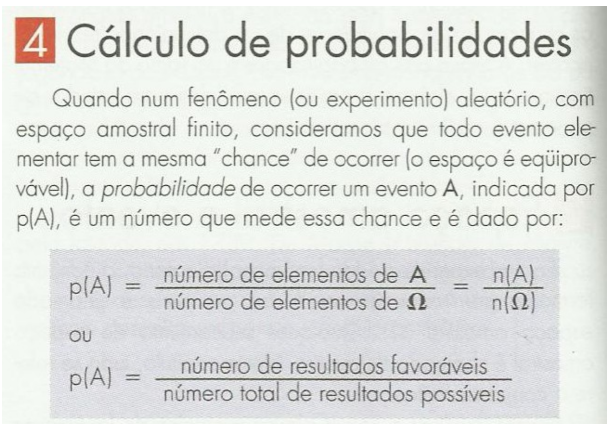
Probabilidade

2.1 Noções Básicas

Normalmente, a probabilidade aparece em livros didáticos da seguinte forma:

$$P(A) = \frac{\text{Número de elementos em } A}{\text{Número total de elementos em } \Omega}.$$

Como podemos verificar em Santos [13]:



4 Cálculo de probabilidades

Quando num fenômeno (ou experimento) aleatório, com espaço amostral finito, consideramos que todo evento elementar tem a mesma "chance" de ocorrer (o espaço é eqüiprovável), a *probabilidade* de ocorrer um evento A, indicada por p(A), é um número que mede essa chance e é dado por:

$$p(A) = \frac{\text{número de elementos de } \mathbf{A}}{\text{número de elementos de } \mathbf{\Omega}} = \frac{n(A)}{n(\Omega)}$$

ou

$$p(A) = \frac{\text{número de resultados favoráveis}}{\text{número total de resultados possíveis}}$$

Em Giovanni[8]:

Considere um experimento aleatório em que para cada um dos n eventos simples, do espaço amostral U, a chance de ocorrência é a mesma. Nesse caso, dizemos que o espaço amostral é um *espaço equiprovável* e que a probabilidade de cada evento simples é $\frac{1}{n}$.

Para um evento simples A, indicamos:

$$P(A) = \frac{1}{n(U)}$$

Podemos ampliar essa definição de probabilidade de um evento simples para a probabilidade de um evento qualquer.

$$P(A) = \frac{n(A)}{n(U)}$$

Na expressão, $n(U)$ é o número de elementos do espaço amostral U e $n(A)$, o número de elementos do evento A.

Em Bianchini [2]:

Probabilidade de ocorrer um evento

Em um experimento aleatório, onde S é um espaço equiprovável, a probabilidade de ocorrer um evento qualquer E é o número $p(E)$, dado por:

$$p(E) = \frac{n(E)}{n(S)}$$

Em Dante [6]:

Parte II

MÓDULO 109

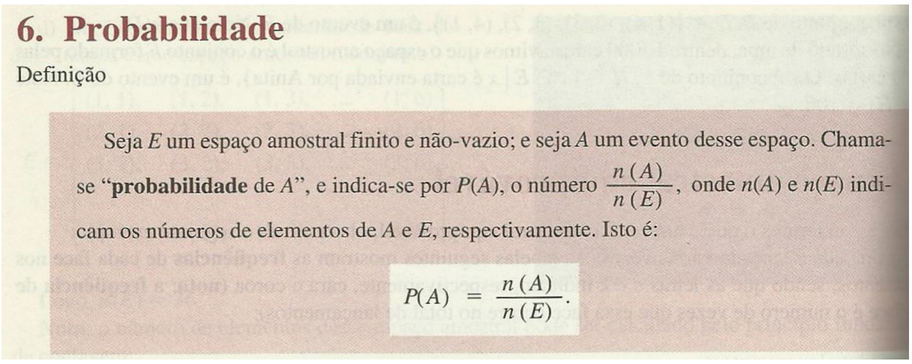
Probabilidade: Definição

Num experimento aleatório eqüiprovável, sendo $n(U)$ o número de elementos do espaço amostral **U** e $n(A)$ o número de elementos do evento **A**, a probabilidade de que ocorra o evento **A** é dada pelo número real:

$$P(A) = \frac{n(A)}{n(U)}$$

Então, a probabilidade de um evento é dada pelo quociente da divisão do número de casos favoráveis pelo número de casos possíveis. Exemplos:

Ou em Paiva [12]:



6. Probabilidade

Definição

Seja E um espaço amostral finito e não-vazio; e seja A um evento desse espaço. Chama-se "**probabilidade** de A", e indica-se por $P(A)$, o número $\frac{n(A)}{n(E)}$, onde $n(A)$ e $n(E)$ indicam os números de elementos de A e E, respectivamente. Isto é:

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

Em geral, a probabilidade é vista como uma função que estabelece uma medida para eventos aleatórios equiprováveis, ou seja, que têm a mesma chance de acontecer.

A função $P(.)$ atribui valores numéricos aos eventos do espaço amostral. Nesse caso, Ω deve ter um número de elementos finito e equiprováveis.

Exemplo 2.1.1 A probabilidade de sair face cara quando se lança uma moeda honesta ao ar será 1/2 já que o espaço amostral é $\Omega = \{\text{cara, coroa}\}$ e os resultados elementares são equiprováveis. Assim, temos que a cardinalidade de $A = \{\text{cara}\}$ é 1 e a de Ω é 2.

Através dessa definição muitos problemas são resolvidos com o uso de técnicas de análise combinatória e de contagem.

Para casos em que Ω não for enumerável e os eventos equiprováveis, podemos aplicar o conceito ao comprimento de intervalos, medida de áreas ou similares, utilizando a *probabilidade geométrica*:

$$P(A) = \frac{\text{Comprimento de } A}{\text{Comprimento total de } \Omega}.$$

Já a *probabilidade frequentista* ou *estatística* considera o limite de frequências relativas como o valor da probabilidade. Se n_A é o número de ocorrências de A em n repetições independentes do experimento, então,

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

Exemplo 2.1.2 A seguir as probabilidades dos eventos do Exemplo 1.3.

Utilizando a definição clássica:

i) $P(A_1) = \frac{1}{2}$ (ocorrer cara em um lançamento de uma moeda);

ii) $P(A_2) = \frac{4}{10} = \frac{2}{5}$ (retirar uma bola com um número primo, de uma urna com 10 bolas numeradas de 1 a 10);

iii) $P(A_3) = \frac{4}{6} = \frac{2}{3}$ (ocorrer um número menor que cinco em um lançamento de um dado);

iv) $P(A_4) = \frac{1}{n}$ (apenas uma peça é defeituosa em n peças);

v) $P(A_5) = \frac{1}{2}$ (escolher uma peça em uma linha de produção, verificar sua funcionalidade, e ela não ser defeituosa);

vi) $P(A_6) = \frac{1}{2}$ (escolher um par de meias de uma gaveta e elas terem a mesma cor).

Utilizando a definição frequentista: Para $P(A_7)$ considere inicialmente um espaço finito, digamos $\Omega_n = \{1, 2, 3, 4, ..., n\}$. Temos:

$$n_A = \begin{cases} \frac{n}{2}, & \text{se } n \text{ é par;} \\ \frac{n+1}{2}, & \text{se } n \text{ é ímpar.} \end{cases}$$

Segue que

vii) $P(A_7) = \lim_{n \to \infty} \frac{n_A}{n} = \frac{1}{2}$.

Utilizando a definição geométrica:

viii) $P(A_8) = \frac{\text{Comprimento de } A}{\text{Comprimento total de } \Omega} = \frac{1}{2}$.

Essas definições fazem uso da intuição e são utilizadas para resolver diversos problemas. Entretanto, elas não são suficientes para uma definição mais abrangente e rigorosa da probabilidade, pois sabemos que há eventos não-equiprováveis.

Por volta de 1930, o russo A. N. Kolmogorov apresentou um conjunto de axiomas matemáticos para definir probabilidade, proporcionando à Teoria da Probabilidade uma base matemática firme, e que incluiu as definições anteriores como casos particulares.

Definição 2.1.1 *Probabilidade é uma função P, definida na σ-álgebra $\mathcal{F}$ de subconjuntos de Ω e com valores entre $[0,1]$, que satisfaz os Axiomas de Kolmogorov:*

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

(ii) Para todo subconjunto $A \in \mathcal{F}, P(A) \geq 0$;

(iii) Para toda sequência A_1, A_2, ... em $\mathcal{F}$, mutuamente exclusivos, temos $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$.

Exemplo 2.1.3 Voltemos ao exemplo E_1, lançar uma moeda honesta e observar a face voltada para cima. Temos $\Omega = \{\text{cara, coroa}\}$ e a σ-álgebra $\mathcal{F}$ pode ser o conjunto das partes de Ω, ou seja, $\mathcal{F} = \{\emptyset, \{\text{cara}\}, \{\text{coroa}\}, \Omega\}$. P é a probabilidade uniforme em todos os pontos de Ω, isto é, $P(\{\omega\}) = 1/2$.

Observe que essa função satisfaz a definição anterior: $P(\Omega)$ é a ocorrência de cara ou coroa, ou seja, $P(\Omega) = 1$, logo o primeiro axioma está satisfeito.

$P(\emptyset) = 0$, $P(\{\text{cara}\}) = 1/2$, $P(\{\text{coroa}\}) = 1/2$ e $P(\Omega) = 1$, satisfazendo o segundo axioma.

Pode-se notar facilmente que o terceiro axioma também é satisfeito.

Definição 2.1.2 *Um espaço de probabilidade é uma trinca $(\Omega, \mathcal{F}, P)$ formada por um conjunto Ω não vazio, uma σ-álgebra $\mathcal{F}$ de subconjuntos em Ω e uma medida positiva P, definida em $\mathcal{F}$, que satisfaz a Definição 2.1.1.*

Propriedade 2.1.1 Se A^C é o complementar do evento A, então $P(A^C) = 1 - P(A)$ (e $P(A) = 1 - P(A^C)$).

Demonstração: A e A^C são disjuntos e $A \bigcup A^C = \Omega$, então $P(\Omega) = P(A \bigcup A^C) = P(A) + P(A^C) = 1$. Logo, $P(A^C) = 1 - P(A)$ (e $P(A) = 1 - P(A^C)$).

Caso particular importante: $P(\emptyset) = 0$, pois $P(\emptyset) = 1 - P(\emptyset^C) = 1 - P(\Omega) = 1 - 1 = 0$.

Propriedade 2.1.2 Se $A \subset B$, então $P(A) \leq P(B)$.

Demonstração: Podemos escrever o evento B da seguinte forma:

$$B = (B \bigcap A) \bigcup (B \bigcap A^C) = A \bigcup (B \bigcap A^C) \text{ (já que } A \subset B).$$

Essa união é disjunta, então $P(B) = P(A) + P(B \bigcap A^C) \geq P(A)$, pois $P(B \bigcap A^C) \geq 0$.

Propriedade 2.1.3 Para quaisquer eventos A e B de Ω, tem-se que

$$P(A \bigcup B) = P(A) + P(B) - P(A \bigcap B).$$

Demonstração: Podemos escrever $(A \bigcup B)$ como uma reunião de dois eventos mutuamente exclusivos: $A \bigcup B = A \bigcup (B \bigcap A^C)$. Segue que

$$P(A \bigcup B) = P(A) + P(B \bigcap A^C). \tag{2.1}$$

Notando que $B = (B \bigcap A) \bigcup (B \bigcap A^C)$, vem

$$P(B) = P(B \bigcap A) + P(B \bigcap A^C). \tag{2.2}$$

Tirando-se o valor de $P(B \bigcap A^C)$ de (2.2) e substituindo em (2.1), obtemos

$$P(A \bigcup B) = P(A) + P(B) - P(A \bigcap B).$$

2.2 Probabilidade Condicional

Em muitas situações práticas, é possível separar o fenômeno aleatório em etapas. O que ocorreu em uma etapa pode influenciar nas probabilidades de ocorrências das etapas seguintes.

Exemplo 2.2.1 Vamos considerar o exemplo de um lote de peças com a seguinte composição: 80 peças não defeituosas e 20 peças defeituosas. Suponha que escolhemos duas peças desse lote: (a) com reposição; (b) sem reposição.

Definamos os eventos:

$A = \{$a primeira peça é defeituosa$\}$; $B = \{$a segunda peça é defeituosa$\}$.

Quando extraímos com reposição temos $P(A) = P(B) = \frac{20}{100} = \frac{1}{5}$, porque em cada vez teremos 20 peças defeituosas em um total de 100. Agora, quando extraímos sem reposição, os resultados não são tão óbvios. O valor de $P(A)$ continua o mesmo, mas e $P(B)$? Para encontrar o valor de $P(B)$ devemos conhecer a composição do lote no momento de extrair a segunda peça, isto é, devemos saber se A ocorreu ou não. Este exemplo mostra a necessidade de se introduzir o conceito de probabilidade condicional.

Definição 2.2 *Dados dois eventos A e B, a probabilidade condicional de B dado que ocorreu A é representada por*

$$P(B|A) = \frac{P(B\bigcap A)}{P(A)}, \text{para } P(A) > 0. \tag{2.3}$$

No exemplo acima temos $P(B|A) = \frac{19}{99}$, pois, se A tiver ocorrido, então para a segunda retirada restarão 99 peças, das quais 19 são defeituosas.

Quando calculamos $P(B|A)$ estamos calculando $P(B)$ em relação ao espaço amostral reduzido A, em lugar de fazê-lo em relação ao espaço amostral original Ω. Vale ressaltar que na fórmula (2.3) os valores de $P(A)$ e $P(A\bigcap B)$ estão relacionados a Ω, pois quando falamos no espaço amostral reduzido A pode-se, por engano, pensar que $P(A) = 1$.

A partir de (2.3) podemos obter as seguintes expressões:

$$P(A\bigcap B) = P(A)P(B|A) \text{ ou } P(A\bigcap B) = P(B)P(A|B). \qquad (2.4)$$

Exemplo 2.2.2 Considere uma urna contendo três bolas azuis e sete bolas vermelhas. Retiram-se duas bolas da urna, sem reposição. Determinar o espaço amostral e as probabilidades associadas a cada ponto amostral.

O espaço amostral é o conjunto $\Omega = \{A_1A_2, A_1V_2, V_1A_2, V_1V_2\}$, que consiste nas combinações possíveis de retiradas das bolas azuis e vermelhas. O evento A_1A_2 corresponde a ocorrer bola azul na primeira retirada e azul na segunda retirada.

Para a primeira retirada temos três bolas azuis em um total de dez bolas, dessa forma $P(A_1) = \frac{3}{10}$. Para a segunda retirada, considerando que já foi retirada uma bola azul, temos duas bolas azuis em um total de nove bolas, assim, $P(A_2|A_1) = \frac{2}{9}$. Utilizando a fórmula vista em (2.4) temos:

$$P(A_1\bigcap A_2) = P(A_1)P(A_2|A_1) = \frac{3}{10}.\frac{2}{9} = \frac{1}{15}.$$

Para os outros pontos do espaço amostral a interpretação é a mesma:

$$P(A_1\bigcap V_2) = P(A_1)P(V_2|A_1) = \frac{3}{10}.\frac{7}{9} = \frac{7}{30}.$$

$$P(V_1\bigcap A_2) = P(V_1)P(A_2|V_1) = \frac{7}{10}.\frac{3}{9} = \frac{7}{30}.$$

$$P(V_1\bigcap V_2) = P(V_1)P(V_2|V_1) = \frac{7}{10}.\frac{6}{9} = \frac{7}{15}.$$

2.3 Independência de Eventos

Definição 2.3 *Dois eventos A e B são independentes se a informação da ocorrência ou não de B não altera a probabilidade da ocorrência de A.* Isto é,

$$P(A|B) = P(A), \text{para } P(B) > 0,$$

ou ainda

$$P(A \bigcap B) = P(A)P(B).$$

Se o evento A é independente do evento B, então B também é independente de A, como podemos verificar a seguir:

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

Exemplo 2.3 Considere a mesma urna do Exemplo 2.2.2, contendo três bolas azuis e sete bolas vermelhas. Retiram-se duas bolas da urna, agora com reposição. Vamos supor que as dez bolas são numeradas de 1 a 10, as azuis recebem os números de 1 a 3 e as vermelhas de 4 a 10. Podemos considerar o seguinte espaço amostral associado a esse experimento:

$$\Omega = \{(i, j) : i, j \in \mathbb{Z}, 1 \leq i \leq 10, 1 \leq j \leq 10\}.$$

Partindo da hipótese de que a probabilidade de qualquer das 10 bolas ser retirada é a mesma, já que há reposição, podemos associar aos pontos do espaço amostral probabilidade $\frac{1}{100}$ (lembrando que o espaço amostral é o conjunto $\Omega = \{(1,1),(1,2),...,(1,10),(2,1),(2,2),...,(2,10),...,(10,1),(10,2),...,(10,10)\}$ pois estamos considerando as bolas numeradas de 1 a 10, totalizando 100 elementos no espaço amostral).

Para a probabilidade do evento A_1A_2 temos três maneiras de retirar bola azul na primeira retirada e na segunda retirada também há três maneiras, pois a primeira bola é reposta. Assim o evento A_1A_2 pode ocorrer de nove maneiras e o espaço amostral tem 100 pontos. Portanto, $P(A_1 \bigcap A_2) = \frac{9}{100}$. Observe que $P(A_1) = \frac{3}{10}$ e $P(A_2) = \frac{3}{10}$, pois em cada retirada temos três bolas azuis em um total de dez bolas, logo:

$$P(A_1 \bigcap A_2) = P(A_1)P(A_2)$$

Para os demais eventos temos os seguintes valores:

$$P(A_1 \bigcap V_2) = \frac{21}{100} = \frac{3}{10}.\frac{7}{10} = P(A_1)P(V_2)$$

$$P(V_1 \bigcap A_2) = \frac{21}{100} = \frac{7}{10}.\frac{3}{10} = P(V_1)P(A_2)$$

$$P(V_1 \bigcap V_2) = \frac{49}{100} = \frac{7}{10}.\frac{7}{10} = P(V_1)P(V_2)$$

Esse exemplo nos mostra que os eventos A_1 e A_2, A_1 e V_2, V_1 e A_2 e V_1 e V_2 são independentes.

2.4 Variáveis Aleatórias em Espaço de Probabilidade

Uma definição mais formal de variável aleatória, que a define em um espaço de probabilidade, é a que segue:

Definição 2.4 *Seja $(\Omega, \mathcal{F}, P)$ um espaço de probabilidade. Denomina-se variável aleatória qualquer função $X : \Omega \to \mathbb{R}$ tal que*

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

para todo intervalo $I \in \mathcal{B}(\mathbb{R})$, onde $\mathcal{B}(\mathbb{R})$ é a σ-álgebra de Borel em $\mathbb{R}$. Em outras palavras, X é tal que sua imagem inversa de intervalos $I \subset \mathbb{R}$ pertencem à σ-álgebra $\mathcal{F}$.

Essa definição é equivalente à Definição 1.5.2.2. Podemos ver essa demonstração em Magalhães [11].

Portanto, uma variável aleatória é uma função do espaço amostral Ω nos reais, para a qual é possível calcular a probabilidade de ocorrência de seus valores.

As variáveis aleatórias que assumem valores em um conjunto enumerável de valores (finito ou infinito) serão denominadas *discretas* e aquelas que assumem valores em um intervalo da reta real serão denominadas *contínuas*.

2.5 Função de Distribuição

Quando admitimos valores para uma variável aleatória, podemos, frequentemente, atribuir probabilidade a cada um desses valores. Se conhecemos todos os valores de uma variável aleatória, juntamente com suas respectivas probabilidades, temos uma distribuição de probabilidades.

A função de distribuição associa uma probabilidade a cada resultado numérico de um experimento, ou seja, dá a probabilidade de cada valor de uma variável aleatória. James [9] e Magalhães [11] citam que alguns autores se referem a ela como função de distribuição acumulada, pois acumula as probabilidades dos valores inferiores ou iguais a algum x ($x \in \mathbb{R}$).

Definição 2.5 *A função de distribuição de uma variável aleatória X é uma função que a cada número real x associa o valor*

$$F(x) = P([X \leq x]), x \in \mathbb{R}.$$

A notação $[X \leq x]$ é usada para designar o conjunto $\{\omega \in \Omega : X(\omega) \leq x\}$, isto é, denota a imagem inversa do intervalo $(-\infty, x]$ pela variável aleatória X.

O conhecimento da função de distribuição permite obter qualquer informação sobre a variável. Mesmo que a variável só assuma valores num subconjunto dos reais, a função de distribuição é definida em toda a reta.

Propriedade 2.5 Valem as seguintes propriedades para a função de distribuição:

F_1. $\lim\limits_{x \to -\infty} F(x) = 0$ e $\lim\limits_{x \to \infty} F(x) = 1$;

F_2. F é contínua à direita em todo ponto de seu domínio;

F_3. F é não descrescente, isto é, $F(x) \leq F(y)$ sempre que $x \leq y$, $\forall x, y \in \mathbb{R}$.

Demonstração: Para verificar F_1 consideramos uma sequência de eventos A_n que cresce para um evento A. Podemos observar que $P(A_n)$ converge para $P(A)$. Se considerarmos uma sequência de números reais x_n que tende a infinito, segue-se que a sequência de eventos $[X \leq x_n]$ tende ao evento $[X < \infty]$ e portanto $P([X \leq x_n])$ converge para $P([X < \infty])$. Mas esse evento é o espaço amostral todo, que tem probabilidade um. Assim, fica demonstrado que $\lim_{x \to \infty} F(x) = 1$. Agora, consideramos uma sequência x_n que tende a $-\infty$. A sequência de eventos $[X \leq x_n]$ tende ao $\emptyset$ e portanto $F(x_n) = P[X \leq x_n]$ tende a zero.

Agora, para verificar F_2, consideramos uma sequência de números reais x_n, que seja decrescente e tenha como limite x, isto é, os x_n's se aproximam de x pela direita ou, em outras palavras, por valores superiores a x. A sequência de eventos $[X \leq x_n]$ é decrescente e se aproxima de $[X \leq x]$ e, assim, P$(X \leq x_n)$ se aproxima de $P(X \leq x)$, valendo o resultado para qualquer x, portanto a propriedade está verificada.

Para F_3 note que $[X \leq x] \subset [X \leq y]$ sempre que $x \leq y$. Logo as probabilidades satisfazem à desigualdade:

$$F(x) = P(X \leq x) \leq P(X \leq y) = F(y).$$

Como x e y são arbitrários, concluímos que F é não decrescente.

Exemplo 2.5 Considere dois lançamentos independentes de uma moeda equilibrada. Seja C cara e K coroa. O espaço amostral deste experimento é $\Omega = \{(C,C), (C,K), (K,C), (K,K)\}$. Podemos definir a variável aleatória X : "número de caras obtidas nos dois lançamentos". Por exemplo, temos que $X((C,C)) = 2$ e $X((K,C)) = 1$.

Vamos encontrar a função de distribuição de X. Os valores que X pode assumir são 0, 1 e 2. Portanto,

$$P(X = 0) = P((K,K)) = \frac{1}{2} . \frac{1}{2} = \frac{1}{4}.$$

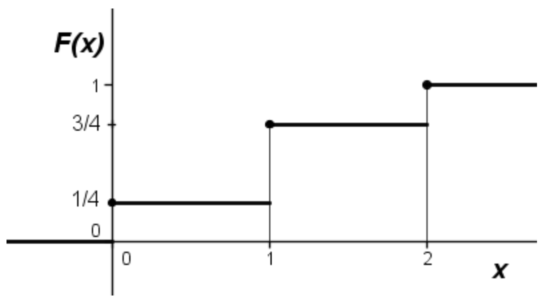
$$P(X = 1) = P((C,K)) + P((K,C)) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}.$$

$$P(X = 2) = P((C,C)) = \frac{1}{4}.$$

Assim, temos que a função de distribuição de X é dada por

$$F(x) = \begin{cases} 0, & \text{se } x < 0; \\ 1/4, & \text{se } 0 \leq x < 1; \\ 3/4, & \text{se } 1 \leq x < 2; \\ 1, & \text{se } x \geq 2. \end{cases}$$

Graficamente:



Podemos verificar facilmente as propriedades de função de distribuição. F_1 é imediata. Para F_2 observe que, exceto nos pontos 0, 1 e 2, F é contínua nos reais. O valor de $F(x)$ é zero para todo $x < 0$. Para $0 \leq x < 1$, $F(x) = \frac{1}{4}$, portanto, o limite à direita no ponto 0, isto é, por valores de x maiores que zero, é igual a $\frac{1}{4}$, que coincide com o valor de $F(0)$, ou seja, $F(x)$ é contínua à direita no ponto zero. Para os outros pontos de descontinuidade, vale uma expressão análoga, como se pode verificar. Observe que $F(x)$ é não decrescente para todo x real e, assim, vale F_3.

2.5.1 Função de Probabilidade

Seja X uma variável aleatória discreta, isto é, o número de valores possíveis de X é finito ou infinito enumerável. O contradomínio de X será formado no máximo por um número infinito enumerável de valores $x_1, x_2, ...$

Definição 2.5.1 *A função de probabilidade de uma variável discreta é uma função que atribui probabilidade a cada um dos possíveis valores assumidos pela variável, ou seja, a cada possível resultado x_i associaremos um número $p(x_i) = P(X = x_i)$, denominado probabilidade de x_i.*

Propriedade 2.5.1 Os números p(x_i), $i = 1, 2, ...$ devem satisfazer às seguintes condições:

(a) $0 \leq p(x_i) \leq 1, \forall i = 1, 2, ...;$

(b) $\sum_i p(x_i) = 1.$

Observe que os valores da função probabilidade estão sempre entre 0 e 1, logo (a) é verdadeira. Cada valor da variável induz a um subconjunto disjunto, cuja união é Ω, ou seja, $[X = x_i] \bigcap [X = x_j] = \emptyset$, se $i \neq j$, e $\bigcup_{i=1}^{\infty} [X = x_i] = \Omega$, o que valida (b).

Exemplo 2.5.1 Uma moeda viciada tem probabilidade de cara igual a 0,4. Para dois lançamentos independentes dessa moeda, vamos estudar o comportamento da variável *número de caras* e fazer o gráfico de sua função de probabilidade.

Os valores que X pode assumir são 0, 1 e 2. Como a moeda é viciada então

$P(C) = 0,4 = \dfrac{4}{10}$ e $P(K) = 0,6 = \dfrac{6}{10}$. Portanto,

$$P(X = 0) = P((K, K)) = \frac{6}{10}.\frac{6}{10} = \frac{36}{100} = 0,36.$$

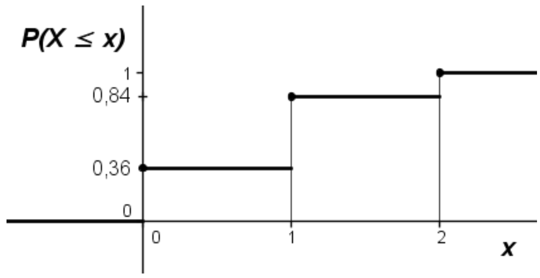
$$P(X = 1) = P((C, K)) + P((K, C)) = \frac{4}{10}.\frac{6}{10} + \frac{6}{10}.\frac{4}{10} = \frac{48}{100} = 0,48.$$

$$P(X = 2) = P((C, C)) = \frac{4}{10}.\frac{4}{10} = \frac{16}{100} = 0,16.$$

Assim, temos que a função de distribuição de X é dada por:

$$P(X \leq x) = \begin{cases} 0; & \text{se } x < 0; \\ 0,36; & \text{se } 0 \leq x < 1; \\ 0,84; & \text{se } 1 \leq x < 2; \\ 1; & \text{se } x \geq 2. \end{cases}$$

Graficamente:



2.5.2 Função Densidade

Seja X uma variável aleatória contínua, isto é, pode tomar qualquer valor numérico em um determinado intervalo ou coleção de intervalos.

Definição 2.5.2 *A função que representa a distribuição de probabilidade de uma variável aleatória contínua X é uma função f não negativa chamada de função densidade. Utilizamos para esse efeito a integral, sendo a função f:*

$$F(x) = \int_{-\infty}^{x} f(\omega)d\omega, \text{ para todo } x \in \mathbb{R}.$$

Propriedade 2.5.2 A função densidade de X em um espaço de probabilidade satisfaz as seguintes condições:

(a) $f(x) \geq 0, \forall x \in \mathbb{R}$;

(b) $\displaystyle\int_{-\infty}^{\infty} f(\omega)d\omega = 1$.

Por definição a função f é não negativa, logo (a) é verdadeira. Temos que (b) é verdadeira como consequência de que $\lim\limits_{x\to\infty} F(x) = 1$.

Exemplo 2.5.2 Uma variável aleatória contínua X possui a seguinte função:

$$f(x) = \begin{cases} k, & \text{para } 0 \leq x < 1; \\ k(2-x), & \text{para } 1 \leq x < 2; \\ 0, & \text{para outros valores de } x. \end{cases}$$

Vamos verificar para quais valores da constante k a função $f(x)$ é uma função densidade de probabilidade.

Para atender à segunda propriedade devemos ter $\int_{-\infty}^{\infty} f(x)dx = 1$, logo

$$\int_{-\infty}^{0} 0dx + \int_{0}^{1} kdx + \int_{1}^{2} k(2-x)dx + \int_{2}^{\infty} 0dx = 1,$$

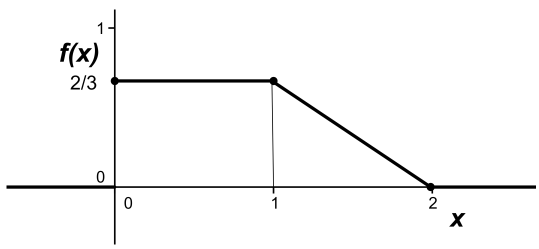
que resulta em

$$kx|_0^1 + k\frac{4x - x^2}{2}\bigg|_1^2 = 1 \rightarrow k + k\frac{1}{2} = 1 \rightarrow k = \frac{2}{3}$$

Para esse valor a segunda propriedade está satisfeita e temos a função

$$f(x) = \begin{cases} \frac{2}{3}, & \text{para } 0 \leq x < 1; \\ \frac{4-x}{3}, & \text{para } 1 \leq x < 2; \\ 0, & \text{para outros valores de } x. \end{cases}$$

Graficamente:



A primeira propriedade também está satisfeita, portanto a função f é uma função densidade de probabilidade.

Capítulo 3

Sugestão de Aula

3.1 Introdução

Inicialmente, podemos fazer um experimento com a turma. Utilizaremos uma moeda honesta e pediremos que cada aluno a lance uma vez (para a aula ficar participativa), anotanto os resultados dos lançamentos. Com os resultados em mãos, perguntamos: "Qual a chance de sair cara?".

Podemos pedir a um aluno que faça o quociente do número de lançamentos que deram cara, pelo total de lançamentos e dizer que o valor encontrado é uma estimativa da chance. Devemos ressaltar que, como o experimento é aleatório, pode ocorrer um número diferente de 0,5 (já que é de conhecimento popular que as chances em um lançamento de uma moeda honesta são de 50% para cada uma das faces).

Em seguida, podemos dar mais um exemplo, como o de um sorteio:

Carlos está participando do sorteio de uma moto. Podem ser enviadas quantas cartas quiser para esse sorteio, mas será sorteada apenas uma, dentre todas as cartas enviadas por todos os participantes.

Sabendo-se que Carlos enviou 100 cartas e que o total de cartas enviadas para o sorteio foi de 10.000, qual é a chance dele ser sorteado, supondo que todas as cartas têm a mesma chance de sair?

Carlos enviou 100 cartas, então dizemos que sua chance de ser sorteado é de 100 em 10.000 e indicamos pela razão $\dfrac{100}{10.000} = \dfrac{1}{100}$.

A razão $\frac{1}{100}$ é chamada de **probabilidade** de Carlos ser sorteado. Podemos dizer que probabilidade é um valor numérico que mede a possibilidade de um resultado ocorrer.

Dando continuidade, devemos introduzir conceitos importantes no estudo de probabilidades (experimento aleatório, espaço amostral, eventos equiprováveis, etc).

3.2 Experimento Aleatório

É qualquer experimento cujo resultado depende exclusivamente do acaso.

Exemplo 3.2
a) O lançamento de uma moeda.
b) O lançamento de um dado.

Podemos pedir que cada aluno escreva um experimento aleatório em seu caderno, diferente dos exemplos dados (isso pode ser feito em duplas ou trios).

3.3 Espaço Amostral

É o conjunto de todos os resultados possíveis de um experimento aleatório. Ele será representado pela letra grega Ω (ômega).

Vamos considerar os exemplos anteriores (Exemplo 3.2) e descrever o espaço amostral para cada um deles.

Exemplo 3.3
a) $\Omega_1 = \{\text{cara, coroa}\}$.
b) $\Omega_2 = \{1, 2, 3, 4, 5, 6\}$.

Neste momento, pedir que cada aluno (ou grupo de alunos) escreva o espaço amostral para os experimentos aleatórios escolhidos anteriormente.

3.4 Evento de um Espaço Amostral

É qualquer subconjunto do espaço amostral.

Exemplo 3.4
a) No lançamento de uma moeda, $A_1 = \{\text{cara}\}$ é um evento de Ω_1.
b) No lançamento de um dado, $A_2 = \{1, 2, 3, 4\}$ é um evento de Ω_2.

Mais uma vez, pedir que os alunos escrevam eventos para os espaços amostrais anteriores.

3.5 Espaço Amostral Equiprovável

Espaço equiprovável é aquele em que, sob mesmas condições, todos os elementos têm a mesma chance de acontecer. Veja que está relacionado com o experimento! Analisar em conjunto com a turma se os eventos e seus respectivos espaços são equiprováveis.

Podemos retomar o exemplo inicial, da moeda honesta. Supor que uma moeda é lançada várias vezes e os resultados foram os mostrados a seguir (se houver tempo suficiente, podemos realizar esse experimento em sala e não utilizar os valores dados aqui):

i) em dez lançamentos: 6 caras e 4 coroas.
ii) em vinte lançamentos: 8 caras e 12 coroas.
iii) em trinta lançamentos: 14 caras e 16 coroas.

Quando aumenta-se indefinidamente o número de lançamentos, as frequências das faces tenderão a um mesmo número. Por isso dizemos que o espaço amostral $\Omega = \{\text{cara, coroa}\}$ é equiprovável.

3.6 Probabilidade

Seja Ω um espaço amostral finito, não-vazio e equiprovável e seja A um evento desse espaço. Chamamos de *probabilidade* de A, e indicamos por $P(A)$, o número $\dfrac{n(A)}{n(\Omega)}$, onde $n(A)$ e $n(\Omega)$ indicam os números de elementos de A e Ω, respectivamente. Isto é:

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

Exemplo 3.6.1

a) $P(A_1) = \dfrac{n(A_1)}{n(\Omega_1)} = \dfrac{1}{2}$.

b) $P(A_2) = \dfrac{n(A_2)}{n(\Omega_2)} = \dfrac{4}{6} = \dfrac{2}{3}$.

Pedir que os alunos calculem as probabilidades dos eventos escolhidos, nos casos em que o espaço amostral for equiprovável.

Como dito anteriormente, um espaço amostral finito Ω é equiprovável quando seus eventos elementares têm probabilidades iguais de ocorrência, ou seja, se $\Omega = \{e_1, e_2, ..., e_n\}$, dizemos que os eventos elementares $\{e_i\}$ são equiprováveis se todos têm a mesma probabilidade de ocorrer:

$$P(\{e_i\}) = \frac{1}{n}.$$

Exemplo 3.6.2

a) No lançamento de uma moeda honesta, os elementos do espaço amostral $\Omega = \{$cara, coroa$\}$ são equiprováveis, pois cada elemento do espaço amostral tem a mesma chance de ocorrer, ou seja, a chance de sair cara é a mesma chance de sair coroa. Portanto,

$$P(\{\text{cara}\}) = P(\{\text{coroa}\}) = \frac{1}{2}.$$

b) No lançamento de um dado honesto, os elementos do espaço amostral $\Omega = \{1, 2, 3, 4, 5, 6\}$ também são equiprováveis, pois cada elemento do espaço amostral tem a mesma chance de ocorrer, ou seja, a chance de sair 1 é a mesma de sair 2, que é a mesma de sair 3, e assim por diante. Portanto,

$$P(\{1\}) = P(\{2\}) = P(\{3\}) = P(\{4\}) = P(\{5\}) = P(\{6\}) = \frac{1}{6}.$$

3.7 Função de Probabilidade

Podemos corresponder a cada evento possível de uma experiência aleatória a probabilidade do evento ocorrer. Em outras palavras, podemos definir da seguinte forma: seja $\Omega = \{x_1, x_2, ..., x_n\}$ os possíveis resultados do experimento. Escrevemos $P(\{x_i\})$ para representar a probabilidade do evento x_i acontecer. Assim, para simplificar escrevemos $P(\{x_i\}) = p(x_i)$ e tem as seguintes propriedades:

(a) $0 \leq p(x_i) \leq 1, \forall i = 1, 2, ...,$ onde $p(x_i)$ é a probabilidade associada a cada evento x_i;

(b) $\sum_i p(x_i) = 1$.

Exemplo 3.7.1 Considere o lançamento de uma moeda honesta. Seja C cara e K coroa. O espaço amostral deste experimento é $\Omega = \{C,\ K\}$. Temos que $P(\{C\}) = \frac{1}{2}$ e $P(\{K\}) = \frac{1}{2}$.

Vamos verificar se ao definirmos $P(\{C\}) = P(\{K\}) = \frac{1}{2}$ então $P(\{x_i\}) = p(x_i)$ satisfaz às duas propriedades da definição.

A propriedade (a) é satisfeita, pois $0 \leq p(x_i) \leq 1, i = 1, 2$. A propriedade (b) é satisfeita, pois $\sum_i p(x_i) = p(x_1) + p(x_2) = P(\{C\}) + P(\{K\}) = \frac{1}{2} + \frac{1}{2} = 1$. Portanto, a função definida é uma função de probabilidade.

Note que, a rigor, para a definição de função de probabilidade, seria necessário falar em variável aleatória, conforme Definição 2.5.1, mas, para simplificar o assunto nessa sugestão de aula, nos referimos aos x_i como possíveis resultados do experimento.

3.8 Eventos Não Equiprováveis

Vai chover amanhã? Se formos pensar nas alternativas de respostas para essa pergunta teríamos apenas duas opções: a de chover e a de não chover. Poderíamos, então, dizer que a probabilidade de chover amanhã é de 50%?

A resposta a essa pergunta é não, pois o universo de possíveis eventos futuros representado pelo espaço amostral {chover, não chover} não é equiprovável. Apesar de o conjunto possuir apenas dois elementos o acontecimento de cada um depende de uma série de fatores externos, fora do nosso alcance.

Podemos, para exemplicar em sala de aula, elaborar um dado viciado, e fazer a experiência com a turma. Para isso precisaríamos confeccionar dois dados, um honesto e um viciado, para efetuarmos uma mesma quantidade de lançamentos com os dois e verificar os resultados.

Para viciar o dado, prenderíamos, com fita adesiva, uma moeda de R$ 0,50 (cinquenta centavos) na parte interna de uma das faces, no momento de sua confecção. Na página seguinte segue o modelo para confecção dos dados.

Repetir o experimento do início da aula, utilizando o dado viciado e o dado honesto. Pedir para que cada aluno jogue os dois dados e anotar os resultados. Os alunos observarão que as chances de sair os valores mudam.

Calcular as chances do dado viciado por estimativa, isto é, lançar 20 vezes o dado e fazer o quociente do número de vezes em que cada número saiu, dividido por 20. Com esses dados definir a função de probabilidade.

Supondo que o dado viciado apresente 50% de chance de obter a face 1 voltada para cima em um lançamento, e as outras faces tenham chances iguais e somem 50%, a função de probabilidade poderia ser representada como:

$$p(x) = \begin{cases} 1/2, & \text{se } x = 1; \\ 1/10, & \text{se } x = 2, 3, 4, 5, 6. \end{cases}$$

Vamos verificar se esta função satisfaz às duas propriedades da definição.

A propriedade (a) é satisfeita, pois $0 \leq p(x_i) \leq 1, i = 1, 2, 3, 4, 5, 6$.

A propriedade (b) também é satisfeita, pois $\sum_i p(x_i) = p(x_1) + p(x_2) + ... + p(x_6) = P(\{1\}) + P(\{2\}) + ... + P(\{6\}) = \frac{1}{2} + \frac{1}{10} + \frac{1}{10} + \frac{1}{10} + \frac{1}{10} + \frac{1}{10} = 1$.

Portanto, a função definida é uma função de probabilidade.

Curiosidade: É possível encontrar na internet algumas formas de viciar um dado. Por exemplo, o *site* da revista Mundo Estranho afirma que basta aquecer o dado, com o número escolhido para cima, por um minuto no microondas. O calor derrete o dado sem deformá-lo e faz o material interno se acumular na parte de baixo, que fica mais densa, deixando a face viciada voltada para cima. O *site* afirma que o dado viciado cai mais vezes - mas nem sempre - apontando o mesmo número[1].

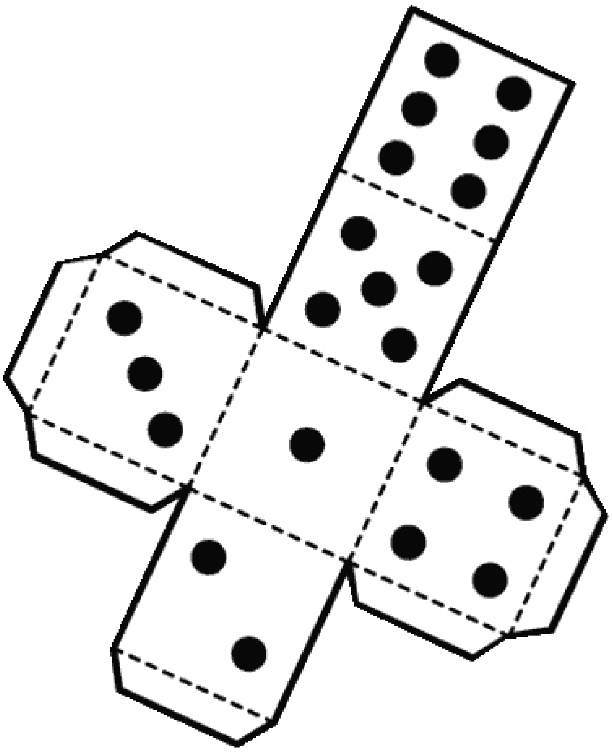


Figura 3.1: Modelo para confecção de dado.

[1]Disponível em *http://mundoestranho.abril.com.br/materia/e-possivel-viciar-um-dado*

Referências

[1] Berlinghoff, William P., Gouvêa, Fernando Q. *A matemática através dos tempos*: um guia fácil e prático para professores e entusiastas. Trad. Elza Gomide, Helena Castro. - São Paulo: Edgard Blücher, 2008.

[2] Bianchini, Edwaldo., Paccola, Herval. *Matemática*, volume 2: versão alfa. 2. ed. rev. e ampl. - São Paulo: Moderna, 1995.

[3] Boyer, Carl B. *História da matemática*. Trad. Elza Gomide. 2. ed., 4. reimpr. - São Paulo: Edgard Blücher, 2006.

[4] Coutinho, Cileda Q. S. *Introdução ao conceito de probabilidade por uma visão frequentista*: estudo epistemológico e didático. São Paulo: PCU, 1994. 150 p.

[5] Cramér, Harald. *Elementos da teoria da probabilidade*: e algumas de suas aplicações. Trad. Luiz A. Caruso. - São Paulo: Editora Mestre Jou, 1973.

[6] Dante, Luiz R. *Matemática*, volume único. 1. ed. - São Paulo: Ática, 2005.

[7] Garding, Lars. *Encontro com a matemática*. Trad. de Célio Alvarenga e Maria M. Alvarenga. - Brasília : Editora Universidade de Brasília, 2. ed., 1997.

[8] Giovanni, José R., Bonjorno, José R. *Matemática*: uma nova abordagem, vol. 2: versão progressões. - São Paulo: FTD, 2000.

[9] James, Barry R. *Probabilidade*: um curso em nível intermediário. 2. ed. - Rio de Janeiro: Instituto de Matemática Pura e Aplicada, 1996.

[10] Lima, César A. V. *Probabilidade para o ensino médio*. São Cristóvão: Universidade Federal de Sergipe, 2013. 45 p.

[11] Magalhães, Marcos N. *Probabilidade e variáveis aleatórias*. 2. ed. - São Paulo: Editora da Universidade de São Paulo, 2006.

[12] Paiva, Manoel R. *Matemática*, volume 2. 1. ed. - São Paulo: Moderna, 1995.

[13] Santos, Carlos A. M., Gentil, Nelson., Greco, Sérgio E. *Matemática*, volume único, Série Novo Ensino Médio. 6. ed., 1. reimpr. - São Paulo: Ática, 2000.