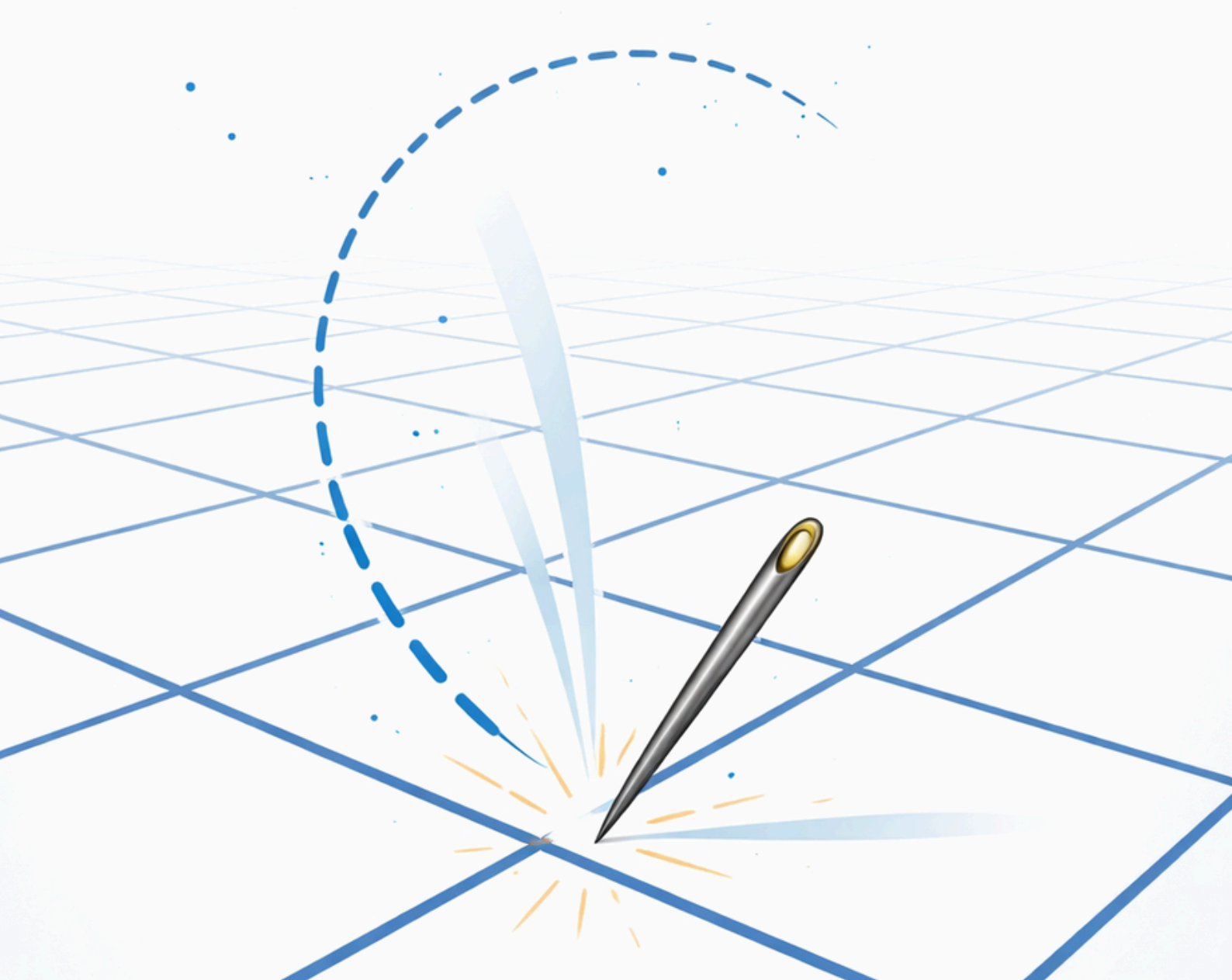


UNIVERSIDADE REGIONAL DO CARIRI – URCA
DEPARTAMENTO DE MATEMÁTICA
PROGRAMA DE MESTRADO PROFISSIONAL EM MATEMÁTICA EM REDE NACIONAL

Alefe de Sousa Carvalho

Probabilidade Geométrica sob a Perspectiva da Teoria da Medida: formalização, modelagem e transposição didática

Manual do professor e do estudante



Sumário

1	INTRODUÇÃO	4
2	PROBABILIDADE CLÁSSICA	6
2.1	História da Probabilidade	6
2.2	Conjuntos e Definições	9
2.3	Probabilidade	14
2.4	Probabilidade Condicional	21
2.5	O Método de Monte Carlo	30
3	PROBABILIDADE GEOMÉTRICA E TEORIA DA MEDIDA	31
3.1	Probabilidade Geométrica e suas Modelagens Intuitivas	31
3.2	Medida - Espaço mensurável	34
3.2.1	O Teorema de Vitali e a impossibilidade de medir todos os conjuntos	36
3.2.2	Medida Exterior	37
3.2.3	Crítério de mensurabilidade	38
3.2.4	Teorema da Extensão de Carathéodory	40
3.3	Retângulos em \mathbb{R}^n	41
3.4	Pré-medida em retângulos	42
3.4.1	Consistência da Pré-medida Geométrica	42
3.4.2	Construção da Medida de Lebesgue em \mathbb{R}^n	44
3.4.3	sigma-álgebra de Lebesgue e medida de Lebesgue	45
3.4.4	Compatibilidade com a sigma-álgebra de Borel	45
3.4.5	Invariância por translações da medida de Lebesgue	45
3.4.6	Homogeneidade (escalonamento) da medida de Lebesgue	47
3.4.7	Conjuntos de Medida Zero	49
3.5	Interpretação probabilística dos conjuntos de medida zero	49
3.5.0.1	Espaço de probabilidade contínuo	49
3.5.0.2	Normalização da medida em conjuntos de medida finita	49
3.5.0.3	Probabilidade induzida por medida (probabilidade geométrica)	49
3.6	Probabilidade Geométrica - Definição Formal	50
3.6.1	Espaço amostral geométrico	50
3.6.2	Definição formal de probabilidade geométrica em \mathbb{R}^n	50
3.6.3	Modelo Linear - Formal	51
3.6.4	Modelo Superficial - Formal	52
3.6.5	Modelo Volumétrico - Formal	52
3.6.6	Condições de validade da probabilidade geométrica	53

3.6.7	Como encarar a probabilidade geométrica daqui para frente	53
3.6.8	O que ocorre ao relaxarmos os axiomas da medida de Lebesgue?	54
3.7	Modelo de Coordenadas Geométricas	55
3.7.1	O Problema do Encontro	55
3.7.2	O Problema do Espaguete	56
3.8	Espaço Amostral Infinito Enumerável	58
3.9	Problemas Clássicos de Probabilidade Geométrica	60
3.9.1	A Agulha de Buffon e de Buffon-Laplace	60
3.9.2	Paradoxo de Bertrand	66
4	PROBLEMAS E SOLUÇÕES	74
	REFERÊNCIAS	95
	ANEXO A – DEMONSTRAÇÕES DOS TEOREMAS	97
A.1	O Teorema de Vitali	97
A.2	O Teorema de Carathéodory	98
A.3	O Teorema da Extensão de Carathéodory	101
A.4	Compatibilidade com a Sigma-Álgebra de Borel	106

1 Introdução

A probabilidade geométrica constitui um ramo da teoria das probabilidades que trata de experimentos aleatórios definidos em espaços contínuos, nos quais a atribuição de probabilidades está associada à razão entre medidas geométricas. Apesar de sua relevância teórica e de sua presença recorrente em problemas de olimpíadas, esse tema costuma ser apresentado de forma excessivamente intuitiva, sem o devido esclarecimento de seus fundamentos matemáticos.

Neste material, apresenta-se uma sistematização da probabilidade geométrica a partir da teoria da medida, com ênfase na construção da medida de Lebesgue e em suas propriedades essenciais, como invariância por translações e homogeneidade, além da análise de problemas que desafiam a intuição geométrica. Também é apresentada uma breve contextualização histórica do tema, bem como os tópicos básicos necessários à compreensão da abordagem mais abstrata.

Com base nesse arcabouço teórico, discutem-se diferentes modelagens da probabilidade geométrica em \mathbb{R}^n , culminando em uma definição formal compatível com os axiomas de Kolmogorov. Paralelamente, o material propõe uma transposição didática desse conteúdo, voltada a professores e estudantes do ensino médio com interesse acadêmico e olímpico, buscando conciliar rigor matemático e acessibilidade conceitual.

Para isso, são apresentados problemas clássicos da probabilidade geométrica, bem como situações que exigem modelagem cuidadosa do espaço amostral, incluindo o paradoxo de Bertrand, discutido à luz da teoria da medida. O material elaborado busca contribuir para a compreensão conceitual da probabilidade geométrica e para o fortalecimento da formação matemática de professores e estudantes.

Neste material, os capítulos seguintes são dedicados à fundamentação teórica, organizada em diferentes eixos. Inicialmente, apresenta-se um capítulo voltado aos pré-requisitos necessários à plena compreensão da probabilidade geométrica. Esse capítulo inicia-se com um panorama histórico da probabilidade, desde suas origens até desenvolvimentos contemporâneos. Em seguida, são discutidas as definições e os principais teoremas da teoria clássica da probabilidade, incluindo o conceito de probabilidade condicional, finalizando com uma breve descrição do método de Monte Carlo.

Na sequência, apresenta-se um capítulo dedicado à probabilidade geométrica e à teoria da medida, no qual se introduz, inicialmente, uma modelagem intuitiva da probabilidade geométrica. Posteriormente, são abordados tópicos de teoria da medida, com o objetivo de fornecer a base formal necessária à sua definição rigorosa, culminando na formalização desse conceito. O capítulo prossegue com a apresentação dos principais

modelos de probabilidade geométrica e se encerra com a discussão de casos de espaços amostrais enumeráveis infinitos, ainda que não geométricos.

Por fim, apresenta-se um capítulo dedicado à resolução de problemas clássicos e contemporâneos relacionados a essa teoria. Encerrada essa etapa, cuja composição envolve os capítulos de probabilidade clássica, probabilidade geométrica, teoria da medida e problemas e soluções, apresenta-se as referências utilizadas. As demonstrações de teoremas que demandam maior aprofundamento técnico são reunidas nos anexos, a fim de preservar a fluidez da leitura.

2 Probabilidade Clássica

2.1 História da Probabilidade

A teoria das probabilidades tem suas origens ligadas aos jogos de azar, emergindo como disciplina matemática no século XVII. Um marco inicial ocorreu com as correspondências trocadas entre Blaise Pascal e Pierre de Fermat, em que ambos investigaram problemas sobre apostas justas em jogos interrompidos. Nessas cartas, motivadas por questões propostas pelo jogador Antoine Gombaud (Cavaleiro de Méré), Pascal e Fermat formularam os princípios fundamentais da probabilidade matemática, resolvendo o chamado “problema dos pontos” (como dividir o prêmio de forma justa quando a partida é interrompida antes do término). Não se sabe o quanto Gombaud ganhou com os conselhos de Pascal, mas a matemática certamente ganharia muito com a nova área que seria desenvolvida (JOHNSON, 1995). Embora pensadores renascentistas como Girolamo Cardano já tivessem analisado o acaso, foi com Pascal e Fermat que se estabeleceu uma metodologia sistemática para o cálculo de probabilidades. Poucos anos depois, em 1657, Christiaan Huygens publicou *De Ratiociniis in Ludo Aleae* (Sobre O Raciocínio em Jogos de Azar), o primeiro tratado de probabilidade, no qual apresentou de forma estruturada as ideias de Pascal-Fermat. Posteriormente, Jan De Witt, em seu *Tratado sobre Anuidades Vitalícias* (1671), dá continuidade aos estudos de probabilidade, introduzindo o conceito de esperança matemática (diferente de esperança moral, ideia proposta por Daniel Bernoulli). Esses desenvolvimentos iniciais lançaram as bases da Teoria das Probabilidades, definindo regras para situações discretas (como lançamentos de dados e cartas) e demonstrando que o acaso podia ser descrito de maneira quantitativa e racional. (BOYER; MERZBACH, 2012)

Nos séculos seguintes, a teoria progrediu significativamente pelas mãos de diversos matemáticos. Dos diversos devotos à probabilidade, um deles foi Abraham De Moivre. Em 1711 publicou a obra *Philosophical Transactions* (Transações Filosóficas), um longo trabalho sobre as leis do acaso, e o expandiu no volume *Doctrine of Chances* (A Doutrina das Probabilidades), em 1718. Nesse trabalho foram apresentados diversos problemas sobre tirar bolas de cores diferentes de um saco, sobre dados e vários outros tipos de jogos. Jacques Bernoulli, em sua obra póstuma *Ars Conjectandi* (A Arte de Conjecturar) (1713), ampliou os fundamentos ao provar a Lei dos Grandes Números, resultado que estabeleceu que as frequências observadas tendem às probabilidades teóricas quando o número de ensaios é grande. Esse foi um passo crucial ao conectar a teoria abstrata com a experiência empírica, mostrando matematicamente por que a probabilidade funciona como uma medida de tendência em longo prazo. Ao longo do século XVIII e início do XIX,

outros estudiosos consolidaram e expandiram o campo. Pierre-Simon Laplace (1749–1827) publicou *Théorie Analytique des Probabilités* (Teoria Analítica das Probabilidades) (1812), uma obra monumental que sistematizou métodos de cálculo probabilístico (inclusive técnicas integrais, uso de funções geradoras e a transformada de Laplace) e aplicou a teoria a problemas reais em astronomia, física e estatística moral. Laplace via a probabilidade como “bom senso expresso em números”, sublinhando seu duplo papel teórico e prático (HAZZAN, 2013). Por volta de meados do século XIX, a teoria das probabilidades já se estabelecia com um corpo conceitual sólido, englobando tanto casos discretos quanto contínuos, e mostrava utilidade em diversos campos do conhecimento, da física astronômica à fiabilidade de seguros, preparando terreno para refinamentos teóricos subsequentes.

Paralelamente à consolidação da probabilidade clássica, surgiu o conceito de probabilidade geométrica, motivado por problemas que envolviam medições de comprimento, área ou volume em vez de contagens de casos finitos. A ideia de atribuir probabilidades por meio de razões geométricas ganhou destaque a partir do século XVIII, quando o naturalista e matemático francês Georges-Louis Leclerc (Conde de Buffon) propôs um problema célebre que se tornaria o marco inicial desse tópico. Em 1777, Buffon apresentou o problema da agulha: qual é a chance de uma agulha, quando lançada aleatoriamente sobre um piso com tábuas ou linhas paralelas igualmente espaçadas, cair de modo que atravesse uma dessas linhas? A solução requer integrar sobre todas as possíveis posições e orientações da agulha, levando em conta relações geométricas. Daí, culmina em um cálculo cujo resultado envolve o número π . A resolução da agulha de Buffon inaugurou formalmente a probabilidade geométrica, pois demonstrou como probabilidades podiam ser obtidas por proporções de comprimentos ou áreas (no caso, o comprimento da agulha em relação à distância entre as linhas) em um espaço de resultados contínuo. Esse trabalho pioneiro evidenciou a íntima ligação entre probabilidade e cálculo integral: problemas de acaso podiam ser resolvidos por métodos da geometria e da análise (BOYER; MERZBACH, 2012). Nos anos que se seguiram, outros problemas geométricos clássicos foram investigados, consolidando o novo ramo. Por exemplo, questões como “qual a probabilidade de, ao quebrar um segmento em três partes, formar-se um triângulo?” ou “qual a probabilidade de três pontos escolhidos ao acaso no plano formarem um triângulo acutângulo?” passaram a ser exploradas. Um destaque no final do século XIX foi o paradoxo de Bertrand (1889), que discute a probabilidade de uma corda aleatória num círculo ser maior que o lado do triângulo equilátero inscrito. Bertrand mostrou que, sem um critério rigoroso para definir o que é “escolher uma corda ao acaso”, respostas contraditórias (como $\frac{1}{2}$ ou $\frac{1}{3}$) podem surgir. Esse paradoxo evidenciou a necessidade de fundamentação mais rigorosa para a probabilidade geométrica: tornou-se claro que era imprescindível definir com precisão o método de seleção aleatória em um contínuo, sob pena de resultados inconsistentes. Nas próximas seções, será trabalhado de forma mais aprofundada esse paradoxo.

No alvorecer do século XX, a matemática forneceu as ferramentas conceituais

para resolver essas questões de rigor e formalizar de vez a teoria da probabilidade em espaços contínuos. Em 1902, Henri Lebesgue introduziu a teoria da medida e a integral de Lebesgue, aprofundando o tratamento de grandezas contínuas. Com essa nova abordagem, tornou-se possível definir rigorosamente o “tamanho” (medida) de conjuntos em \mathbb{R}^n , seja pelo comprimento de subconjuntos da reta, área de regiões do plano ou volume em espaços de dimensão superior, mesmo em situações altamente abstratas. A aplicação da teoria da medida à probabilidade foi imediata e poderosa: ela forneceu a base técnica para definir probabilidades geométricas de forma não ambígua, atribuindo a eventos contínuos medidas compatíveis com nosso intuitivo “cálculo de áreas”. Ambiguidades como as do paradoxo de Bertrand foram eliminadas ao se entender que cada procedimento aleatório corresponde a uma medida específica no espaço geométrico considerado. Consolidando essa visão, em 1933 Andrey Kolmogorov publicou *Grundbegriffe der Wahrscheinlichkeitsrechnung* (Fundamentos da Teoria das Probabilidades), no qual apresentou os axiomas modernos da probabilidade. Kolmogorov definiu probabilidade como uma medida definida em uma σ -álgebra de subconjuntos de um espaço amostral, satisfazendo propriedades de não-negatividade, aditividade (contável) e normalização (probabilidade total igual a 1). Essa estrutura axiomática, firmemente baseada na teoria de Lebesgue, unificou os tratamentos discreto e contínuo num só arcabouço teórico. Com isso, a probabilidade geométrica adquiriu status formal dentro da teoria geral, deixando de ser apenas uma coleção de problemas curiosos para se tornar parte integrante do conhecimento probabilístico. A partir de então, qualquer situação envolvendo seleção “ao acaso” de pontos em um segmento, região ou volume pode, e deve, ser modelada indicando-se a distribuição (medida) adotada, assegurando respostas bem definidas. Vale notar que, posteriormente, ramificações especializadas surgiram deste contexto, como a geometria integral e a geometria estocástica, que estudam respectivamente medidas invariantes em espaços geométricos e modelos aleatórios de estruturas geométricas; essas áreas refletem o desdobramento natural do amadurecimento da probabilidade como um todo ao longo do século XX.

Além de seu interesse histórico, a probabilidade geométrica destaca-se por sua relevância teórica e prática na atualidade. Do ponto de vista teórico, ela exemplifica de forma elegante a interface entre probabilidade e análise matemática, fornecendo intuições geométricas para resultados de teoria da medida e integrabilidade de funções (por exemplo, calcular probabilidades equivalentes a determinar áreas sob curvas ou volumes em regiões). Já do ponto de vista aplicado, inúmeros problemas contemporâneos fazem uso de princípios de probabilidade geométrica. Métodos de simulação por Monte Carlo, amplamente empregados em ciência de dados e engenharia, baseiam-se justamente em amostragem aleatória de pontos em domínios geométricos para estimar valores numéricos (como áreas, integrais ou soluções de equações complexas) de forma aproximada, uma estratégia que remonta à ideia de Buffon de estimar π lançando agulhas ao acaso.

Igualmente importante, no âmbito educacional, a probabilidade geométrica fornece

problemas ricos e visualmente atrativos, que têm sido incorporados tanto no ensino médio quanto em competições matemáticas (olimpíadas). Problemas de olimpíadas frequentemente exploram conceitos de probabilidade contínua disfarçados em desafios de geometria, exigindo dos estudantes raciocínio aprofundado sobre áreas e comprimentos. Essa característica interdisciplinar faz da probabilidade geométrica uma ferramenta pedagógica valiosa: ao mesmo tempo em que estimula a intuição geométrica, ela reforça os fundamentos de probabilidade e cálculo, mostrando na prática como conceitos abstratos se aplicam em situações concretas e muitas vezes surpreendentes. Estudar probabilidade geométrica hoje, portanto, é relevante tanto pelo seu valor teórico, ao conectar duas grandes áreas da matemática, fornecendo técnicas para resolução de integrais e compreensão de medidas, quanto pelo seu valor prático, nas aplicações modernas mencionadas e na formação de alunos e resolutores de problemas.

Motivado pelo panorama apresentado, esta dissertação centra-se na aplicação da probabilidade geométrica na resolução de problemas em nível de ensino médio e de olimpíadas de Matemática, explorando como os conceitos clássicos podem ser ensinados de forma acessível e utilizados para solucionar desafios não triviais. No capítulo seguinte, apresenta-se o referencial teórico fundamental sobre probabilidades e medidas geométricas que dará suporte às discussões posteriores, seguido de uma estruturação dos capítulos subsequentes de modo a conectar o contexto histórico e conceitual aqui delineado com os objetivos e metodologia do presente trabalho. Além disso, serão apresentados problemas interessantes de probabilidade geométrica e os métodos utilizados para solucioná-los.

Em suma, partindo das raízes históricas até as aplicações atuais, buscamos evidenciar nesta introdução a evolução e a importância da probabilidade geométrica, estabelecendo o contexto para os desenvolvimentos que serão aprofundados nos capítulos seguintes da dissertação.

2.2 Conjuntos e Definições

Antes de se falar de probabilidade, é necessário elencar algumas propriedades de conjuntos que serão utilizadas ao longo do trabalho para resolver alguns problemas e demonstrar certos teoremas. Além disso, serão definidas uma série de propriedades matemáticas que aparecerão ao longo do texto.

1. **Experimento aleatório:** é um processo que, mesmo repetido diversas vezes sob as mesmas condições, produz resultados imprevisíveis. Lançamento de dados não enviesados de seis faces e verificação da face superior, lançamento de moedas, por exemplo, constituem experimentos aleatórios.
2. **Espaço amostral (Ω):** Conjunto de possíveis resultados de um experimento

aleatório. No caso do dado, as possibilidades são de cair, com a face voltada para cima, os números $\{1, 2, 3, 4, 5, 6\} = \Omega$. Outro tipo de experimento aleatório seria a escolha de um número qualquer em um intervalo real $[a, b]$, com $a < b$. Nesse caso, $[a, b] = \Omega$. O primeiro e segundo exemplos constituem espaços amostrais enumeráveis e não-enumeráveis, respectivamente.

3. **Conjunto Vazio:** Denotado por \emptyset , ele representa o que chamamos de evento impossível.
4. Os elementos (ou pontos) de um espaço amostral podem ser representados genericamente por $w \in \Omega$. Nesse sentido, os pontos w serão associados aos possíveis resultados do experimento. Normalmente, a denominação “ponto” será utilizada quando se tratar de probabilidade geométrica.
5. Qualquer subconjunto de Ω será representado por letras maiúsculas, como $A \subset \Omega$. A ele daremos o nome de “evento”. Diremos que Ω é o evento certo, \emptyset é o evento impossível, e $\{w\}$ o evento elementar.
6. **Números de Elementos:** Caso um espaço amostral seja enumerável e finito, denotaremos por $|\Omega|$ o número de elementos de Ω . Ocasionalmente, pode ser necessário usar $\#\Omega$ para essa finalidade. No caso da probabilidade geométrica, denotaremos a área de uma região (ou volume de um sólido) Ω por $|\Omega|$.
7. **Complementar:** O conjunto A^c é o complementar de A , isto é, todos os elementos $w \in \Omega$, tais que $w \notin A$. Pode ser reescrito como $A^c = \Omega - A$.
8. **União:** $\bigcup_{i=1}^n A_i = A_1 \cup A_2 \cup \dots \cup A_n$ representa os pontos de Ω tal que eles pertencem a, pelo menos, um dos A_i , com $1 \leq i \leq n$.
9. **Intersecção:** $\bigcap_{i=1}^n A_i = A_1 \cap A_2 \cap \dots \cap A_n$ representa os pontos de Ω tal que eles pertencem, simultaneamente, a todos os A_i , com $1 \leq i \leq n$.
10. **Conjuntos Disjuntos:** Caso A e B sejam subconjuntos de Ω e se tenha $A \cap B = \emptyset$, diremos que esses conjuntos são disjuntos, ou ainda, não comungam de elementos em comum.
11. **Diferença:** Dado $A, B \subset \Omega$, define-se a diferença como o conjunto dos elementos que pertencem a A e não pertencem a B , isto é,

$$A - B = \{w \in \Omega \mid w \in A \text{ e } w \notin B\}.$$

12. **Diferença simétrica:** Dado $A, B \subset \Omega$, define-se a diferença simétrica como o conjunto dos elementos que pertencem a um dos dois conjuntos, mas não a ambos, isto é,

$$A \Delta B = (A - B) \cup (B - A) = (A \cup B) - (A \cap B).$$

13. **Conjunto das Partes de Ω :** Denotado por $\mathcal{P}(\Omega)$, para não confundir posteriormente com a probabilidade ($P(\Omega)$), é o conjunto que tem todos os subconjuntos de Ω .

Um conceito que será amplamente utilizado neste trabalho é a definição de classe de subconjuntos, a qual definiremos agora.

Definição 2.2.1 (Classe de Subconjuntos:). *Seja Ω um conjunto não vazio. Denomina-se classe de subconjuntos de Ω qualquer coleção \mathcal{C} cujos elementos pertencem à $\mathcal{P}(\Omega)$, isto é*

$$\mathcal{C} \subseteq \mathcal{P}(\Omega)$$

em que $\mathcal{P}(\Omega)$ é o conjunto das partes de Ω .

A proposição à seguir ajudará a provar alguns fatos ocasionalmente. Ela é conhecida como “Leis de Morgan”.

Proposição 2.2.2 (Leis de Morgan). *Sejam A_1, \dots, A_n subconjuntos de Ω (esse sendo o conjunto universo). Então*

$$\left(\bigcup_{j=1}^n A_j \right)^c = \bigcap_{j=1}^n A_j^c \text{ e } \left(\bigcap_{j=1}^n A_j \right)^c = \bigcup_{j=1}^n A_j^c$$

Demonstração: *Para provar uma relação de igualdade entre conjuntos, basta que* $\left(\bigcup_{j=1}^n A_j \right)^c \subset \bigcap_{j=1}^n A_j^c$ *e* $\bigcap_{j=1}^n A_j^c \subset \left(\bigcup_{j=1}^n A_j \right)^c$. *Se* $w \in \left(\bigcup_{j=1}^n A_j \right)^c \Rightarrow w \notin \bigcup_{j=1}^n A_j$. *Como*

$w \notin A_j$, *para todo* $1 \leq j \leq n$, *logo* w *pertence a todos* A_j^c , *com* $1 \leq j \leq n$. *Daí* $w \in \bigcap_{j=1}^n A_j^c$

e $\left(\bigcup_{j=1}^n A_j \right)^c \subset \bigcap_{j=1}^n A_j^c$. *Por outro lado, se* $w \in \bigcap_{j=1}^n A_j^c \Rightarrow w \notin \bigcup_{j=1}^n A_j \Rightarrow w \in \left(\bigcup_{j=1}^n A_j \right)^c$.

Logo $\bigcap_{j=1}^n A_j^c \subset \left(\bigcup_{j=1}^n A_j \right)^c$ *e, portanto,* $\left(\bigcup_{j=1}^n A_j \right)^c = \bigcap_{j=1}^n A_j^c$. *Utilizando o mesmo método para a segunda igualdade, teremos:*

$$w \in \left(\bigcap_{j=1}^n A_j \right)^c \Rightarrow w \notin \bigcap_{j=1}^n A_j \Rightarrow w \in \bigcup_{j=1}^n A_j^c \Rightarrow \left(\bigcap_{j=1}^n A_j \right)^c \subset \bigcup_{j=1}^n A_j^c$$

Por outro lado,

$$w \in \bigcup_{j=1}^n A_j^c \Rightarrow w \notin \bigcap_{j=1}^n A_j \Rightarrow w \in \left(\bigcap_{j=1}^n A_j \right)^c \Rightarrow \bigcup_{j=1}^n A_j^c \subset \left(\bigcap_{j=1}^n A_j \right)^c$$

Teremos, portanto, $\left(\bigcap_{j=1}^n A_j\right)^c = \bigcup_{j=1}^n A_j^c$.

Por fim, se tivermos uma sequência $(A_n)_{n \in \mathbb{N}}$, utilizando as argumentações anteriores de forma análoga, obtemos

$$\left(\bigcup_{n=1}^{\infty} A_n\right)^c = \bigcap_{n=1}^{\infty} A_n^c \text{ e } \left(\bigcap_{n=1}^{\infty} A_n\right)^c = \bigcup_{n=1}^{\infty} A_n^c.$$

□

A formalização da probabilidade em espaços contínuos exige a introdução de uma estrutura matemática que garanta fechamento sob operações como união enumerável e complemento. Essa estrutura é dada pela noção de σ -álgebra, fundamental para a construção rigorosa da teoria da medida e, conseqüentemente, da probabilidade geométrica. Segue:

Definição 2.2.3 (σ -álgebra). *Uma classe de subconjuntos de Ω , representada por \mathcal{F} , é denominada uma σ -álgebra se satisfaz as seguintes propriedades:*

1. $\Omega \in \mathcal{F}$;
2. Se $A \in \mathcal{F}$, então $A^c \in \mathcal{F}$;
3. Se $A_i \in \mathcal{F}$, $i \geq 1$, então $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

Na definição 2.2.3, percebemos que, para uma classe ser uma σ -álgebra (lê-se “sigma álgebra”), essa classe deve satisfazer os três requisitos, pois nem toda classe de Ω satisfará os três simultaneamente. No exemplo 2.2.4, são apresentadas duas situações, uma em que se tem uma σ -álgebra, e outra que não é uma σ -álgebra.

Exemplo 2.2.4. *Seja o conjunto $\Omega = \{a, b, c, d\}$ e considere:*

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

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

Perceba que $\Omega \in \mathcal{F}_1$, para cada conjunto pertencente a \mathcal{F}_1 , o complementar também pertence e quaisquer uniões de subconjuntos pertencentes a \mathcal{F}_1 também pertence à ele. Logo, \mathcal{F}_1 é uma σ -álgebra. Por outro lado, percebe-se facilmente que 1. e 2. da definição 2.2.3 é válida na classe \mathcal{F}_2 . No entanto, perceba que $\{a\} \cup \{c, d\} \notin \mathcal{F}_2$. Logo, \mathcal{F}_2 não é uma σ -álgebra.

No exemplo 2.2.4, perceba que, dadas as condições estabelecidas, \mathcal{F}_1 é a menor σ -álgebra que possui o elemento $\{a\}$. Sendo mais específico, qualquer outra σ -álgebra que tenha $\{a\}$ terá mais elementos que \mathcal{F}_1 . Por isso, de forma mais genérica, se tivermos

$A \subset \Omega$ e $\mathcal{F} = \{\emptyset, \Omega, A, A^c\}$, essa será a σ -álgebra gerada por A . Outros que sejam maiores, diremos que são \mathcal{F} geradas por A e denotaremos por $\sigma(A)$. Por fim, perceba que os dois primeiros passos são simples de verificar e o terceiro é o mais trabalhoso.

Nos exemplos apresentados, foram trabalhados conjuntos enumeráveis. Um conjunto X é dito enumerável quando ele é finito ou quando existe uma bijeção $f : \mathbb{N} \rightarrow X$. Exemplos de conjuntos enumeráveis são os naturais, os inteiros, os racionais e o conjunto trabalhado no exemplo 2.2.4. Caso contrário, diremos que ele é não enumerável. A definição apresentada está presente em (LIMA, 2013), além do fato de demonstrar alguns teoremas que serão muito úteis ao longo deste texto, como, por exemplo, que o conjunto dos números reais \mathbb{R} é não enumerável, assim como intervalos reais não degenerados. Perceba que, ao trabalhar com segmentos no plano cartesiano, eles constituem, assim, conjuntos não enumeráveis. Estendendo esse raciocínio para o \mathbb{R}^2 e \mathbb{R}^3 , as regiões do plano e do espaço associadas a áreas e volumes não nulos também formam conjuntos não enumeráveis. Logo, na probabilidade geométrica, os espaços amostrais considerados serão, por via de regra, não enumeráveis.

No exemplo 2.2.5 é apresentada uma σ -álgebra cujo espaço amostral será um conjunto não enumerável, a saber, \mathbb{R} .

Exemplo 2.2.5. *Seja $\Omega = \mathbb{R}$ e seja*

$$\mathcal{F} = \{\emptyset, \Omega, (-\infty, 2], (2, +\infty), (-\infty, 1), [1, +\infty), (-\infty, 1) \cup (2, +\infty), [1, 2]\}.$$

A classe \mathcal{F} é uma σ -álgebra. Para verificar, basta perceber que todos os requisitos da definição 2.2.3 são válidos.

Teorema 2.2.6. *Seja \mathcal{F} uma σ -álgebra de subconjuntos de Ω . Se $A_1, A_2, \dots \in \mathcal{F}$, então*

$$\bigcap_{n=1}^{\infty} A_n \in \mathcal{F}.$$

Demonstração: *Tendo $A_n \in \mathcal{F}$, pelo fechamento por complemento (A_2), temos $A_n^c \in \mathcal{F}$ para todo n . Pelo fechamento por uniões enumeráveis (A_3), temos $\bigcup_{n=1}^{\infty} A_n^c \in \mathcal{F}$. Novamente pelo fechamento por complemento, temos*

$$\left(\bigcup_{n=1}^{\infty} A_n^c \right)^c \in \mathcal{F}.$$

Por outro lado, pelas Leis de Morgan 2.2.2, temos $\left(\bigcup_{n=1}^{\infty} A_n^c \right)^c = \bigcap_{n=1}^{\infty} A_n$. Daí, obtemos $\bigcap_{n=1}^{\infty} A_n \in \mathcal{F}$. Lembre-se que $(A^c)^c = A$. □

Na definição 2.2.7, será apresentada uma σ -álgebra muito importante, conhecida como σ -álgebra de Borel.

Definição 2.2.7. A menor σ -álgebra em $\Omega = \mathbb{R}$ que contém todos os os intervalos abertos de \mathbb{R} é chamada de σ -álgebra de Borel, denotada por $\mathcal{B}(\mathbb{R})$. Um elemento dessa σ -álgebra é denominado por boreliano.

O interessante da σ -álgebra de Borel é que ela pode ser gerada pelos intervalos da forma $(-\infty, x)$, com $x \in \mathbb{R}$. O fato principal dessa afirmação é que qualquer intervalo real pode ser obtido por meio de operações de união ou intersecção com intervalos desse tipo. Segue a:

Proposição 2.2.8. A σ -álgebra de Borel em \mathbb{R} (denotada por $\mathcal{B}(\mathbb{R})$) pode ser gerada pelos intervalos da forma $(-\infty, x)$, com $x \in \mathbb{R}$, isto é,

$$\mathcal{B}(\mathbb{R}) = \sigma\left(\{(-\infty, x) : x \in \mathbb{R}\}\right).$$

Demonstração: Seja $\mathcal{F} = \sigma\left(\{(-\infty, x) : x \in \mathbb{R}\}\right)$. Como \mathcal{F} é uma σ -álgebra, ela é fechada por complementos, uniões e intersecções enumeráveis. Assim, para cada $x \in \mathbb{R}$,

$$(-\infty, x] = \bigcap_{n=1}^{\infty} \left(-\infty, x + \frac{1}{n}\right) \in \mathcal{F}, \text{ e } (x, +\infty) = ((-\infty, x])^c \in \mathcal{F}.$$

Logo, para $a < b$,

$$(a, b) = (-\infty, b) \cap (a, +\infty) \in \mathcal{F}.$$

Como todo aberto $U \subset \mathbb{R}$ pode ser escrito como união enumerável de intervalos abertos (a, b) , segue que $U \in \mathcal{F}$ e, portanto, $\sigma(\text{abertos}) \subset \mathcal{F}$.

Por outro lado, cada $(-\infty, x)$ é um aberto, logo $\mathcal{F} \subset \sigma(\text{abertos})$. Concluimos, então, que

$$\mathcal{B}(\mathbb{R}) = \sigma(\text{abertos}) = \sigma\left(\{(-\infty, x) : x \in \mathbb{R}\}\right).$$

□

Na prática, isso significa que todos os intervalos, semirretas, conjuntos fechados, e mesmo uniões ou intersecções contáveis desses conjuntos pertencem a $\mathcal{B}(\mathbb{R})$. Portanto, é um conjunto suficientemente amplo para lidar com qualquer situação que apareça em aplicações reais.

Uma coleção não-vazia de subconjuntos de um Ω dado que seja fechado para finitas operações de teoria dos conjuntos é chamada de *álgebra de subconjuntos de Ω* . O que a diferencia de σ -álgebra é o fato de que, essa última, relaciona infinitas e enumeráveis operações, configurando-se como uma coleção mais restritiva. (STONE, 1978).

2.3 Probabilidade

O estudo desenvolvido até aqui estabeleceu a base conceitual necessária para o tratamento rigoroso de situações aleatórias no contexto matemático. Na seção anterior,

foram apresentados elementos fundamentais da Teoria dos Conjuntos e estruturas algébricas, bem como a construção formal de σ -álgebras, em especial a σ -álgebra de Borel, que servem como domínio de eventos nos quais a probabilidade poderá ser definida de forma consistente. Essa etapa preliminar não é apenas formal, ela assegura que os objetos matemáticos manipulados possuem propriedades que permitem operações bem definidas, como união contável, interseção e complementar, além da mensurabilidade necessária para a atribuição de valores numéricos coerentes. Nesse último caso, a teoria complementar será vista na seção [3.2](#).

A partir desta fundamentação, torna-se possível evoluir da simples noção intuitiva de acaso para uma formulação axiomática de probabilidade. Neste capítulo, portanto, será introduzida a definição formal de probabilidade, com base nos axiomas propostos por Andrei Kolmogorov em 1933, marco que consolidou a probabilidade como ramo rigoroso da matemática. Serão discutidas suas propriedades básicas, consequências lógicas do modelo axiomático e a estrutura geral de um espaço de probabilidade.

Ao longo do capítulo, a abordagem alternará entre o rigor matemático necessário ao nível universitário e interpretações didáticas que permitam compreender como tais conceitos se conectam a problemas reais e a aplicações que serão exploradas posteriormente, especialmente no contexto da probabilidade geométrica. Assim, do ponto de vista estrutural, este capítulo constitui a ponte entre a formalização dos objetos matemáticos apresentada anteriormente e os modelos probabilísticos que fundamentarão as análises desenvolvidas nos capítulos seguintes. Os conceitos apresentados a seguir foram vistos em ([MAGALHÃES, 2024](#)) e ([JAMES, 2010](#)).

Antes de apresentarmos o modelo axiomático de Kolmogorov, é útil retomar a noção intuitiva de probabilidade em contextos simples, nos quais o conjunto de resultados possíveis é finito e todos os resultados são considerados igualmente prováveis. Esse é o caso clássico de experimentos como o lançamento de um dado, o sorteio de uma carta ou o giro de uma roleta com divisões idênticas.

Nessas situações, a probabilidade pode ser entendida como uma razão entre quantidades: conta-se quantos resultados favoráveis ao evento desejado ($|A|$) existem e divide-se pelo total de resultados possíveis no experimento ($|\Omega|$).

$$P(A) = \frac{|A|}{|\Omega|}$$

Essa interpretação oferece uma porta de entrada natural para o conceito formal, ajudando a construir a intuição necessária antes de generalizá-lo para contextos mais abstratos, contínuos ou não enumeráveis.

Abaixo, apresentaremos a definição formal de probabilidade, baseada em um conjunto de axiomas propostos por A. N. Kolmogorov. Perceba que a definição informal

de probabilidade deriva da definição formal.

Definição 2.3.1. *Uma função P , definida na σ -álgebra \mathcal{F} de subconjuntos de Ω e com valores em $[0, 1]$, é uma probabilidade se satisfaz os axiomas de Kolmogorov:*

$$A1 \quad P(\Omega) = 1$$

$$A2 \quad \text{Para todo subconjunto } A \in \mathcal{F}, P(A) \geq 0;$$

A3 Para toda sequência $A_1, A_2, \dots \in \mathcal{F}$, mutuamente exclusivos, temos

$$P\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} P(A_j)$$

A trinca (Ω, \mathcal{F}, P) é denominada *espaço de probabilidade*. Os subconjuntos em \mathcal{F} são denominados eventos, e somente a eles se atribui probabilidade. O espaço de probabilidade forma a base para o tratamento de fenômenos aleatórios (fenômenos aleatórios são experimentos cujos resultados não podem ser previstos com certeza, mesmo quando repetidos nas mesmas condições, como o lançamento de uma moeda ou um dado). Na prática, o A1 afirma que a probabilidade do evento certo é 1. O A2 afirma que a probabilidade de qualquer evento acontecer é positiva ou nula, nunca será negativa. Por fim, o A3 afirma que se os eventos possuem interseção vazia, basta somar suas probabilidades individualmente.

Exemplo 2.3.2. *Considere o lançamento de um dado de seis lados honesto e a observação da face superior. As possibilidades observáveis seriam $\Omega = \{1, 2, 3, 4, 5, 6\}$. Definamos a σ -álgebra $\mathcal{F} = \mathcal{P}(\Omega)$. Não é difícil perceber que o conjunto das partes é uma σ -álgebra. Por fim, definamos a função $P(A) = \frac{|A|}{|\Omega|} = \frac{|A|}{6}$. Fazemos isso partindo do princípio da indiferença, que, na probabilidade, afirma que, na ausência de informações, todos os resultados possíveis de um evento devem ter probabilidades iguais. Logo $P(i) = \frac{1}{6} \forall i \in \Omega$. Chamaremos essa situação de resultados equiprováveis, como explica [James \(2010\)](#). Observe que, definido assim, a trinca (Ω, \mathcal{F}, P) configurará um espaço de probabilidade. Verifiquemos os axiomas um a um.*

$$A1 \quad P(\Omega) = \frac{|\Omega|}{6} = 1.$$

A2 Perceba que $|A| \geq 0$ para todo $A \in \mathcal{F}$. Logo, $P(A) \geq 0$, para todo $A \in \mathcal{F}$.

A3 Seja $(A_n)_{n \geq 1} \subset \mathcal{F}$ uma família de eventos dois a dois disjuntos. Como Ω é finito, no máximo seis dos A_n podem ser não vazios (pois, se $A_n \neq \emptyset$, podemos escolher um elemento distinto $a_n \in A_n$, produzindo uma coleção de elementos distintos de Ω). Portanto, o conjunto de índices $I = \{n \geq 1 : A_n \neq \emptyset\}$ é finito, com $|I| \leq 6$. Assim,

$$\bigcup_{n \geq 1} A_n = \bigcup_{n \in I} A_n \quad e \quad \left| \bigcup_{n \in I} A_n \right| = \sum_{n \in I} |A_n|,$$

pois os A_n são disjuntos. Assim, calculando a probabilidade da união, obtemos

$$P\left(\bigcup_{n \geq 1} A_n\right) = \frac{1}{6} \left| \bigcup_{n \in I} A_n \right| = \frac{1}{6} \sum_{n \in I} |A_n| = \sum_{n \in I} \frac{|A_n|}{6} = \sum_{n \geq 1} P(A_n).$$

Tendo mostrado a validade dos axiomas, podemos calcular a probabilidade de eventos específicos. Por exemplo:

1. $A_1 = \{2, 4, 6\}$. Esse conjunto seria sinônimo de calcular a probabilidade de a face ser par. Vale ressaltar que deve-se parar de contar no 6, pois Ω é o espaço amostral e seus valores variam apenas de 1 a 6. Perceba também que $A_1 \in \mathcal{F}$, logo podemos calcular sua probabilidade. Daí, obtemos que $P(A_1) = \frac{3}{6} = \frac{1}{2}$.
2. $A_2 = \{3, 5\}$. Pode-se dizer que esse conjunto é formado pelos ímpares maiores do que 2 e sua probabilidade será $P(A_2) = \frac{2}{6} = \frac{1}{3}$.
3. Vale ressaltar que, no caso de A_1 e A_2 , o axioma A3 é válido, já que $A_1 \cap A_2 = \emptyset$. Assim, $P(A_1 \cup A_2) = \frac{1}{2} + \frac{1}{3} = \frac{5}{6}$, que equivaleria à probabilidade da face ser par ou um valor ímpar maior do que 2.

Neste momento, uma pergunta pertinente que o leitor pode se fazer é: qual a razão da construção de uma σ -álgebra para a definição formal de probabilidade? Se pensarmos um pouco, de fato, ela não é necessária para casos em que o Ω seja finito ou enumerável. Isso se deve ao fato de que o conjunto das partes de Ω é uma σ -álgebra (JAMES, 2010). Os casos em que é de fato relevante a construção de uma σ -álgebra é quando temos conjuntos não enumeráveis, como a reta real, o plano ou o espaço, que farão parte do que chamamos de probabilidade geométrica. Por isso, a justificativa da construção da σ -álgebra se dará na seção 3.1. Ademais, apresentaremos agora algumas propriedades da probabilidade que servirão para qualquer espaço de probabilidade, seja com conjuntos discretos ou contínuos. Segue:

Proposição 2.3.3. *Dado o espaço de probabilidade $(\Omega, \mathcal{F}, \mathcal{P})$, considere que os conjuntos mencionados abaixo são eventos desse espaço de probabilidade. Temos:*

P1. $P(A) = 1 - P(A^c)$.

Demonstração: Pelo axioma 3 e 1, temos que

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

Logo $P(A) = 1 - P(A^c)$. Um detalhe interessante é que

$$P(\Omega) = 1 - P(\Omega^c) = 1 - P(\emptyset)$$

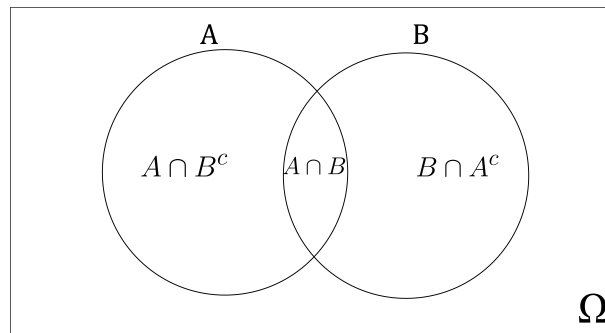
$$P(\emptyset) = 0.$$

□

P2. Sendo A e B dois eventos quaisquer, vale $P(B) = P(B \cap A) + P(B \cap A^c)$.

Demonstração: Para dois eventos quaisquer A e B , sempre é possível reescrever B como a união disjunta $B = (B \cap A) \cup (B \cap A^c)$, como mostrado na figura 1. Supondo que essa união não fosse disjunta, se $w \in A$, teríamos $w \in B$ e conseqüentemente $w \in A^c$, já que existiria pelo menos um elemento $w \in B \cap A$ e $w \in B \cap A^c$, que é um absurdo, já que um elemento não pode pertencer simultaneamente a um conjunto e ao seu complementar. Assim, teremos, pelo axioma A3, que $P(B) = P(B \cap A) + P(B \cap A^c)$. \square

Figura 1 – União disjunta $B = (B \cap A) \cup (B \cap A^c)$.



Fonte: Elaborada pelo autor (2026).

P3. Se $A \subset B$, então $P(A) \leq P(B)$.

Demonstração: Pela propriedade anterior, temos $B = (B \cap A) \cup (B \cap A^c) = A \cup (B \cap A^c)$. Essa última sendo disjunta. Pelo axioma A3, temos que $P(B) = P(A) + P(B \cap A^c)$. Como pelo A2 toda probabilidade deve ser não negativa, temos, portanto, $P(B) \geq P(A)$ ou $P(A) \leq P(B)$. \square

P4. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Demonstração: Utilizaremos o fato de que $A \cup B = B \cup (A \cap B^c)$ é uma união disjunta. Pela P2, sabemos que

$$P(A) = P(A \cap B) + P(A \cap B^c) \text{ e } P(B) = P(B \cap A) + P(B \cap A^c)$$

Organizando a equação, obtemos

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

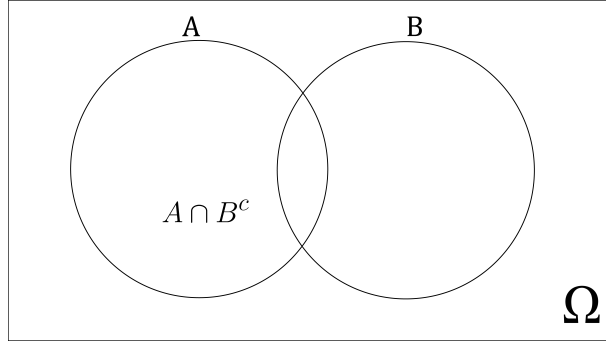
Pela P2, sabemos que $P(A \cap B) + P(B \cap A^c) = P(B)$. Substituindo, obtemos

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

Perceba que $A \cup B$ pode ser reescrita como a união disjunta $B \cup (A \cap B^c)$, como é observável na figura 2. Assim, do A3, temos $P(A \cup B) = P(B) + P(A \cap B^c)$. Substituindo, obteremos, portanto,

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

Figura 2 – $A - B = A \cap B^c$ considerando o espaço amostral Ω .



Fonte: Elaborada pelo autor (2026).

Um detalhe importante é que, se $A \cap B = \emptyset$, teremos $P(A \cup B) = P(A) + P(B)$. \square

P5. Para quaisquer eventos A_j , temos

$$P\left(\bigcup_{j=1}^{\infty} A_j\right) \leq \sum_{j=1}^{\infty} P(A_j).$$

Demonstração: Perceba que podemos reescrever a soma $\bigcup_{j=1}^{\infty} A_j$ como uma união disjunta da seguinte forma:

$$\bigcup_{j=1}^{\infty} A_j = A_1 \cup (A_1^c \cap A_2) \cup (A_1^c \cap A_2^c \cap A_3) \cup \dots$$

Daí, pelo A3, temos

$$P\left(\bigcup_{j=1}^{\infty} A_j\right) = P(A_1) + P(A_1^c \cap A_2) + P(A_1^c \cap A_2^c \cap A_3) + \dots$$

Perceba que $A_1^c \cap A_2^c \cap A_3^c \cap \dots \cap A_n \subset A_n$, para todo n . Assim, pela P3, obtemos

$$P\left(\bigcup_{j=1}^{\infty} A_j\right) = P(A_1) + P(A_1^c \cap A_2) + P(A_1^c \cap A_2^c \cap A_3) + \dots \leq \sum_{j=1}^{\infty} P(A_j).$$

\square

P6. Se $A_n \uparrow A$, então $P(A_n) \uparrow P(A)$. De forma similar, se $A_n \downarrow A$, então $P(A_n) \downarrow P(A)$.

Demonstração: Mostraremos separadamente os casos crescente e decrescente.

(i) **Caso crescente.** Seja $(A_n)_{n \geq 1}$ uma sequência de eventos tal que

$$A_1 \subseteq A_2 \subseteq A_3 \subseteq \dots \quad e \quad A = \bigcup_{n=1}^{\infty} A_n.$$

Defina os conjuntos

$$D_1 := A_1, \quad D_n := A_n \setminus A_{n-1} \quad (n \geq 2).$$

Pela construção, os conjuntos D_k são dois a dois disjuntos e satisfazem

$$A_m = \bigcup_{k=1}^m D_k \quad (m \geq 1), \quad A = \bigcup_{k=1}^{\infty} D_k.$$

Pela aditividade finita da probabilidade (P4), obtemos

$$P(A_m) = \sum_{k=1}^m P(D_k), \quad m \geq 1.$$

Logo, a sequência $(P(A_m))$ é monótona crescente e limitada superiormente por $P(\Omega) = 1$, sendo, portanto, convergente. Denotemos

$$L := \lim_{m \rightarrow \infty} P(A_m) = \sum_{k=1}^{\infty} P(D_k).$$

Como $A_m \subseteq A$ para todo m , pela monotonicidade da probabilidade (P3), tem-se

$$P(A_m) \leq P(A) \quad \text{para todo } m,$$

o que implica $L \leq P(A)$.

Por outro lado, pela subaditividade enumerável (P5),

$$P(A) = P\left(\bigcup_{k=1}^{\infty} D_k\right) \leq \sum_{k=1}^{\infty} P(D_k) = L.$$

Assim, $P(A) = L$, isto é,

$$\lim_{m \rightarrow \infty} P(A_m) = P(A).$$

(ii) Caso decrescente. Seja $(B_n)_{n \geq 1}$ uma sequência de eventos tal que

$$B_1 \supseteq B_2 \supseteq B_3 \supseteq \cdots \quad \text{e} \quad B = \bigcap_{n=1}^{\infty} B_n.$$

Considere os complementares B_n^c . Então

$$B_1^c \subseteq B_2^c \subseteq \cdots \quad \text{e} \quad B^c = \bigcup_{n=1}^{\infty} B_n^c.$$

Logo, a sequência (B_n^c) é crescente. Pelo item (i),

$$\lim_{n \rightarrow \infty} P(B_n^c) = P(B^c).$$

Utilizando a propriedade do complementar (P1) e passando ao limite, obtemos

$$\lim_{n \rightarrow \infty} P(B_n) = \lim_{n \rightarrow \infty} (1 - P(B_n^c)) = 1 - P(B^c) = P(B).$$

Isso conclui a prova. □

2.4 Probabilidade Condicional

Na teoria das probabilidades, frequentemente lidamos com situações em que a informação disponível restringe o universo de resultados possíveis. Por exemplo, ao saber que um aluno pertence ao Clube Olímpico de Matemática da escola, a chance de ele ser medalhista não é mais avaliada sobre todo o conjunto de estudantes, mas apenas sobre o subconjunto dos que pertencem ao clube. Esse tipo de raciocínio exige uma reformulação da medida de probabilidade, levando ao conceito de **probabilidade condicional**.

Para ilustrar essa ideia, considere o seguinte exemplo.

Exemplo 2.4.1. *Em uma escola estadual com 200 estudantes, 10 foram premiados com medalhas na Olimpíada Mandacaru de Matemática. Dentre esses 10 medalhistas, 8 pertencem ao Clube Olímpico de Matemática da escola, que é composto exatamente por 10 estudantes.*

1. *Ao escolher um estudante aleatoriamente na escola, qual é a probabilidade de ele ser medalhista?*
2. *Ao escolher um estudante sabendo que ele pertence ao Clube Olímpico de Matemática, qual é a probabilidade de ele ter sido medalhista na Olimpíada Mandacaru?*

Solução:

1. *Como há 10 medalhistas em um total de 200 estudantes, a probabilidade de que um estudante escolhido ao acaso na escola seja medalhista é*

$$P(\text{medalhista}) = \frac{10}{200} = \frac{1}{20}.$$

2. *Agora, ao sabermos que o estudante escolhido pertence ao Clube Olímpico de Matemática, o universo de referência deixa de ser o conjunto de todos os estudantes da escola e passa a ser o conjunto dos 10 membros do clube. Dentre eles, 8 foram medalhistas. Assim, a probabilidade procurada é*

$$P(\text{medalhista} \mid \text{clube}) = \frac{8}{10} = \frac{4}{5}.$$

Esse exemplo evidencia que a probabilidade de um mesmo evento pode variar significativamente quando informações adicionais são incorporadas ao modelo probabilístico. Em particular, ao restringirmos o espaço amostral a um subconjunto relevante, torna-se necessário redefinir a forma como a probabilidade é calculada. Esse procedimento conduz naturalmente ao conceito de **probabilidade condicional**.

Formalmente, considere um espaço de probabilidade (Ω, \mathcal{F}, P) . Agora, suponha que $B \in \mathcal{F}$ seja um evento com $P(B) > 0$. Queremos medir a probabilidade de outro

evento $A \in \mathcal{F}$, sabendo que B ocorreu. Intuitivamente, restringimos o espaço amostral apenas a B , tornando-o o novo “universo” de referência. Daí, surge a:

Definição 2.4.2. *Considere os eventos A e B em $(\Omega, \mathcal{F}, \mathcal{P})$. Sendo $P(B) > 0$, a probabilidade condicional de A , dado que ocorreu B , é dada por*

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

Caso $P(B) = 0$, definimos $P(A|B) = P(A)$.

Segundo [Magalhães \(2024\)](#), no caso $P(B) = 0$, alguns autores preferem igualar a probabilidade condicional à zero ou considerá-la indefinida.

Proposição 2.4.3. *Seja (Ω, \mathcal{F}, P) um espaço de probabilidade e $A_1, A_2, \dots, A_n \in \mathcal{F}$ eventos tais que $P\left(\bigcap_{j=1}^n A_j\right) > 0$. Então, vale a Regra do Produto:*

$$P(A_1 \cap A_2 \cap \dots \cap A_n) = P(A_1) \cdot P(A_2 | A_1) \cdot P(A_3 | A_1 \cap A_2) \cdots P(A_n | A_1 \cap A_2 \cap \dots \cap A_{n-1}).$$

Essa igualdade expressa a decomposição da probabilidade conjunta em termos de probabilidades condicionais sucessivas.

Demonstração: Utilizaremos indução em n . Observemos, primeiramente, o caso $n = 2$. Pela definição de probabilidade condicional,

$$P(A_1 \cap A_2) = P(A_1) \frac{P(A_1 \cap A_2)}{P(A_1)} = P(A_1) P(A_2 | A_1),$$

já que $P\left(\bigcap_{j=1}^n A_j\right) > 0 \implies P(A_1) > 0$.

Assuma agora que tal proposição é válida para $n - 1$ eventos, isto é,

$$P\left(\bigcap_{i=1}^{n-1} A_i\right) = P(A_1) \prod_{k=2}^{n-1} P\left(A_k | A_1 \cap \dots \cap A_{k-1}\right).$$

Considere n eventos. Pelo caso $n = 2$ aplicado aos eventos $\bigcap_{i=1}^{n-1} A_i$ e A_n ,

$$P\left(\bigcap_{i=1}^n A_i\right) = P\left(\bigcap_{i=1}^{n-1} A_i\right) P\left(A_n | \bigcap_{i=1}^{n-1} A_i\right).$$

Substituindo a hipótese de indução obtemos

$$P\left(\bigcap_{i=1}^n A_i\right) = P(A_1) \prod_{k=2}^{n-1} P\left(A_k | A_1 \cap \dots \cap A_{k-1}\right) \cdot P\left(A_n | \bigcap_{i=1}^{n-1} A_i\right),$$

que é exatamente a fórmula desejada para n eventos. Isso completa a prova por indução.

□

Exemplo 2.4.4. Numa lagoa de um pesque e pague há 6 tilápias, 3 bagres, 10 tambacus e 4 carpas, totalizando 23 peixes. Um pescador captura aleatoriamente 4 peixes sem reposição, todas as extrações sendo equiprováveis. Qual a probabilidade de que os quatro peixes capturados sejam tilápias?

Solução. Considere que cada A_i corresponde ao evento "o i -ésimo peixe capturado é uma tilápia", para $i = 1, 2, 3, 4$, e tome como A o evento "pescar quatro tilápias". Então temos

$$P(A) = P(A_1) P(A_2 | A_1) P(A_3 | A_1 \cap A_2) P(A_4 | A_1 \cap A_2 \cap A_3)$$

$$P(A) = \frac{6}{23} \cdot \frac{5}{22} \cdot \frac{4}{21} \cdot \frac{3}{20} = \frac{3}{1771}.$$

Alternativamente, usando a abordagem combinatória (contagem de casos favoráveis sobre casos possíveis), obtemos a mesma resposta:

$$P(A) = \frac{\binom{6}{4} \binom{23}{0}}{\binom{23}{4}} = \frac{15}{8855} = \frac{3}{1771}.$$

A inspiração para esse problema veio pela leitura de [James \(2010\)](#).

Teorema 2.4.5. Seja (B_1, B_2, \dots, B_n) uma partição de Ω , isto é, eventos em \mathcal{F} tais que

$$B_i \cap B_j = \emptyset \text{ para } i \neq j, \quad e \quad \bigcup_{i=1}^n B_i = \Omega,$$

com $P(B_i) > 0$ para todo i . Então, para qualquer evento $A \in \mathcal{F}$, vale a Lei da Probabilidade Total:

$$P(A) = \sum_{i=1}^n P(A | B_i) P(B_i).$$

Essa lei permite calcular $P(A)$ somando as probabilidades condicionais de A em cada subconjunto B_i ponderadas pelas probabilidades de cada B_i .

Demonstração: Como $\{B_i\}_{i=1}^n$ é uma partição de Ω , temos a decomposição

$$A = A \cap \Omega = A \cap \left(\bigcup_{i=1}^n B_i \right) = \bigcup_{i=1}^n (A \cap B_i),$$

onde os conjuntos $A \cap B_i$ e $A \cap B_j$ são disjuntos se $i \neq j$. Pelo A3, segue

$$P(A) = P\left(\bigcup_{i=1}^n (A \cap B_i) \right) = \sum_{i=1}^n P(A \cap B_i).$$

Para cada i com $P(B_i) > 0$ aplica-se a definição de probabilidade condicional:

$$P(A \cap B_i) = P(B_i) P(A | B_i).$$

Substituindo na soma obtemos a identidade desejada:

$$P(A) = \sum_{i=1}^n P(B_i) P(A | B_i).$$

□

Observação 2.4.6. Se algum $P(B_i) = 0$, então $P(A \cap B_i) = 0$ e o termo correspondente na soma é zero; por isso a fórmula continua válida interpretando-se $P(B_i)P(A | B_i) = 0$ quando $P(B_i) = 0$.

Exemplo 2.4.7. Durante um jogo de tabuleiro, um personagem está amaldiçoado e para quebrar a maldição a próxima rolagem de dado feita deve resultar em um número menor ou igual a 3. Tal jogo utiliza dados justos de 6, 8, 10 e 20 lados, que possuem diferentes chances de serem os próximos rolados:

tipo do dado	D_1	D_2	D_3	D_4
número de faces d	6	8	10	20
probabilidade de rolagem $P(B_i)$	$\frac{1}{10}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{2}$

Qual a probabilidade de que, após a rolagem, o personagem quebre a maldição?

Solução. Seja B_i , com $i = 1, 2, 3, 4$, o evento “foi escolhido o dado D_i ” e A o evento “a rolagem resulta em um valor ≤ 3 ”.

Pela Lei da Probabilidade Total, usando a partição $\{B_1, B_2, B_3, B_4\}$, temos

$$P(A) = \sum_{i=1}^4 P(B_i) P(A | B_i).$$

Para um dado com d faces (e $d \geq 3$), a probabilidade condicional de se obter um número menor ou igual a 3 é $P(A | B_i) = \frac{3}{d}$. Logo

$$P(A | B_1) = \frac{3}{6} = \frac{1}{2},$$

$$P(A | B_2) = \frac{3}{8},$$

$$P(A | B_3) = \frac{3}{10},$$

$$P(A | B_4) = \frac{3}{20}.$$

Substituindo na soma ponderada,

$$\begin{aligned} P(A) &= \frac{1}{10} \cdot \frac{1}{2} + \frac{1}{5} \cdot \frac{3}{8} + \frac{1}{5} \cdot \frac{3}{10} + \frac{1}{2} \cdot \frac{3}{20} \\ &= \frac{1}{20} + \frac{3}{40} + \frac{3}{50} + \frac{3}{40} \\ &= \frac{52}{200} = \frac{13}{50}. \end{aligned}$$

Teorema 2.4.8. Nas mesmas condições da Lei da Probabilidade Total, a **Regra de Bayes** permite inverter a condição de um evento. Para $A, B_i \in \mathcal{F}$, com $P(B_i) > 0$ e $P(A) > 0$, vale:

$$P(B_i | A) = \frac{P(A | B_i) P(B_i)}{\sum_{j=1}^n P(A | B_j) P(B_j)}.$$

Demonstração: Pela definição de probabilidade condicional,

$$P(B_i | A) = \frac{P(B_i \cap A)}{P(A)}, \quad P(A | B_i) = \frac{P(A \cap B_i)}{P(B_i)}.$$

Da segunda igualdade obtemos $P(A \cap B_i) = P(A | B_i) P(B_i)$. Substituindo na fórmula de $P(B_i | A)$ segue

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

Aplicando a Lei da Probabilidade Total ao denominador $P(A)$ (usando a partição $\{B_j\}_{j=1}^n$) obtemos

$$P(A) = \sum_{j=1}^n P(A \cap B_j) = \sum_{j=1}^n P(A | B_j) P(B_j).$$

Substituindo essa expressão em $P(B_i | A)$ resulta exatamente na fórmula reivindicada:

$$P(B_i | A) = \frac{P(A | B_i) P(B_i)}{\sum_{j=1}^n P(A | B_j) P(B_j)}.$$

Isto conclui a prova. □

Exemplo 2.4.9. Num pântano existem três bruxas que preparam poções por encomenda. Quando uma encomenda é feita, a probabilidade é de $\frac{1}{2}$ de que ela seja preparada pela bruxa B_1 , de $\frac{3}{10}$ pela bruxa B_2 e de $\frac{1}{5}$ pela bruxa B_3 .

Um barão com problemas conjugais encomendou uma poção com o objetivo de reacender a paixão entre ele e sua esposa; porém, há o risco indesejado de que a poção faça com que o casal nunca mais possa ficar junto. Seja A o evento “o casal é separado”, e sejam as probabilidades condicionais desse efeito

$$P(A | B_1) = \frac{3}{20} \quad P(A | B_2) = \frac{3}{10} \quad P(A | B_3) = \frac{2}{5}.$$

Após a confecção e utilização da poção, o casal de fato foi separado. Sabendo que o evento A ocorreu, o Barão prendeu a bruxa cuja probabilidade de ter feito a poção defeituosa era maior. Qual bruxa foi presa pelo barão?

Solução. Pela Lei da Probabilidade Total, temos:

$$\begin{aligned} P(A) &= P(B_1)P(A | B_1) + P(B_2)P(A | B_2) + P(B_3)P(A | B_3) \\ &= \frac{1}{2} \cdot \frac{3}{20} + \frac{3}{10} \cdot \frac{3}{10} + \frac{1}{5} \cdot \frac{2}{5} \\ &= \frac{3}{40} + \frac{9}{100} + \frac{2}{25} = \frac{15}{200} + \frac{18}{200} + \frac{16}{200} = \frac{49}{200}. \end{aligned}$$

Como, $P(A) = \sum_{i=1}^3 P(A | B_i) P(B_i)$, aplicando a Regra de Bayes para $i = 1, 2, 3$, temos:

$$P(B_1 | A) = \frac{P(A | B_1)P(B_1)}{P(A)} = \frac{\frac{3}{20} \cdot \frac{1}{2}}{\frac{49}{200}} = \frac{3/40}{49/200} = \frac{3 \cdot 200}{40 \cdot 49} = \frac{15}{49},$$

$$P(B_2 | A) = \frac{P(A | B_2)P(B_2)}{P(A)} = \frac{\frac{3}{10} \cdot \frac{3}{10}}{\frac{49}{200}} = \frac{9/100}{49/200} = \frac{9 \cdot 200}{100 \cdot 49} = \frac{18}{49},$$

$$P(B_3 | A) = \frac{P(A | B_3)P(B_3)}{P(A)} = \frac{\frac{2}{5} \cdot \frac{1}{5}}{\frac{49}{200}} = \frac{2/25}{49/200} = \frac{2 \cdot 200}{25 \cdot 49} = \frac{16}{49}.$$

Tais resultados são interpretados da seguinte forma: se a encomenda tivesse sido repetida de maneira independente um número muito grande de vezes, então a bruxa B_2 seria responsável por aproximadamente $\frac{18}{49}$ das poções defeituosas.

Assim, as probabilidades B_i podem ser chamadas de probabilidades “anteriores” (ou “a priori”), enquanto as $P(B_i | A)$ podem ser chamadas de probabilidades “posteriores” (ou “a posteriori”). Como $P(B_2 | A)$ é a maior das três probabilidades posteriores, o Barão prendeu a bruxa B_2 .

Uma análise superficial desse problema indicaria erroneamente que B_1 seria a bruxa presa pois, embora B_1 tivesse a maior probabilidade a priori, a combinação da probabilidade a priori de B_2 com seu risco condicional faz com que B_2 seja a candidata mais provável após observar A .

Definição 2.4.10. Seja (Ω, \mathcal{F}, P) um espaço de probabilidade. Dois eventos $A, B \in \mathcal{F}$ são ditos independentes se a ocorrência de um não influencia a probabilidade de ocorrência do outro, isto é, se vale:

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

De forma equivalente, sempre que $P(B) > 0$, a condição acima implica:

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

e, simetricamente, se $P(A) > 0$,

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

Definição 2.4.11. Os eventos A_1, \dots, A_n em (Ω, \mathcal{F}, P) são independentes se, para toda coleção de índices $1 \leq i_1 \leq i_2 \leq \dots \leq i_k \leq n$ e $2 \leq k \leq n$, tivermos

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

Exemplo 2.4.12. Considere o espaço de probabilidade obtido ao lançar, uma de cada vez, duas moedas justas e independentes. Seja C a representação da face “cara” para cima, e K a representação de “coroa” para cima. Daí, temos

$$\Omega = \{CC, CK, KC, KK\},$$

com cada resultado elementar equiprovável.

Definimos os eventos

$$A_1 = \{\text{primeira moeda mostra } C\} = \{CC, CK\},$$

$$A_2 = \{\text{segunda moeda mostra } C\} = \{CC, KC\},$$

$$A_3 = \{\text{as duas moedas dão o mesmo resultado}\} = \{CC, KK\}.$$

Observe que:

$$P(A_1) = P(A_2) = P(A_3) = \frac{2}{4} = \frac{1}{2}.$$

Agora, calculemos as interseções de pares:

$$A_1 \cap A_2 = \{CC\}, \quad P(A_1 \cap A_2) = \frac{1}{4} = P(A_1)P(A_2),$$

$$A_1 \cap A_3 = \{CC\}, \quad P(A_1 \cap A_3) = \frac{1}{4} = P(A_1)P(A_3),$$

$$A_2 \cap A_3 = \{CC\}, \quad P(A_2 \cap A_3) = \frac{1}{4} = P(A_2)P(A_3).$$

Portanto os três eventos são par-a-par independentes.

Entretanto, a interseção dos três eventos é

$$A_1 \cap A_2 \cap A_3 = \{CC\}, \quad P(A_1 \cap A_2 \cap A_3) = \frac{1}{4}.$$

Se os eventos fossem independentes em conjunto deveríamos ter

$$P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8},$$

o que não ocorre ($\frac{1}{4} \neq \frac{1}{8}$). Assim, embora sejam independentes par-a-par, as variáveis não são independentes em conjunto.

Lema 2.4.13. Seja $(A_n)_{n \in \mathbb{N}}$ uma sequência de eventos em $(\Omega, \mathcal{F}, \mathcal{P})$ com $p_n = P(A_n)$ para todo $n \geq 1$. Temos:

1. Se $\sum_{n=1}^{\infty} p_n < \infty$, então $P(\limsup A_n) = 0$. (Primeiro Lema de Borel–Cantelli)

2. Se os eventos (A_n) são independentes e $\sum_{n=1}^{\infty} p_n = \infty$, então $P(\limsup A_n) = 1$.
(Segundo Lema de Borel–Cantelli)

Demonstração: Recorde que

$$\limsup_{n \rightarrow \infty} A_n = \bigcap_{k=1}^{\infty} \bigcup_{n \geq k} A_n, \quad (\limsup_n A_n)^c = \bigcup_{k=1}^{\infty} \bigcap_{n \geq k} A_n^c.$$

(i) Se $\sum p_n < \infty$. Defina $B_k = \bigcup_{n \geq k} A_n$. Então (B_k) é decrescente e

$$\limsup A_n = \bigcap_k B_k.$$

Pela continuidade de cima,

$$P(\limsup A_n) = \lim_{k \rightarrow \infty} P(B_k).$$

Pela σ -subaditividade,

$$P(B_k) \leq \sum_{n \geq k} p_n.$$

Como $\sum p_n$ converge, temos

$$\lim_{k \rightarrow \infty} \sum_{n=k}^{\infty} p_n = 0$$

logo $P(\limsup A_n) = 0$.

(ii) Se $\sum p_n = \infty$ e (A_n) independentes. Defina

$$C_k = \bigcap_{n \geq k} A_n^c.$$

Então (C_k) é crescente e

$$(\limsup A_n)^c = \bigcup_k C_k.$$

Para $m \geq k$, defina

$$C_{k,m} = \bigcap_{n=k}^m A_n^c.$$

Pela independência,

$$P(C_{k,m}) = \prod_{n=k}^m (1 - p_n).$$

Passando ao limite,

$$P(C_k) = \lim_{m \rightarrow \infty} \prod_{n=k}^m (1 - p_n).$$

Como $1 - x \leq e^{-x}$,

$$P(C_{k,m}) \leq e^{-\sum_{n=k}^m p_n}.$$

Logo,

$$P(C_k) \leq \exp\left(-\sum_{n=k}^{\infty} p_n\right).$$

Como $\sum_{n=1}^{\infty} p_n$ converge, segue que $\lim_{k \rightarrow \infty} \sum_{n=k}^{\infty} p_n = 0$.

portanto

$$\lim_{k \rightarrow \infty} P(C_k) = 0.$$

Pela continuidade de baixo,

$$P((\limsup A_n)^c) = \lim_{k \rightarrow \infty} P(C_k) = 0.$$

Logo $P(\limsup A_n) = 1$.

□

Observação 2.4.14. *A independência é essencial no item (ii): sem ela a conclusão pode falhar (por exemplo, se todos os A_n coincidirem com um mesmo evento A de probabilidade $0 < p < 1$, então $\sum_n p_n = +\infty$ mas $\limsup_n A_n = A$ tem probabilidade $p \neq 1$).*

Exemplo 2.4.15. *Considere, a cada dia $n \geq 1$, que um carteiro embaralha aleatoriamente um baralho padrão de 52 cartas e coloca o baralho na sua mesa. Assuma que, em dias distintos, os embaralhamentos são independentes entre si e que cada permutação do baralho é igualmente provável.*

Fixe uma carta específica, por exemplo o Ás de Espadas, e defina o evento

$$A_n = \{\text{no dia } n \text{ o Ás de Espadas está na primeira posição (topo) do baralho}\}.$$

Como cada permutação é equiprovável, temos a probabilidade constante

$$p_n = \mathbb{P}(A_n) = \frac{1}{52} \quad \text{para todo } n.$$

Aplicando o Lema de Borel–Cantelli (parte (ii)), temos que

$$\sum_{n=1}^{\infty} p_n = \sum_{n=1}^{\infty} \frac{1}{52} = +\infty,$$

e, pelo fato de os eventos A_n serem mutuamente independentes (cada embaralhamento independe dos demais), segue que

$$\mathbb{P}\left(\limsup_{n \rightarrow \infty} A_n\right) = 1.$$

Em palavras: com probabilidade 1 o Ás de Espadas aparecerá na posição de topo infinitas vezes ao longo dos infinitos dias.

Ao longo desta seção, foram apresentados os fundamentos da probabilidade em seu contexto clássico, com ênfase em espaços amostrais finitos ou enumeráveis, nos quais a atribuição de probabilidades se dá por meio da contagem e da razão entre casos favoráveis e possíveis. Ainda que essa abordagem seja suficiente para uma ampla classe de problemas, ela se mostra limitada quando o espaço amostral deixa de ser discreto. A partir deste ponto do trabalho em questão, interrompe-se a análise de modelos baseados exclusivamente na contagem e passa-se à investigação de situações em que o espaço de resultados é contínuo, exigindo novas ferramentas conceituais. Nesse contexto, o método de Monte Carlo surge como ponte intuitiva entre a interpretação frequencial (que foge do escopo do trabalho, por isso será citada brevemente a seguir) e os modelos contínuos, preparando o terreno para a formalização via teoria da medida desse último modelo.

2.5 O Método de Monte Carlo

Durante a Segunda Guerra Mundial, físicos e matemáticos buscavam compreender o comportamento dos nêutrons dentro dos reatores nucleares. Realizar experimentos reais era inviável, tanto pelo alto custo quanto pelo risco envolvido, o que motivou a criação de modelos matemáticos de simulação estocástica que reproduzissem de forma aproximada o fenômeno observado. Essas simulações, baseadas em procedimentos aleatórios e repetitivos, receberam o nome de Método de Monte Carlo, em referência à cidade europeia famosa por seus cassinos, devido à analogia com jogos de azar (*gambling games*). O termo passou, então, a designar uma classe de métodos numéricos e probabilísticos que utilizam experimentos aleatórios (reais ou simulados) para estimar valores ou resolver problemas complexos, sobretudo aqueles que envolvem integração, otimização ou distribuição de probabilidades em espaços contínuos. Em linhas gerais, o método consiste em realizar um grande número de repetições de um experimento aleatório e utilizar a frequência relativa observada para aproximar probabilidades teóricas ou valores esperados. Essa abordagem é particularmente útil quando a análise matemática direta é impraticável, e as simulações computacionais se tornam uma alternativa viável (JOHNSON, 1995). Embora o foco deste trabalho não esteja na aplicação computacional ou estatística do método de Monte Carlo, é relevante mencioná-lo, pois o termo surge ocasionalmente em contextos de probabilidade geométrica e em discussões sobre estimativas probabilísticas obtidas por simulação. Assim, a presente menção visa apenas situar o leitor quanto ao significado e à origem desse conceito, sem adentrar nas técnicas específicas de sua implementação. Caso haja interesse se aprofundar em probabilidade geométrica por meio de simulações, consulte Araújo (2017).

3 Probabilidade Geométrica e Teoria da Medida

3.1 Probabilidade Geométrica e suas Modelagens Intuitivas

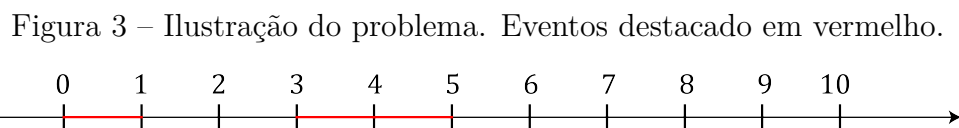
Primeiramente, vamos analisar três experimentos aleatórios e, utilizando um pouco de senso comum e lógica, calcular as probabilidades de cada evento. Será apresentada a definição intuitiva do que seria a probabilidade geométrica em \mathbb{R} , \mathbb{R}^2 e \mathbb{R}^3 .

Definição 3.1.1. Quando o espaço amostral Ω é um intervalo $[a, b] \in \mathbb{R}$, com $a \leq b$, a probabilidade geométrica mede a razão entre o comprimento da parte favorável (evento) e o comprimento total. Em outras palavras, se um ponto é escolhido ao acaso em um intervalo $\Omega = [a, b]$, e o evento A corresponde a um subintervalo de $[a, b]$, então:

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

Exemplo 3.1.2. Escolhe-se, ao acaso, um ponto no intervalo $[0, 10]$. Qual a probabilidade desse ponto pertencer a $A = [0, 1]$? E a $B = [3, 5]$? E a $A \cup B$?

Solução: Seja $\Omega = [0, 10]$ o espaço amostral, com comprimento $|\Omega| = 10$. Na figura 3, está representado o espaço amostral e os dois eventos (destacados em vermelho). Primeiramente,



Fonte: Elaborada pelo autor (2026).

calcularemos a probabilidade de pertencer a $A = [0, 1]$. Como o comprimento desse intervalo é $|A| = 1$, teremos $P(A) = \frac{|A|}{|\Omega|} = \frac{1}{10} = 0,1$.

A probabilidade de pertencer ao intervalo $B = [3, 5]$ será: Comprimento do intervalo $|B| = 2$, logo $P(B) = \frac{|B|}{|\Omega|} = \frac{2}{10} = 0,2$.

Por fim, calcularemos a probabilidade de pertencer a $A \cup B = [0, 1] \cup [3, 5]$. Como $A \cap B = \emptyset$, temos pela propriedade 2.3.3 que

$$P(A \cup B) = P(A) + P(B) = \frac{1}{10} + \frac{2}{10} = \frac{3}{10} = 0,3.$$

Portanto,

$$P(A) = 0,1 \quad P(B) = 0,2 \quad P(A \cup B) = 0,3$$

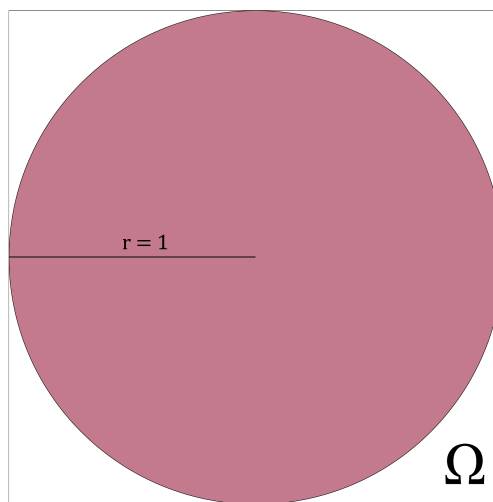
Definição 3.1.3. No plano, a probabilidade geométrica é a razão entre a área da região favorável e a área da região total. Se um ponto é escolhido ao acaso em uma região Ω do plano, e $A \subseteq \Omega$, então:

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

Exemplo 3.1.4. Considere Ω como um quadrado de lado 2, e A o círculo de raio 1 inscrito nesse quadrado. Qual a probabilidade de se escolher um ponto aleatoriamente no quadrado de modo que ele também pertença ao círculo?

Solução:

Figura 4 – Quadrado com um círculo inscrito, com coloração destacando o evento.



Fonte: Elaborada pelo autor (2026).

Seja Ω o quadrado de lado 2 e A o círculo de raio 1 inscrito em Ω , como mostra a figura 4. A probabilidade geométrica é dada por

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

A área do quadrado é

$$|\Omega| = 2^2 = 4.$$

A área do círculo inscrito é

$$|A| = \pi \cdot 1^2 = \pi.$$

Portanto,

$$P(A) = \frac{|A|}{|\Omega|} = \frac{\pi}{4}.$$

$$P(A) = \frac{\pi}{4}.$$

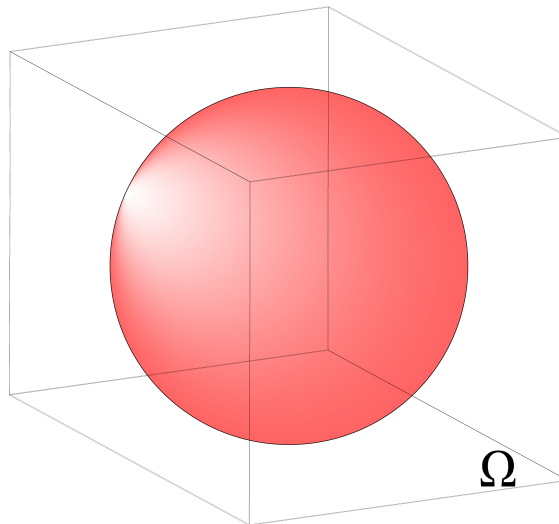
Definição 3.1.5. No espaço tridimensional, a probabilidade geométrica é a razão entre o volume da parte favorável e o volume total. Ou seja, se um ponto é escolhido ao acaso em uma região espacial Ω e $A \subseteq \Omega$ é uma sub-região, então:

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

Exemplo 3.1.6. Considere Ω como um cubo de aresta 1, e A a esfera inscrita nesse cubo (raio $r = \frac{1}{2}$). Ao se escolher um ponto aleatoriamente pertencente ao cubo, qual a probabilidade dele pertencer à esfera?

Solução: Seja Ω o cubo de aresta 1, e A a esfera inscrita de raio $r = \frac{1}{2}$, como ilustra a figura 5. A probabilidade geométrica é dada por $P(A) = \frac{\text{volume}(A)}{\text{volume}(\Omega)} = \frac{|A|}{|\Omega|}$.

Figura 5 – Esfera inscrita em um cubo, destacando-se o evento.



Fonte: Elaborada pelo autor (2026).

O volume do cubo é

$$|\Omega| = 1^3 = 1.$$

O volume da esfera é

$$|A| = \frac{4}{3}\pi r^3 = \frac{4}{3}\pi \left(\frac{1}{2}\right)^3 = \frac{4}{3}\pi \cdot \frac{1}{8} = \frac{\pi}{6}.$$

Portanto,

$$P(A) = \frac{|A|}{|\Omega|} = \frac{\frac{\pi}{6}}{1} = \frac{\pi}{6}.$$

$$P(A) = \frac{\pi}{6}.$$

Em James (2010), o autor explica, de forma sucinta, que não é possível definir $P(A)$ para todo subconjunto $A \subset \Omega$. Em outras palavras, existem conjuntos cujo comprimento

(no sentido de Lebesgue), área ou volume não são bem definidos. A demonstração dessa impossibilidade está relacionada ao *Axioma da Escolha*, razão pela qual o autor estabelece que as probabilidades serão calculadas apenas para comprimentos, áreas e volumes que possuam medida bem definida.

3.2 Medida - Espaço mensurável

Ao falar sobre probabilidade, é comum pensar em espaços amostrais discretos, como o lançamento de um dado ou de uma moeda, nos quais basta listar todos os resultados possíveis e contar quantos satisfazem um determinado evento. Contudo, quando passamos para situações contínuas, como escolher um ponto ao acaso em um segmento, em um quadrado ou em um cubo, surge uma questão mais sutil: como garantir que todos os eventos possíveis tenham uma “área”, “comprimento” ou “volume” bem definidos?

Essa questão é o ponto de partida para o uso de σ -álgebras e, em especial, da σ -álgebra de Borel. Em espaços contínuos, nem todo subconjunto de \mathbb{R} , \mathbb{R}^2 ou \mathbb{R}^3 tem uma “medida” coerente. Existem conjuntos problemáticos, como o conjunto de Vitali 3.2.3, que não admitem uma definição consistente de medida. Assim, é necessário restringir os conjuntos àqueles que são “bem comportados”, isto é, que podem ser medidos de maneira compatível com as propriedades desejadas que veremos a seguir. Esses conjuntos “bem comportados” são, justamente, os elementos da σ -álgebra de Borel, que posteriormente será estritamente ampliada na construção da medida de Lebesgue.

Antes de iniciar esta seção, foram apresentadas algumas definições que serão importantes ao longo do texto. Para preservar a fluidez da leitura, as demonstrações excessivamente longas foram deslocadas para o Anexo A e, ao longo do texto principal, serão indicadas por um asterisco (*). Dito isso, segue o desenvolvimento da seção.

Definição 3.2.1. *O conjunto dos **números reais estendidos** é definido por*

$$\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\} = [-\infty, +\infty].$$

A ordem usual de \mathbb{R} é estendida a $\overline{\mathbb{R}}$ impondo-se que, para todo $x \in \mathbb{R}$,

$$-\infty < x < +\infty.$$

Para os nossos objetivos, nesse trabalho, consideraremos principalmente o subconjunto de $\overline{\mathbb{R}}$ dado por $[0, +\infty]$, pois estaremos trabalhando com tamanhos não negativos. Enfatizamos que os símbolos $-\infty$ e $+\infty$ não são números reais, trata-se de um símbolo formal utilizado para descrever comportamento de não limitação, como em limites, ou medidas infinitas.

Um *espaço mensurável* é um par (Ω, \mathcal{F}) , onde Ω é um conjunto não-vazio e \mathcal{F} é uma σ -álgebra em Ω .

Definição 3.2.2. (*Medidas e espaços de medida*). Seja (Ω, \mathcal{F}) um dado espaço mensurável. Uma função $\mu : \mathcal{F} \rightarrow [0, +\infty]$ é chamada medida em (Ω, \mathcal{F}) se

1. $\mu(\emptyset) = 0$,
2. $\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$ para toda sequência $(A_n)_n$ de conjuntos em \mathcal{F} disjuntos. Esta propriedade chama-se σ -aditividade.

O trio $(\Omega, \mathcal{F}, \mu)$ é chamado *espaço de medida*. Dizemos que μ é uma *medida finita* se $\mu(\Omega) < \infty$. Dizemos que μ é uma *medida σ -finita* se existem conjuntos $(A_n)_{n \in \mathbb{N}}$ em \mathcal{F} tais que $\bigcup_{n \in \mathbb{N}} A_n = \Omega$ e $\mu(A_n) < \infty$ para todo n .

Quando $\mu(\Omega) = 1$, a medida μ é uma *medida de probabilidade*, como definida na seção anterior.

Tabela 1 – Tabela-resumo sobre os conceitos que podem confundir em um primeiro momento.

Objeto	Definição	Papel conceitual
Espaço mensurável	Par (Ω, \mathcal{F})	Diz <i>quais</i> $A \in \Omega$ podem receber medida.
Medida	Função $\mu : \mathcal{F} \rightarrow [0, +\infty]$	Atribui um <i>tamanho</i> aos conjuntos mensuráveis.
Espaço de medida	Tríplice $(\Omega, \mathcal{F}, \mu)$	Estrutura completa para medir conjuntos.

Dado um intervalo $I \subset \mathbb{R}$ definimos $\ell(I)$ como o seu comprimento, isto é, $\ell(I) = +\infty$ se I for ilimitado e $\ell(I) = b - a$ se $a, b \in \mathbb{R}$, a for seu extremo inferior, b seu extremo superior e $a \leq b$.

De acordo com [Rolla e Lima \(2025\)](#), o problema de medida na reta é encontrar uma função m a valores reais estendidos não negativos, isto é, tomando valores em $[0, +\infty]$ definida nos subconjuntos de \mathbb{R} que satisfaça as seguintes propriedades:

1. se $E_n \subset \mathbb{R}$, $n \in \mathbb{N}$, é uma sequência de subconjuntos dois a dois disjuntos então

$$m\left(\bigcup_{n \geq 1} E_n\right) = \sum_{n \geq 1} m(E_n)$$

(σ -aditividade);

2. se I é um intervalo então $m(I) = \ell(I)$;
3. se E é um conjunto e $x \in \mathbb{R}$ então $m(E + x) = m(E)$, sendo

$$E + x = \{y + x; y \in E\}.$$

Os conjuntos para os quais essa construção é possível são chamados conjuntos mensuráveis à Lebesgue, e a função assim obtida recebe o nome de medida de Lebesgue¹.

Um detalhe curioso é que, de acordo com Royden e Fitzpatrick (2010), não há sequer uma função que satisfaça as condições 2 e 3 e que seja finitamente aditiva para todos os subconjuntos dos reais. Por isso a motivação por construir uma função sobre um conjunto específico que possua tais propriedades.

Segundo Fernandez (2002), esse problema é conhecido por alguns autores como “problema difícil da medida” e ele não tem solução, isto é, não é possível estabelecer uma função que consiga “medir” todos os subconjuntos dos reais, seguindo as propriedades naturais desejáveis citadas anteriormente. Esse fato é evidenciado pela construção do conjunto de Vitali, apresentada no Teorema 3.2.3.

3.2.1 O Teorema de Vitali e a impossibilidade de medir todos os conjuntos

Uma questão natural ao tentar generalizar a noção de comprimento é saber se é possível atribuir medida a *todos* os subconjuntos da reta real, preservando propriedades geométricas desejáveis, como a invariância por translações e a aditividade enumerável. O Teorema de Vitali demonstra que nem todo subconjunto de \mathbb{R} admite medida compatível com invariância por translação e aditividade, reforçando a necessidade de restringir a probabilidade geométrica a conjuntos mensuráveis, que serão definidos em breve. Dito isso, segue:

Teorema 3.2.3 (Vitali*). *Existe um subconjunto $V \subset [0, 1]$ tal que não existe nenhuma função $m : \mathcal{P}(\mathbb{R}) \rightarrow [0, +\infty]$ satisfazendo simultaneamente:*

1. $m(\emptyset) = 0$;
2. m é σ -aditiva;
3. $m(I) = \ell(I)$ para todo intervalo $I \subset \mathbb{R}$;
4. $m(E + x) = m(E)$ para todo $E \subset \mathbb{R}$ e todo $x \in \mathbb{R}$.

Em particular, o conjunto V não é mensurável no sentido de Lebesgue.

Observação 3.2.4. *O Teorema de Vitali mostra que não é possível definir uma medida em todos os subconjuntos de \mathbb{R} que seja simultaneamente compatível com a noção geométrica de comprimento e invariável por translações. Esse resultado justifica a necessidade de restringir o domínio da medida a uma σ -álgebra apropriada. É exatamente por isso que não podemos considerar $\mathcal{P}(\Omega)$ como a σ -álgebra (mesmo ela sendo uma σ -álgebra*

¹ Henri Lebesgue (1875–1941) foi um matemático francês que introduziu a medida e a integral que levam seu nome, fundamentais para a análise matemática moderna.

propriamente dita e a maior possível, inclusive), pois ela contém conjuntos aos quais não é possível atribuir uma medida com tais propriedades.

3.2.2 Medida Exterior

Antes de introduzir formalmente o conceito de medida exterior, é importante esclarecer sua motivação. Em muitos contextos, especialmente em \mathbb{R}^n , deseja-se atribuir um tamanho a subconjuntos arbitrários, mesmo quando não se sabe à priori se eles pertencem a uma σ -álgebra adequada. A medida exterior surge como um primeiro passo nesse processo: trata-se de uma função definida em todos os subconjuntos do espaço, que fornece uma estimativa do tamanho de um conjunto por meio de aproximações externas (por isso o nome de “medida exterior” ou “*outer measure*”, em inglês). Somente em um segundo momento será possível selecionar, a partir desse objeto, os conjuntos para os quais essa estimativa é compatível com as propriedades de uma medida propriamente dita.

Definição 3.2.5. *Seja Ω um conjunto. Uma função*

$$\mu^* : \mathcal{P}(\Omega) \longrightarrow [0, +\infty]$$

*é chamada de **medida exterior** em Ω se satisfaz as seguintes propriedades:*

1. $\mu^*(\emptyset) = 0$;
2. (**Monotonicidade**) Se $A \subset B \subset \Omega$, então

$$\mu^*(A) \leq \mu^*(B);$$

3. (**Subaditividade enumerável**) Para toda família enumerável $(A_n)_{n \in \mathbb{N}} \subset \mathcal{P}(\Omega)$,

$$\mu^* \left(\bigcup_{n=1}^{\infty} A_n \right) \leq \sum_{n=1}^{\infty} \mu^*(A_n).$$

Diferentemente de uma medida, a medida exterior é definida sobre todo subconjunto de Ω , não estando, em geral, associada a uma σ -álgebra, nem satisfazendo a propriedade de aditividade. Seu papel fundamental consiste em servir como ponto de partida para a seleção dos conjuntos que podem ser considerados mensuráveis, por meio do critério de mensurabilidade de Carathéodory.

A introdução da medida exterior permite, assim, dissociar dois problemas conceitualmente distintos: o de atribuir um “tamanho” a subconjuntos de um espaço e o de determinar quais desses subconjuntos devem ser admitidos como mensuráveis. A partir de uma medida exterior, define-se uma classe especial de subconjuntos, chamados conjuntos mensuráveis no sentido de Carathéodory, para os quais a restrição da medida exterior passa a satisfazer as propriedades de uma medida propriamente dita.

No contexto da probabilidade geométrica, a medida exterior é construída a partir de coberturas por intervalos, no caso da reta real, ou por retângulos, em \mathbb{R}^n . Essa construção conduz, via o Teorema de Carathéodory, à definição da medida de Lebesgue, que fornece uma noção rigorosa de comprimento, área e volume compatível com as propriedades naturais desejadas. Observe que a definição dada não é, ainda, medida exterior de Lebesgue, mas sim uma definição geral.

3.2.3 Critério de mensurabilidade

De acordo com [Royden e Fitzpatrick \(2010\)](#), existem várias maneiras de definir o que significa mensurabilidade de um conjunto. Para os nossos objetivos, seguimos uma abordagem devida a Constantine Caratheodory², conhecida como critério de mensurabilidade de Carathéodory.

Definição 3.2.6. *Seja μ^* uma medida exterior em um conjunto Ω . Um subconjunto $E \subset \Omega$ é dito **mensurável** se, para todo conjunto $A \subset \Omega$, vale*

$$\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E^c).$$

A condição acima expressa que o conjunto E é compatível com o processo de medição induzido pela medida exterior. Mais precisamente, para todo conjunto $A \subset \Omega$, a medida exterior de A pode ser decomposta exatamente nas partes contidas em E e em seu complementar, sem perda ou excesso de medida. Dessa forma, conjuntos mensuráveis no sentido de Carathéodory são aqueles que permitem decomposições aditivas consistentes da medida exterior.

Observação 3.2.7. *Algumas literaturas, como ([ROLLA; LIMA, 2025](#)), dizem que a condição supracitada é chamada de μ^* -mensurável. Por padronização, continuaremos chamando de “mensurável à Carathéodory”.*

Teorema 3.2.8 (Carathéodory*). *Seja μ^* uma medida exterior em Ω e denote por \mathcal{M} a coleção de todos os subconjuntos de Ω que são mensuráveis no sentido de Carathéodory. Então:*

1. \mathcal{M} é uma σ -álgebra;
2. a restrição $\mu = \mu^*|_{\mathcal{M}}$ é uma medida em (Ω, \mathcal{M}) .

O Teorema de Carathéodory fornece o mecanismo central para a construção de medidas. A ideia fundamental consiste em partir de uma medida exterior definida sobre

² Constantin Carathéodory (1873–1950) foi um matemático que formulou o critério de mensurabilidade que fundamenta a construção rigorosa da medida de Lebesgue.

a coleção de todas as partes de um conjunto Ω , selecionar, por meio de um critério de mensurabilidade, uma σ -álgebra natural de subconjuntos e, em seguida, restringir a medida exterior a essa classe. Essa restrição satisfaz as propriedades exigidas de uma medida propriamente dita, resultando, assim, em um espaço de medida bem definido.

No caso particular em que $\Omega = \mathbb{R}^n$, a medida exterior é construída a partir de coberturas por retângulos. A aplicação do critério de mensurabilidade de Carathéodory conduz, então, à σ -álgebra dos conjuntos mensuráveis de Lebesgue e à correspondente medida de Lebesgue, que fornece uma noção rigorosa de comprimento, área e volume. Essa construção constitui a base teórica da probabilidade geométrica.

Observação 3.2.9. *O critério de Carathéodory não impõe previamente uma σ -álgebra; ao contrário, a σ -álgebra surge naturalmente como consequência da compatibilidade dos conjuntos com a medida exterior. Esse fato reforça o caráter estrutural, e não arbitrário, da escolha do domínio da medida. Como mostrado anteriormente, a propriedade de σ -aditividade aplica-se apenas a uniões enumeráveis. O exemplo a seguir destaca as consequências dessa restrição:*

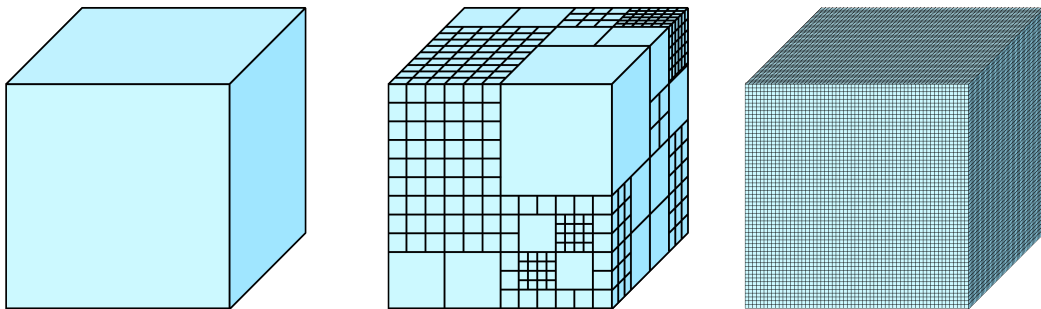
Exemplo 3.2.10. *Um cubo de volume $V \neq 0$ pode ser particionado em um número finito ou enumerável de cubos menores e mensuráveis, de modo que o volume total seja dado pela soma dos volumes das partes.*

No entanto, o cubo também pode ser escrito como a união de todos os seus pontos, cada um deles de volume zero, o que levaria à conclusão equivocada de que $V = 0$.

Ambas as situações são demonstradas na figura 6.

Por que não é legítimo tentar calcular o volume do cubo como a soma dos volumes de seus pontos?

Figura 6 – Cubo sendo decomposto em vários outros cubos e depois em infinitos (não enumeráveis) pontos.



Fonte: Elaborada pelo autor (2026).

Solução: O problema do raciocínio está na aplicação indevida da propriedade de aditividade da medida. Quando um cubo é particionado em um número finito ou enumerável de cubos menores mensuráveis, a medida satisfaz a propriedade de σ -aditividade, isto é,

$$\mu\left(\bigcup_n V_n\right) = \sum_n \mu(V_n),$$

o que permite recuperar corretamente o volume total como soma (ou limite de somas) dos volumes das partes.

Por outro lado, ao escrever o cubo como a união de todos os seus pontos,

$$C = \bigcup_{x \in C} \{x\},$$

temos uma decomposição em conjuntos de medida nula, mas cuja união é não enumerável. A propriedade de σ -aditividade não se aplica a uniões não enumeráveis, de modo que a expressão

$$\mu(C) = \sum_{x \in C} \mu(\{x\})$$

não é válida, pois a propriedade de σ -aditividade aplica-se apenas a uniões enumeráveis, e não a uniões não enumeráveis.

Assim, não há contradição: o volume de um conjunto não é obtido como soma dos volumes de seus pontos, mas como um limite de volumes de decomposições finitas ou enumeráveis por conjuntos mensuráveis.

3.2.4 Teorema da Extensão de Carathéodory

Em muitas situações, uma função de medida é inicialmente definida apenas sobre uma classe restrita de conjuntos simples, como intervalos ou retângulos. O Teorema da Extensão de Carathéodory garante que, sob hipóteses adequadas, essa definição pode ser estendida de maneira única a uma medida definida em uma σ -álgebra maior.

Teorema 3.2.11 (Extensão de Carathéodory*). *Seja \mathcal{A} uma álgebra de subconjuntos de um conjunto Ω e seja $\mu_0 : \mathcal{A} \rightarrow [0, +\infty]$ uma pré-medida, isto é, uma função tal que $\mu_0(\emptyset) = 0$ e que é σ -aditiva, no seguinte sentido: se $(A_k)_{k \geq 1} \subset \mathcal{A}$ é uma família de conjuntos dois a dois disjuntos tal que*

$$\bigcup_{k=1}^{\infty} A_k \in \mathcal{A},$$

então

$$\mu_0\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \mu_0(A_k).$$

Então existe uma medida exterior μ^* em Ω tal que:

1. todo conjunto de \mathcal{A} é mensurável no sentido de Carathéodory;
2. a restrição de μ^* à σ -álgebra \mathcal{M} dos conjuntos mensuráveis é uma medida μ que estende μ_0 .

Além disso, se μ_0 é σ -finita, essa extensão é única.

Observação 3.2.12. No caso de \mathbb{R}^n , toma-se como álgebra inicial a coleção dos retângulos, com a pré-medida definida pelo produto dos comprimentos dos lados. O Teorema da Extensão de Carathéodory garante que essa pré-medida se estende de forma consistente à medida de Lebesgue, definida sobre a σ -álgebra dos conjuntos mensuráveis.

Os resultados apresentados nesta seção completam o esquema abstrato de construção de medidas: parte-se de uma pré-medida simples, constrói-se uma medida exterior, seleciona-se uma σ -álgebra natural de conjuntos mensuráveis e obtém-se, por restrição, uma medida propriamente dita. Esse esquema será aplicado diretamente na definição da probabilidade geométrica e nas demonstrações das proposições da próxima seção.

3.3 Retângulos em \mathbb{R}^n

Os fundamentos para criar essa seção vieram de Royden e Fitzpatrick (2010), Folland (1999) e Marginean (2015). Como o objetivo é definir a probabilidade geométrica na reta, no plano e no espaço, as demonstrações serão desenvolvidas diretamente em \mathbb{R}^n . Assim, os casos particulares correspondentes a \mathbb{R} , \mathbb{R}^2 e \mathbb{R}^3 decorrem naturalmente das proposições gerais apresentadas a seguir.

Definição 3.3.1. Em \mathbb{R}^n , chama-se **retângulo** qualquer conjunto da forma

$$R = \prod_{i=1}^n [a_i, b_i), \quad \text{com } a_i < b_i.$$

No caso $n = 1$, tais conjuntos correspondem a intervalos semiabertos da reta real.

Os retângulos constituem uma classe simples de subconjuntos de \mathbb{R}^n , estável por interseções finitas e adequada para gerar a σ -álgebra de Borel. Eles servirão como blocos fundamentais para a construção da medida de Lebesgue. A escolha de retângulos semiabertos garante decomposições sem ambiguidades de fronteira e assegura a aditividade da pré-medida, evitando sobreposições artificiais nos extremos. Segue um exemplo dessa afirmação:

Exemplo 3.3.2. Em \mathbb{R}^3 , considere o retângulo semiaberto

$$R = [0, 2) \times [0, 1) \times [0, 1).$$

Esse conjunto pode ser decomposto como a união de dois retângulos semiabertos disjuntos,

$$R = [0, 1) \times [0, 1) \times [0, 1) \cup [1, 2) \times [0, 1) \times [0, 1),$$

cuja união coincide exatamente com R .

Se fossem utilizados intervalos fechados, teríamos

$$[0, 2] \times [0, 1] \times [0, 1] = [0, 1] \times [0, 1] \times [0, 1] \cup [1, 2] \times [0, 1] \times [0, 1],$$

mas os conjuntos do lado direito não seriam disjuntos, pois compartilham a face $\{1\} \times [0, 1] \times [0, 1]$, que seria contada duas vezes. Assim, a escolha de intervalos semiabertos evita a contagem dupla de fronteiras e garante a aditividade da pré-medida.

3.4 Pré-medida em retângulos

Definição 3.4.1. Define-se a **pré-medida geométrica** em retângulos de \mathbb{R}^n pela função μ_0 dada por

$$\mu_0\left(\prod_{i=1}^n [a_i, b_i)\right) = \prod_{i=1}^n (b_i - a_i).$$

Observação 3.4.2. Embora ambos os lados da igualdade envolvam o símbolo \prod , o termo $\prod_{i=1}^n [a_i, b_i)$ denota um produto cartesiano de conjuntos, ao passo que $\prod_{i=1}^n (b_i - a_i)$ denota um número real, interpretado como o volume do retângulo.

No caso $n = 1$, μ_0 coincide com o comprimento de intervalos; para $n = 2$, com a área de retângulos; e para $n = 3$, com o volume de paralelepípedos.

3.4.1 Consistência da Pré-medida Geométrica

A função μ_0 definida sobre a coleção dos retângulos de \mathbb{R}^n possui a propriedade fundamental de aditividade enumerável quando uma decomposição de um retângulo em retângulos disjuntos é considerada. A demonstração desse fato em sua forma mais geral exige um formalismo técnico mais elaborado, envolvendo propriedades estruturais da família de retângulos como um semianel de conjuntos. Uma exposição completa pode ser encontrada, por exemplo, em [Royden e Fitzpatrick \(2010, p. 424\)](#), na seção sobre medida de Lebesgue no espaço euclídeo \mathbb{R}^n .

No contexto deste trabalho, será apresentada uma versão mais acessível desse resultado, suficiente para os objetivos da construção da medida geométrica, considerando o caso em que a decomposição do retângulo ocorre ao longo de uma única coordenada, que é possível ser provada usando os fatos já conhecidos.

Proposição 3.4.3. A função μ_0 definida sobre a coleção dos retângulos de \mathbb{R}^n é uma pré-medida, isto é,

1. $\mu_0(\emptyset) = 0$;

2. Se

$$R = \prod_{i=1}^n [a_i, b_i)$$

pode ser escrito como união enumerável de retângulos dois a dois disjuntos

$$R = \bigcup_{k=1}^{\infty} R_k,$$

com

$$R_k = \prod_{i=1}^n [a_{ik}, b_{ik}),$$

e tal decomposição ocorre ao longo de uma única coordenada, então

$$\mu_0(R) = \sum_{k=1}^{\infty} \mu_0(R_k).$$

Demonstração: Primeiramente, observe que $\mu_0(\emptyset) = 0$, pois o conjunto vazio pode ser identificado com um retângulo degenerado

$$\prod_{i=1}^n [a_i, a_i),$$

para o qual

$$\mu_0(\emptyset) = \mu_0\left(\prod_{i=1}^n [a_i, a_i)\right) = \prod_{i=1}^n (a_i - a_i) = 0.$$

Agora, seja

$$R = \prod_{i=1}^n [a_i, b_i)$$

um retângulo de \mathbb{R}^n e suponha que

$$R = \bigcup_{k=1}^{\infty} R_k,$$

onde $(R_k)_{k \geq 1}$ é uma família de retângulos dois a dois disjuntos cuja união é novamente um retângulo. Para cada k , escrevemos

$$R_k = \prod_{i=1}^n [a_{ik}, b_{ik}),$$

onde i indica a coordenada e k identifica o retângulo da decomposição.

Nessas condições, a decomposição só pode ocorrer ao longo de uma única coordenada, digamos a coordenada j , de modo que

$$[a_j, b_j) = \bigcup_{k=1}^{\infty} [a_{jk}, b_{jk}),$$

com os intervalos do lado direito dois a dois disjuntos, enquanto, para todo $i \neq j$, vale

$$a_{ik} = a_i \quad e \quad b_{ik} = b_i \quad \text{para todo } k.$$

Assim,

$$\mu_0(R) = \prod_{i=1}^n (b_i - a_i) = \left(\prod_{i \neq j} (b_i - a_i) \right) \sum_{k=1}^{\infty} (b_{jk} - a_{jk})$$

Como $\prod_{i \neq j} (b_i - a_i)$ representa o produto dos comprimentos dos intervalos que permanecem fixos e

$$\sum_{k=1}^{\infty} (b_{jk} - a_{jk}) = b_j - a_j,$$

além de, para cada k ,

$$\mu_0(R_k) = \left(\prod_{i \neq j} (b_i - a_i) \right) (b_{jk} - a_{jk}),$$

segue que

$$\left(\prod_{i \neq j} (b_i - a_i) \right) \sum_{k=1}^{\infty} (b_{jk} - a_{jk}) = \sum_{k=1}^{\infty} \mu_0(R_k) = \mu_0(R).$$

Logo, μ_0 é aditiva para decomposições enumeráveis de retângulos dois a dois disjuntos quando a decomposição ocorre ao longo de uma única coordenada. Conclui-se que μ_0 é uma pré-medida sob as condições dadas.

Sobre a última igualdade, poderia ter sido escrita alternativamente como:

$$\left(\prod_{i \neq j} (b_i - a_i) \right) \sum_{k=1}^{\infty} (b_{jk} - a_{jk}) = \sum_{k=1}^{\infty} \prod_{i=1}^n (b_{ik} - a_{ik}) = \sum_{k=1}^{\infty} \mu_0(R_k) = \mu_0(R).$$

□

A diferença do teorema 3.4.3 para a sua generalização é que a decomposição não precisa ser feita em uma única coordenada. A consistência da pré-medida garante que ela pode ser estendida, via o Teorema de Carathéodory, a uma medida definida em uma σ -álgebra maior.

3.4.2 Construção da Medida de Lebesgue em \mathbb{R}^n

Definição 3.4.4. A *medida exterior de Lebesgue* λ_n^* em \mathbb{R}^n é definida, para todo conjunto $A \subset \mathbb{R}^n$, por

$$\lambda_n^*(A) = \inf \left\{ \sum_{k=1}^{\infty} \mu_0(R_k) : A \subset \bigcup_{k=1}^{\infty} R_k, R_k \text{ retângulos} \right\}.$$

A função λ_n^* é uma medida exterior. Aplicando o critério de Carathéodory, obtém-se uma σ -álgebra de conjuntos mensuráveis sobre a qual a restrição de λ_n^* é uma medida.

3.4.3 sigma-álgebra de Lebesgue e medida de Lebesgue

Definição 3.4.5. Chama-se **σ -álgebra de Lebesgue** em \mathbb{R}^n a coleção $\mathcal{L}(\mathbb{R}^n)$ dos conjuntos mensuráveis no sentido de Carathéodory em relação à medida exterior λ_n^* . A **medida de Lebesgue** λ_n é a restrição de λ_n^* a essa σ -álgebra.

A medida de Lebesgue estende a pré-medida geométrica definida em retângulos, é invariante por translações e constitui a base formal da probabilidade geométrica em \mathbb{R}^n .

3.4.4 Compatibilidade com a sigma-álgebra de Borel

Proposição 3.4.6 ($\mathcal{B}(\mathbb{R}^n) \subset \mathcal{L}(\mathbb{R}^n)^*$). A σ -álgebra de Borel $\mathcal{B}(\mathbb{R}^n)$ está contida na σ -álgebra de Lebesgue $\mathcal{L}(\mathbb{R}^n)$. Além disso, a medida de Lebesgue $\lambda = \lambda^*|_{\mathcal{L}(\mathbb{R}^n)}$ estende a pré-medida geométrica λ_0 , isto é,

$$\lambda(R) = \lambda_0(R), \quad \forall R \in \mathcal{R}.$$

Observação 3.4.7. A σ -álgebra de Lebesgue é estritamente maior que a σ -álgebra de Borel. De fato, embora todo conjunto boreliano seja mensurável à Lebesgue, existem conjuntos mensuráveis à Lebesgue que não pertencem à σ -álgebra de Borel. Um exemplo típico é dado por subconjuntos não borelianos do conjunto de Cantor (para uma explicação completa desse fato, que é demasiado longo, consulte [Royden e Fitzpatrick \(2010, p. 53\)](#)), o qual é um conjunto fechado e de medida nula. Como todo subconjunto de um conjunto de medida zero é mensurável à Lebesgue, obtêm-se assim conjuntos mensuráveis que não são borelianos. Por outro lado, a σ -álgebra de Lebesgue é estritamente menor que o conjunto das partes de \mathbb{R}^n , pois existem subconjuntos de \mathbb{R}^n que não são mensuráveis no sentido de Lebesgue, como os conjuntos de Vitali.

3.4.5 Invariância por translações da medida de Lebesgue

Teorema 3.4.8. Seja $A \in \mathcal{L}(\mathbb{R}^n)$ e seja $x \in \mathbb{R}^n$. Então

$$\lambda_n(A + x) = \lambda_n(A),$$

onde $A + x = \{a + x : a \in A\}$.

Demonstração: Para demonstrar a invariância por translações da medida de Lebesgue, procede-se da seguinte forma: prova-se inicialmente que a medida exterior associada à pré-medida geométrica é invariante por translações. Em seguida, verifica-se que a translação preserva a mensurabilidade no sentido de Carathéodory. Como consequência, obtêm-se a invariância da medida de Lebesgue, uma vez que esta é definida como a restrição da medida exterior aos conjuntos mensuráveis.

Seja \mathcal{R} a família dos retângulos semiabertos de \mathbb{R}^n , μ_0 a pré-medida geométrica em \mathcal{R} , μ^* é a medida exterior definida para todo $E \subset \mathbb{R}^n$ e que $\lambda_n = \mu^*|_{\mathcal{L}(\mathbb{R}^n)}$.

Passo 1: Para todo $E \subset \mathbb{R}^n$ e todo $x \in \mathbb{R}^n$, vale $\lambda^*(E + x) = \lambda^*(E)$.

De fato, note primeiro que, se $R = \prod_{i=1}^n [a_i, b_i] \in \mathcal{R}$ e $x = (x_1, \dots, x_n)$, então

$$R + x = \prod_{i=1}^n [a_i + x_i, b_i + x_i] \in \mathcal{R} \quad e \quad \mu_0(R + x) = \prod_{i=1}^n ((b_i + x_i) - (a_i + x_i)) = \mu_0(R).$$

Agora fixe $\varepsilon > 0$ e escolha uma cobertura $E \subset \bigcup_{k=1}^{\infty} R_k$ com $R_k \in \mathcal{R}$ tal que

$$\sum_{k=1}^{\infty} \mu_0(R_k) \leq \lambda^*(E) + \varepsilon.$$

Transladando, obtemos $E + x \subset \bigcup_{k=1}^{\infty} (R_k + x)$ e, pela invariância de μ_0 em retângulos,

$$\lambda^*(E + x) \leq \sum_{k=1}^{\infty} \mu_0(R_k + x) = \sum_{k=1}^{\infty} \mu_0(R_k) \leq \lambda^*(E) + \varepsilon.$$

Como $\varepsilon > 0$ é arbitrário, segue $\lambda^*(E + x) \leq \lambda^*(E)$. Aplicando a desigualdade anterior ao conjunto $E + x$ e ao ponto $-x$, obtemos

$$\lambda^*(E) = \lambda^*((E + x) + (-x)) \leq \lambda^*(E + x),$$

logo $\lambda^*(E + x) = \lambda^*(E)$.

Passo 2: Se $A \in \mathcal{L}(\mathbb{R}^n)$, então $A + x \in \mathcal{L}(\mathbb{R}^n)$.

Seja $E \subset \mathbb{R}^n$ arbitrário e ponha $F = E - x$; então $E = F + x$. Como translações comutam com interseção e complemento,

$$(F \cap A) + x = (F + x) \cap (A + x) = E \cap (A + x),$$

$$(F \cap A^c) + x = (F + x) \cap (A + x)^c = E \cap (A + x)^c.$$

Pelo Passo 1,

$$\lambda^*(E) = \lambda^*(F + x) = \lambda^*(F),$$

$$\lambda^*(E \cap (A + x)) = \lambda^*((F \cap A) + x) = \lambda^*(F \cap A),$$

$$\lambda^*(E \cap (A + x)^c) = \lambda^*((F \cap A^c) + x) = \lambda^*(F \cap A^c).$$

Como A é mensurável à Carathéodory (isto é, $A \in \mathcal{L}(\mathbb{R}^n)$), temos, para todo $F \subset \mathbb{R}^n$,

$$\lambda^*(F) = \lambda^*(F \cap A) + \lambda^*(F \cap A^c).$$

Substituindo as igualdades acima, obtemos

$$\lambda^*(E) = \lambda^*(E \cap (A + x)) + \lambda^*(E \cap (A + x)^c).$$

Como E foi arbitrário, concluímos que $A + x$ é mensurável no sentido de Carathéodory, isto é, $A + x \in \mathcal{L}(\mathbb{R}^n)$.

Passo 3: Como $A \in \mathcal{L}(\mathbb{R}^n)$ e $A + x \in \mathcal{L}(\mathbb{R}^n)$, segue que

$$\lambda_n(A + x) = \lambda^*(A + x) = \lambda^*(A) = \lambda_n(A),$$

onde usamos o Passo 1 e a definição $\lambda_n = \lambda^*|_{\mathcal{L}(\mathbb{R}^n)}$.

□

Observação 3.4.9. A invariância por translações expressa a homogeneidade do espaço euclidiano do ponto de vista da medida e é essencial para a definição de uma probabilidade geométrica uniforme, pois, ao deslocar um conjunto em seu respectivo espaço, sua medida não é alterada.

3.4.6 Homogeneidade (escalonamento) da medida de Lebesgue

Proposição 3.4.10. Seja $A \in \mathcal{L}(\mathbb{R}^n)$ e seja $\alpha > 0$. Então

$$\lambda_n(\alpha A) = \alpha^n \lambda_n(A),$$

onde $\alpha A = \{\alpha a : a \in A\}$.

Demonstração: Seja \mathcal{R} a família dos retângulos semiabertos de \mathbb{R}^n e μ_0 a pré-medida geométrica definida em \mathcal{R} por

$$\mu_0\left(\prod_{i=1}^n [a_i, b_i)\right) = \prod_{i=1}^n (b_i - a_i).$$

Passo 1: comportamento da pré-medida sob escalonamento.

Se $R = \prod_{i=1}^n [a_i, b_i) \in \mathcal{R}$ e $\alpha > 0$, então

$$\alpha R = \prod_{i=1}^n [\alpha a_i, \alpha b_i) \in \mathcal{R}$$

e

$$\mu_0(\alpha R) = \prod_{i=1}^n (\alpha b_i - \alpha a_i) = \alpha^n \prod_{i=1}^n (b_i - a_i) = \alpha^n \mu_0(R).$$

Passo 2: comportamento da medida exterior sob escalonamento.

Seja $E \subset \mathbb{R}^n$. Fixe $\varepsilon > 0$ e escolha uma cobertura $E \subset \bigcup_{k=1}^{\infty} R_k$ com $R_k \in \mathcal{R}$ tal que

$$\sum_{k=1}^{\infty} \mu_0(R_k) \leq \lambda^*(E) + \varepsilon.$$

Então

$$\alpha E \subset \bigcup_{k=1}^{\infty} (\alpha R_k)$$

e, pelo Passo 1,

$$\lambda^*(\alpha E) \leq \sum_{k=1}^{\infty} \mu_0(\alpha R_k) = \alpha^n \sum_{k=1}^{\infty} \mu_0(R_k) \leq \alpha^n (\lambda^*(E) + \varepsilon).$$

Como $\varepsilon > 0$ é arbitrário,

$$\lambda^*(\alpha E) \leq \alpha^n \lambda^*(E).$$

Aplicando essa desigualdade ao conjunto αE e ao fator $1/\alpha$, obtemos

$$\lambda^*(E) = \lambda^*\left(\frac{1}{\alpha}(\alpha E)\right) \leq \left(\frac{1}{\alpha}\right)^n \lambda^*(\alpha E),$$

isto é,

$$\alpha^n \lambda^*(E) \leq \lambda^*(\alpha E).$$

Portanto,

$$\lambda^*(\alpha E) = \alpha^n \lambda^*(E).$$

Passo 3: preservação da mensurabilidade.

Se $A \in \mathcal{L}(\mathbb{R}^n)$ e $E \subset \mathbb{R}^n$, ponha $F = \frac{1}{\alpha}E$. Então $E = \alpha F$ e, usando o Passo 2,

$$\lambda^*(E) = \alpha^n \lambda^*(F),$$

$$\lambda^*(E \cap (\alpha A)) = \lambda^*(\alpha(F \cap A)) = \alpha^n \lambda^*(F \cap A),$$

$$\lambda^*(E \cap (\alpha A)^c) = \lambda^*(\alpha(F \cap A^c)) = \alpha^n \lambda^*(F \cap A^c).$$

Como A é mensurável à Carathéodory,

$$\lambda^*(F) = \lambda^*(F \cap A) + \lambda^*(F \cap A^c),$$

logo,

$$\lambda^*(E) = \lambda^*(E \cap (\alpha A)) + \lambda^*(E \cap (\alpha A)^c).$$

Como E é arbitrário, $\alpha A \in \mathcal{L}(\mathbb{R}^n)$.

Passo 4: Como $A, \alpha A \in \mathcal{L}(\mathbb{R}^n)$, segue que

$$\lambda_n(\alpha A) = \lambda^*(\alpha A) = \alpha^n \lambda^*(A) = \alpha^n \lambda_n(A).$$

□

Observação 3.4.11. Essa propriedade justifica a interpretação geométrica da medida de Lebesgue como comprimento, área ou volume, conforme a dimensão do espaço, isto é, expressa o comportamento esperado da medida de Lebesgue sob mudanças de escala, formalizando, por exemplo, a conversão entre diferentes unidades de medida, como metros cúbicos e centímetros cúbicos.

3.4.7 Conjuntos de Medida Zero

Definição 3.4.12. Um conjunto $A \subset \mathbb{R}^n$ é dito de **medida zero** se $\lambda_n(A) = 0$.

Observação 3.4.13. Conjuntos de medida zero podem ser geometricamente não vazios, mas são desprezíveis do ponto de vista da medida e desempenham papel central na teoria da probabilidade contínua.

3.5 Interpretação probabilística dos conjuntos de medida zero

Observação 3.5.1. No contexto probabilístico, eventos correspondentes a conjuntos de medida zero são interpretados como eventos de probabilidade nula. Isso justifica, por exemplo, que a escolha aleatória de um ponto em um intervalo tenha probabilidade zero de coincidir com um ponto fixo.

3.5.0.1 Espaço de probabilidade contínuo

Definição 3.5.2. Um **espaço de probabilidade contínuo** é um espaço de probabilidade (Ω, \mathcal{F}, P) no qual Ω é um subconjunto mensurável de \mathbb{R}^n , \mathcal{F} é uma σ -álgebra de subconjuntos de Ω e P é obtida pela normalização da medida de Lebesgue em Ω .

Observação 3.5.3. A estrutura do espaço de probabilidade contínuo permite modelar experimentos aleatórios com infinitos resultados possíveis.

3.5.0.2 Normalização da medida em conjuntos de medida finita

Definição 3.5.4. Seja $\Omega \in \mathcal{L}(\mathbb{R}^n)$ tal que $0 < \lambda_n(\Omega) < \infty$. Define-se a função

$$P(A) = \frac{\lambda_n(A)}{\lambda_n(\Omega)}, \quad A \subset \Omega, A \in \mathcal{L}(\mathbb{R}^n),$$

como a normalização da medida de Lebesgue em Ω .

Observação 3.5.5. A normalização transforma a medida de Lebesgue em uma medida de probabilidade, preservando proporcionalidade geométrica.

3.5.0.3 Probabilidade induzida por medida (probabilidade geométrica)

Definição 3.5.6. A **probabilidade geométrica** em \mathbb{R}^n é a medida de probabilidade P induzida pela medida de Lebesgue, definida por

$$P(A) = \frac{\lambda_n(A)}{\lambda_n(\Omega)},$$

para todo evento mensurável $A \subset \Omega$, onde $\Omega \subset \mathbb{R}^n$ é mensurável e de medida finita.

Observação 3.5.7. Nos casos particulares $n = 1, 2, 3$, a probabilidade geométrica é interpretada, respectivamente, em termos de comprimento, área e volume.

3.6 Probabilidade Geométrica - Definição Formal

3.6.1 Espaço amostral geométrico

Definição 3.6.1. Um **espaço amostral geométrico** é um conjunto

$$\Omega \subset \mathbb{R}^n$$

tal que Ω é mensurável no sentido de Lebesgue e satisfaz

$$0 < \lambda_n(\Omega) < \infty.$$

Nesse contexto, os eventos são os subconjuntos mensuráveis de Ω , isto é, os conjuntos $A \subset \Omega$ pertencentes à σ -álgebra $\mathcal{L}(\mathbb{R}^n)$.

Observação 3.6.2. A exigência de mensurabilidade e de medida finita garante que a noção de probabilidade geométrica seja bem definida e normalizável.

3.6.2 Definição formal de probabilidade geométrica em \mathbb{R}^n

Definição 3.6.3. Seja $\Omega \subset \mathbb{R}^n$ um espaço amostral geométrico. Define-se a **probabilidade geométrica** associada a Ω como a função

$$P : \mathcal{L}(\Omega) \longrightarrow [0, 1]$$

onde

$$\mathcal{L}(\Omega) = \{A \subset \Omega \mid A \in \mathcal{L}(\mathbb{R}^n)\}.$$

E escreve-se:

$$P(A) = \frac{\lambda_n(A)}{\lambda_n(\Omega)}, \quad A \in \mathcal{L}(\Omega).$$

Observação 3.6.4. A função P assim definida é uma medida de probabilidade em $(\Omega, \mathcal{L}(\Omega))$. Teremos, portanto, a trinca $(\Omega, \mathcal{L}(\Omega), P)$ como o espaço de probabilidade que satisfaz os axiomas de Kolmogorov.

Nos casos particulares $n = 1, 2, 3$, a medida de Lebesgue λ_n admite interpretações geométricas clássicas:

- em \mathbb{R}^1 , λ_1 coincide com o comprimento;
- em \mathbb{R}^2 , λ_2 coincide com a área;

- em \mathbb{R}^3 , λ_3 coincide com o volume.

Consequentemente, a probabilidade geométrica nesses espaços é interpretada como a razão entre comprimentos, áreas ou volumes.

3.6.3 Modelo Linear - Formal

Definição 3.6.5. *Seja $\Omega \subset \mathbb{R}$ um conjunto mensurável tal que*

$$0 < \lambda_1(\Omega) < \infty.$$

Define-se a probabilidade geométrica na reta real como a medida de probabilidade

$$P(A) = \frac{\lambda_1(A)}{\lambda_1(\Omega)}, \quad A \subset \Omega, \quad A \in \mathcal{L}(\mathbb{R}),$$

onde λ_1 denota a medida de Lebesgue em \mathbb{R} .

Como λ_1 coincide com o comprimento de intervalos, essa definição pode ser interpretada como a razão entre comprimentos. Em particular, se $\Omega = [a, b]$ e $A \subset [a, b]$ é mensurável, então

$$P(A) = \frac{\text{comprimento de } A}{b - a}.$$

Um exemplo interessante que foi incentivado por um problema de [James \(2010\)](#) da aplicação dos conjuntos mensuráveis à Lebesgue é:

Exemplo 3.6.6. *Qual a probabilidade de, ao escolher um número aleatoriamente no intervalo $[0, 1]$, ele seja irracional?*

Solução: *Considere o espaço de probabilidade*

$$(\Omega, \mathcal{L}([0, 1]), P), \quad P(A) = \lambda_1(A),$$

onde λ_1 é a medida de Lebesgue restrita ao intervalo $[0, 1]$.

O conjunto dos números racionais em $[0, 1]$, dado por $\mathbb{Q} \cap [0, 1]$, é enumerável. Como todo conjunto enumerável possui medida de Lebesgue nula, temos

$$\lambda_1(\mathbb{Q} \cap [0, 1]) = 0.$$

O conjunto dos números irracionais em $[0, 1]$ é o complementar de $\mathbb{Q} \cap [0, 1]$ em $[0, 1]$, isto é,

$$[0, 1] \setminus (\mathbb{Q} \cap [0, 1]).$$

Assim,

$$P(\text{número irracional}) = 1 - P(\text{número racional}) = 1 - \lambda_1(\mathbb{Q} \cap [0, 1]) = 1.$$

Portanto, a probabilidade de que um número escolhido aleatoriamente no intervalo $[0, 1]$ seja irracional é igual a 1.

Uma observação interessante é que, nesse caso, como não se trata de um problema que envolva conjuntos patológicos, poderíamos também utilizar $\mathcal{B}([0, 1])$.

Outro fato interessante observado em [James \(2010\)](#) é que, nos problemas de probabilidade geométrica usualmente considerados, as regiões da reta, do plano ou do espaço, aquelas que podem ser descritas e visualizadas geometricamente, pertencem à σ -álgebra de Borel. Assim, toda a teoria desenvolvida até aqui pode ser aplicada a essas figuras sem a necessidade de lidar com conjuntos problemáticos da teoria da medida.

3.6.4 Modelo Superficial - Formal

Definição 3.6.7. *Seja $\Omega \subset \mathbb{R}^2$ um conjunto mensurável tal que*

$$0 < \lambda_2(\Omega) < \infty.$$

A probabilidade geométrica no plano é definida por

$$P(A) = \frac{\lambda_2(A)}{\lambda_2(\Omega)}, \quad A \subset \Omega, \quad A \in \mathcal{L}(\mathbb{R}^2),$$

onde λ_2 é a medida de Lebesgue bidimensional.

Nesse caso, λ_2 coincide com a área de conjuntos mensuráveis usuais, de modo que a probabilidade geométrica corresponde à razão entre áreas. Assim, quando Ω é uma região plana limitada e $A \subset \Omega$ é um evento mensurável, $P(A)$ representa a proporção da área de Ω ocupada por A .

3.6.5 Modelo Volumétrico - Formal

Definição 3.6.8. *Seja $\Omega \subset \mathbb{R}^3$ um conjunto mensurável tal que*

$$0 < \lambda_3(\Omega) < \infty.$$

Define-se a probabilidade geométrica no espaço tridimensional por

$$P(A) = \frac{\lambda_3(A)}{\lambda_3(\Omega)}, \quad A \subset \Omega, \quad A \in \mathcal{L}(\mathbb{R}^3),$$

onde λ_3 denota a medida de Lebesgue em \mathbb{R}^3 .

Nesse contexto, λ_3 coincide com o volume de sólidos mensuráveis, e a probabilidade geométrica é interpretada como a razão entre volumes. Assim, se Ω representa um sólido de volume finito e $A \subset \Omega$ é um subconjunto mensurável, $P(A)$ mede a fração do volume total de Ω correspondente ao evento A .

Observação 3.6.9. *Nos três casos anteriores, a probabilidade geométrica é obtida pela normalização da medida de Lebesgue no espaço considerado. A exigência de mensurabilidade e de medida total finita é essencial para garantir que a função P seja uma medida de probabilidade bem definida.*

3.6.6 Condições de validade da probabilidade geométrica

Observação 3.6.10. *A definição de probabilidade geométrica pressupõe implicitamente:*

1. *a mensurabilidade dos eventos considerados;*
2. *a finitude e positividade da medida do espaço amostral;*
3. *o uso da medida de Lebesgue como modelo de homogeneidade do espaço;*
4. *a invariância por translações, garantindo ausência de viés geométrico.*

Na ausência dessas condições, a interpretação probabilística pode falhar ou tornar-se ambígua.

Dessa forma, a probabilidade geométrica surge naturalmente como uma normalização da medida de Lebesgue em subconjuntos mensuráveis de \mathbb{R}^n , fornecendo um modelo rigoroso para experimentos aleatórios contínuos.

Segundo [James \(2010\)](#), os conjuntos que surgem em problemas geométricos usuais, definidos a partir de regiões limitadas por curvas ou superfícies regulares, são mensuráveis à Lebesgue. Assim, ao longo deste trabalho, consideraremos apenas conjuntos mensuráveis, evitando construções patológicas que não aparecem no contexto da probabilidade geométrica, principalmente se tratando do Ensino Básico.

3.6.7 Como encarar a probabilidade geométrica daqui para frente

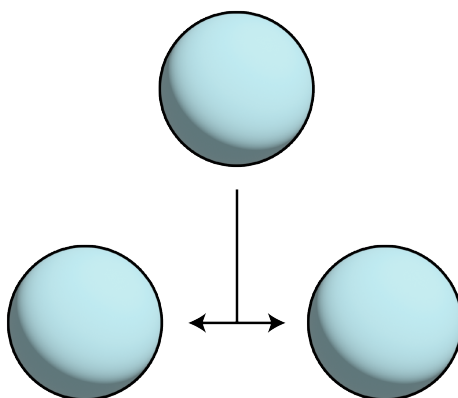
Nos problemas de natureza geométrica que serão considerados ao longo deste trabalho, os conjuntos envolvidos correspondem, em geral, a regiões limitadas por curvas ou superfícies regulares, bem como a combinações finitas dessas regiões. Tais conjuntos pertencem à σ -álgebra de Borel e, conseqüentemente, são mensuráveis no sentido de Lebesgue. Por esse motivo, não será necessário, nos exemplos e aplicações seguintes, verificar explicitamente a mensurabilidade dos conjuntos considerados. A probabilidade geométrica poderá, assim, ser interpretada diretamente como a razão entre medidas de Lebesgue, refletindo a proporcionalidade geométrica entre comprimentos, áreas ou volumes.

3.6.8 O que ocorre ao relaxarmos os axiomas da medida de Lebesgue?

O exemplo a seguir é o famoso Paradoxo de Banach-Tarski, que mostra ser possível dividir uma bola de volume V em duas bolas de mesmo volume V . Esse resultado não contradiz a teoria da medida de Lebesgue, pois ocorre apenas quando se abandonam hipóteses como a mensurabilidade dos conjuntos envolvidos. Segue a descrição do paradoxo.

Exemplo 3.6.11. *O chamado Paradoxo de Banach-Tarski é um resultado da geometria e da teoria dos conjuntos que afirma que, em espaços euclidianos de dimensão maior ou igual a três, é possível decompor uma bola sólida em um número finito de subconjuntos disjuntos e, utilizando apenas rotações e translações, recompor esses pedaços de modo a obter duas bolas sólidas idênticas à original, como na figura 7.*

Figura 7 – Uma bola pode, com movimentos de rotação e translação, se decompor em duas exatamente iguais.



Fonte: Elaborada pelo autor (2026).

Explicação: *À primeira vista, o Paradoxo de Banach-Tarski parece contradizer a noção usual de volume, uma vez que sugere a duplicação de um sólido sem acréscimo de matéria. No entanto, não há contradição com a teoria da medida de Lebesgue. Os subconjuntos envolvidos na decomposição paradoxal não são mensuráveis no sentido de Lebesgue, de modo que a noção de volume simplesmente não se aplica a eles. O paradoxo evidencia que, em dimensões mais altas, conjuntos extremamente irregulares podem surgir quando se admite o Axioma da Escolha, inviabilizando a extensão da noção geométrica de volume a todos os subconjuntos do espaço.*

A menção ao Paradoxo de Banach-Tarski tem por objetivo destacar que as restrições impostas pela teoria da medida não são meramente técnicas, mas refletem limitações

estruturais profundas da geometria dos espaços euclidianos. Esse resultado mostra que, em dimensões maiores ou iguais a três, não é possível definir uma noção de volume que seja simultaneamente invariante por isometrias e aditiva, mesmo quando se exige apenas aditividade finita. Dessa forma, a medida de Lebesgue surge como o limite natural e matematicamente coerente de qualquer tentativa de formalizar a noção de volume e, conseqüentemente, de probabilidade geométrica em \mathbb{R}^n . Como esse paradoxo foge do escopo do trabalho, ele foi citado apenas a título de curiosidade. Para mais, consulte [Tomkiewicz e Wagon \(2016\)](#) e [Banach e Tarski \(1924\)](#).

3.7 Modelo de Coordenadas Geométricas

Além das formulações puramente geométricas, a probabilidade geométrica também pode ser tratada de forma analítica, por meio de um modelo de coordenadas geométricas. Nesse contexto, os pontos do espaço amostral são representados por coordenadas reais, em uma reta, plano ou espaço, e os eventos correspondem a subconjuntos definidos por equações ou desigualdades.

Assim, problemas de probabilidade geométrica passam a ser formulados e resolvidos por meio de comprimentos na reta, áreas ou volumes em regiões do espaço de coordenadas, o que permite uma abordagem mais geral e precisa, mas ainda sim baseando-se em tudo que foi exposto. Esse tipo de formulação é útil em problemas que envolvem condições algébricas entre variáveis, como distâncias, ângulos ou desigualdades envolvendo coordenadas, permitindo uma transição natural entre a interpretação geométrica e a formulação analítica da probabilidade. Vejamos a seguir dois problemas que introduzem bem essa ideia.

3.7.1 O Problema do Encontro

O problema do encontro é um caso clássico do modelo de coordenadas geométricas que será apresentado agora. Algumas pequenas variações podem aparecer no enunciado. O caso que estudamos foi apresentado em [Tunala \(1992\)](#) e transcrito logo a seguir.

Problema 3.7.1. *Duas pessoas decidiram se encontrar em um determinado local entre 11 e 12 horas. Combinou-se previamente que a primeira pessoa a chegar esperará no máximo 15 minutos pela outra. Ache a probabilidade P_4 deste encontro realizar-se neste intervalo, admitindo-se que os instantes de chegada (entre 11 e 12 horas) de cada uma das pessoas provêm do acaso.*

Solução: *Sendo aleatórios os instantes de chegada das duas pessoas em $[11, 12]$, podemos associá-los a pontos de dois segmentos de comprimento unitário (amplitude do intervalo), representados em eixos ortogonais x e y em \mathbb{R}^2 . Cada ponto $(x, y) \in [0, 1] \times [0, 1]$ terá*

coordenadas x, y numericamente iguais às frações de horas dos respectivos instantes de chegada, $11 + x$ e $11 + y$, das duas pessoas.

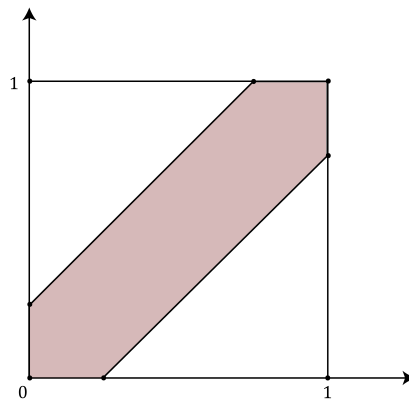
De acordo com o enunciado, o encontro somente terá lugar se

$$|y - x| \leq \frac{1}{4},$$

ou seja, se

$$\begin{cases} y \leq x + \frac{1}{4}, \\ y \geq x - \frac{1}{4}. \end{cases}$$

Figura 8 – O evento do problema do encontro está em destaque.



Fonte: Elaborada pelo autor (2026).

Essas restrições definem a região destacada da figura 8. Por isso, basta calcular a área formada pelos dois triângulos congruentes, cada um com base e altura iguais a $\frac{3}{4}$ e, então, calcular o complementar disso. Assim,

$$P_4 = \frac{\text{Área de } g}{\text{Área de } G} = \frac{1 - 2 \cdot \left(\frac{1}{2} \cdot \frac{3}{4} \cdot \frac{3}{4}\right)}{1} = \frac{7}{16}.$$

Para calcular essa área g , utilizamos o complementar.

3.7.2 O Problema do Espaguete

Um problema bastante conhecido em probabilidade geométrica é o chamado **Problema do Espaguete**. Ele mostra como situações aleatórias simples podem ser estudadas por meio de desenhos geométricos no plano. A ideia central é transformar as possíveis

divisões de um segmento de reta em pontos de uma região geométrica, permitindo que a probabilidade seja calculada a partir de razões entre áreas. A inspiração do enunciado e da resposta foi retirada de [Wagner \(s.d.\)](#) e [Morgado e Carvalho \(2015\)](#). O nome do problema vem pelo fato de um espagete se assemelhar a um segmento de reta, e de que é possível fazer esse problema em sala de aula utilizando esse tipo de objeto.

Problema 3.7.2. *Dividindo aleatoriamente um segmento em três partes, qual é a probabilidade de que esses novos segmentos formem um triângulo?*

Solução: *Considere um segmento de comprimento 1. Escolhem-se dois pontos ao acaso sobre esse segmento, dividindo-o em três partes. A questão que se coloca é: qual é a probabilidade de que esses três segmentos possam formar um triângulo?*

Sejam x e y os dois pontos escolhidos em $[0, 1]$. Sem perda de generalidade, podemos supor

$$0 \leq x \leq y \leq 1.$$

Nesse caso, os comprimentos dos três segmentos formados são

$$x, \quad y - x, \quad 1 - y.$$

Recordemos que três segmentos positivos formam um triângulo se, e somente se, cada um deles é menor do que a soma dos outros dois (desigualdade triangular). Como a soma dos três comprimentos é igual a 1, essas condições são equivalentes às desigualdades

$$x < \frac{1}{2}, \quad y - x < \frac{1}{2}, \quad 1 - y < \frac{1}{2}.$$

A última desigualdade equivale a $y > \frac{1}{2}$. Portanto, a condição necessária e suficiente para que os três segmentos formem um triângulo é

$$\frac{1}{2} < y < x + \frac{1}{2}, \quad 0 \leq x \leq y \leq 1.$$

A figura 9 mostra o espaço amostral (triângulo maior) e o evento (triângulo interno menor).

- *a área do triângulo total (espaço amostral) é $\frac{1}{2}$;*
- *a região favorável (evento) é um triângulo de base $\frac{1}{2}$ e altura $\frac{1}{2}$, cuja área é $\frac{1}{8}$.*

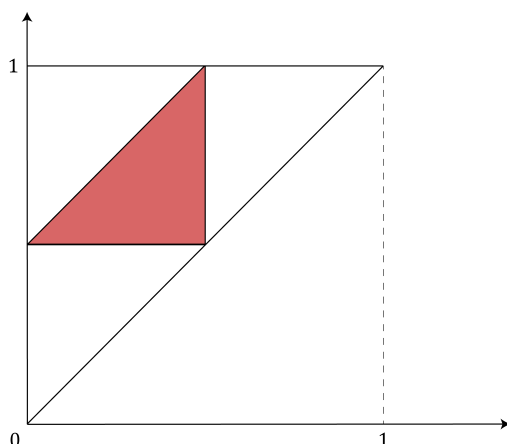
Assim, a probabilidade de que os três segmentos formem um triângulo é

$$\mathbb{P} = \frac{\frac{1}{8}}{\frac{1}{2}} = \frac{1}{4}.$$

Concluimos, portanto, que a probabilidade de que um segmento dividido aleatoriamente em três partes gere três comprimentos capazes de formar um triângulo é igual a

$$\frac{1}{4}.$$

Figura 9 – Espaço amostral e evento do problema do espagete.



Fonte: Elaborada pelo autor (2026).

3.8 Espaço Amostral Infinito Enumerável

Nos exemplos elementares de probabilidade, o espaço amostral costuma ser finito, como no lançamento de um dado ou de uma moeda. Nesses casos, quando todos os resultados são considerados equiprováveis, a probabilidade de um evento é obtida simplesmente pela razão entre o número de casos favoráveis e o número total de casos possíveis. Essa definição, porém, deixa de ser adequada quando o espaço amostral passa a ser infinito, ainda que permaneça enumerável.

Considere, por exemplo, um experimento cujo conjunto de resultados possíveis é $\Omega = \mathbb{N}$, ou o conjunto de todas as seqüências finitas de lançamentos de uma moeda até a ocorrência da primeira face “cara”. Embora tais conjuntos sejam infinitos, eles podem ser colocados em correspondência biunívoca com os números naturais, o que os caracteriza como espaços amostrais enumeráveis. Nesses contextos, não faz sentido falar em “casos igualmente prováveis” no mesmo sentido do caso finito, pois não existe uma maneira consistente de atribuir a mesma probabilidade positiva a cada ponto de Ω sem violar o axioma da soma total igual a 1.

A abordagem correta, portanto, consiste em definir diretamente uma função de probabilidade

$$P : \mathcal{P}(\Omega) \rightarrow [0, 1],$$

em que Ω é um conjunto enumerável, satisfazendo as propriedades fundamentais de não

negatividade, normalização e aditividade enumerável. Em particular, a probabilidade de um evento $A \subset \Omega$ é dada por

$$P(A) = \sum_{w \in A} P(\{w\}),$$

desde que a série seja convergente. Normalmente, para o Ensino Básico, o tratamento dessas séries se dará pela soma dos termos de progressões geométricas de razão $0 < q < 1$, uma vez que elas são sempre convergentes. Assim, a probabilidade em espaços enumeráveis infinitos é construída a partir da atribuição de probabilidades aos eventos elementares, exigindo necessariamente o uso de séries numéricas e da noção de soma infinita. Veja o exemplo abaixo, que foi inspirado em um dos problemas de [GRINSTEAD e SNELL \(2006\)](#):

Exemplo 3.8.1. *Considere um experimento aleatório no qual uma moeda é lançada repetidamente até que ocorra, pela primeira vez, o resultado “cara”. O espaço amostral é*

$$\Omega = \{1, 2, 3, \dots\},$$

em que o número $\omega \in \Omega$ representa a quantidade de lançamentos necessários até a primeira ocorrência de “cara”. Trata-se de um conjunto infinito, porém enumerável.

Se a moeda é honesta, cada lançamento é independente e a probabilidade de “cara” em cada tentativa é $p = \frac{1}{2}$. Definimos a probabilidade pontual por

$$P(\{w\}) = \left(\frac{1}{2}\right)^w, \quad w \in \mathbb{N}$$

Observe que não é um espaço equiprovável: eventos correspondentes a valores maiores de ω possuem probabilidades cada vez menores.

A normalização é garantida pelo fato de que

$$\sum_{w=1}^{\infty} P(\{w\}) = \sum_{w=1}^{\infty} \left(\frac{1}{2}\right)^w = 1$$

ou seja, a soma infinita das probabilidades pontuais converge para 1. Para qualquer evento $A \subset \Omega$, a probabilidade é dada por

$$P(A) = \sum_{w \in A} \left(\frac{1}{2}\right)^w,$$

o que evidencia a necessidade do uso de séries numéricas (nesse caso, soma da PG) e da aditividade enumerável.

Esse exemplo mostra que, em espaços enumeráveis infinitos, a probabilidade não decorre de contagem de casos, mas da atribuição explícita de uma distribuição compatível com a soma total unitária.

Esse caso intermediário, infinito, mas discreto, desempenha um papel conceitual importante: ele evidencia que o cerne da teoria da probabilidade não está na contagem de casos, mas na definição de uma medida que respeite a estrutura do espaço amostral. Essa ideia é essencial para espaços contínuos, nos quais a noção de probabilidade pontual deixa de fazer sentido e exige ferramentas mais sofisticadas, como as medidas em \mathbb{R}^n . Mais exemplos desse tipo de caso estão disponíveis na seção 4.

3.9 Problemas Clássicos de Probabilidade Geométrica

3.9.1 A Agulha de Buffon e de Buffon-Laplace

O problema da agulha de Buffon foi proposto pelo naturalista e matemático francês Georges-Louis Leclerc, Conde de Buffon, em 1777, sendo um dos primeiros problemas conhecidos de probabilidade geométrica. A questão buscava determinar a probabilidade de uma agulha cair sobre uma linha em um plano com linhas paralelas igualmente espaçadas e marcou o início do uso da probabilidade para medir áreas e comprimentos, antecipando conceitos da teoria da medida e do método de Monte Carlo (usado séculos depois em simulações computacionais) (JOHNSON, 1995). No caso desse problema, deve-se ter um conhecimento prévio de cálculo integral para compreender completamente sua solução.

Problema 3.9.1 (Agulha de Buffon). *Considere um plano no qual foram traçadas linhas paralelas igualmente espaçadas, separadas por uma distância d . Lança-se ao acaso uma agulha de comprimento L (com $L \leq d$) sobre esse plano. Deseja-se determinar a probabilidade P de que a agulha intersecte uma das linhas.*

Solução: *Seja θ o ângulo entre a agulha e as linhas paralelas, com $0 \leq \theta \leq \frac{\pi}{2}$, e seja x a distância do centro da agulha à linha mais próxima, variando de 0 a $\frac{d}{2}$, como mostra a figura 10.*

Observando a figura 10, na situação central, percebe-se que

$$\cos \theta = \frac{x}{\frac{L}{2}} = \frac{2x}{L} \Rightarrow x = \frac{L}{2} \cdot \cos \theta.$$

Daí, temos que a agulha intersecta uma linha se:

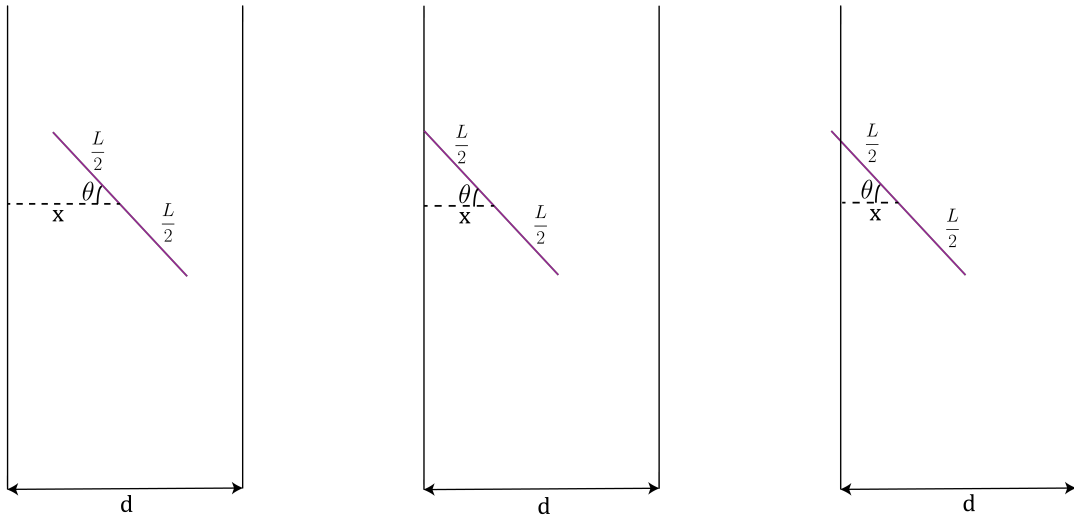
$$x \leq \frac{L}{2} \cos \theta.$$

O espaço amostral é formado por todos os pares (θ, x) possíveis, com área total:

$$\lambda_2(\Omega) = \frac{d}{2} \cdot \frac{\pi}{2} = \frac{\pi d}{4}.$$

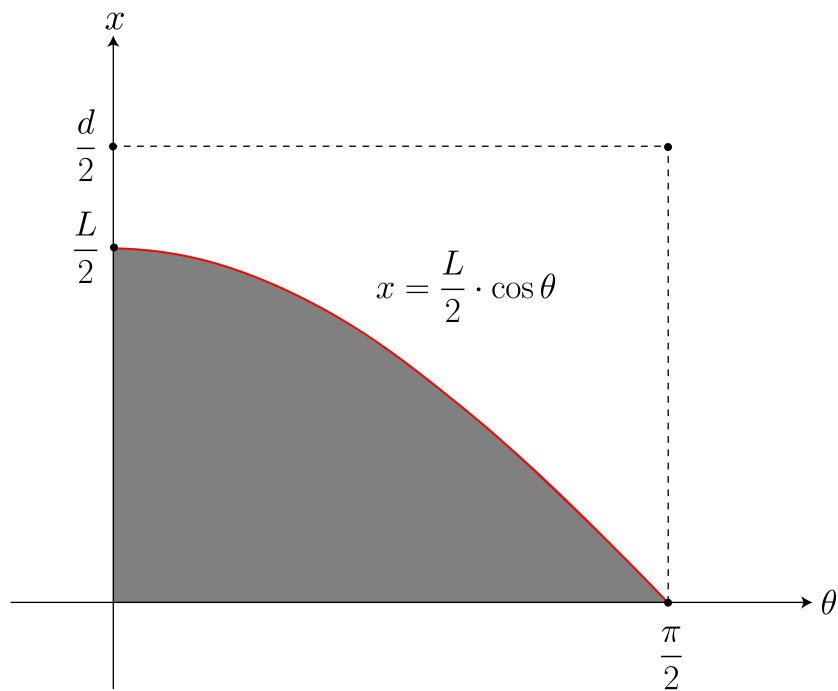
A ilustração no plano cartesiano do espaço amostral e do evento está representado na figura 11.

Figura 10 – Três situações: uma em que a agulha não toca a linha, uma em que ela toca a linha com sua extremidade e outro em que ela atravessa a linha.



Fonte: Elaborada pelo autor (2026).

Figura 11 – representação gráfica do problema.



Fonte: Elaborada pelo autor (2026).

A área favorável (ou seja, o conjunto dos pontos (θ, x) tal que $\theta \in \left[0, \frac{\pi}{2}\right]$, $x \in \left[0, \frac{d}{2}\right]$ e $x \leq \frac{L}{2} \cdot \cos \theta$), pode ser calculada por meio de integral definida. Daí:

$$\lambda_2(A) = \int_0^{\frac{\pi}{2}} \frac{L}{2} \cos \theta \, d\theta = \frac{L}{2}.$$

Assim, a probabilidade é dada por:

$$P(A) = \frac{\lambda_2(A)}{\lambda_2(\Omega)} = \frac{2L}{d\pi}.$$

Perceba que:

1. O resultado é válido apenas quando $L \leq d$.
2. No caso em que $L = d$, temos $P = \frac{2}{\pi}$.
3. Buffon mostrou que este experimento pode ser usado para estimar o valor de π empiricamente.

Em [Boyer e Merzbach \(2012\)](#), enuncia-se uma versão estendida do problema da agulha de Buffon, que foi enunciado posteriormente por Laplace. A ideia seria, além de construir as retas paralelas, criar um conjunto de retas perpendiculares, criando um reticulado. O problema, seguido de sua solução, será apresentado agora, no problema [3.9.2](#). Para resolvermos esse problema, utilizaremos a propriedade de que, dados dois eventos A e B , temos $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. Outros autores resolvem esse problema utilizando a propriedade $P(A) = 1 - P(A^c)$, como é o caso de [Solomon \(1978\)](#).

Problema 3.9.2 (Agulha de Buffon-Laplace). *Considere um plano com linhas verticais espaçadas por $a > 0$ e linhas horizontais espaçadas por $b > 0$, formando retângulos congruentes. Lança-se ao acaso uma agulha (segmento) de comprimento $L \leq \min\{a, b\}$ com posição e orientação uniformes. Qual é a probabilidade P de a agulha intersectar ao menos uma das linhas do reticulado?*

Solução: Sejam:

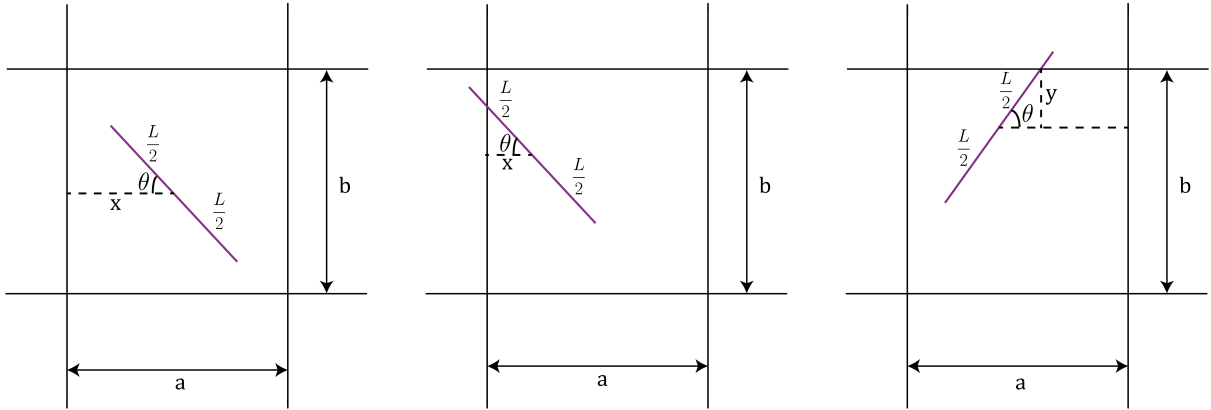
- $\theta \in \left[0, \frac{\pi}{2}\right]$: ângulo entre a agulha e o eixo x ;
- $x \in \left[0, \frac{a}{2}\right]$: distância do centro da agulha à reta vertical mais próxima;
- $y \in \left[0, \frac{b}{2}\right]$: distância do centro à reta horizontal mais próxima.

A agulha cruza:

$$\begin{cases} \text{alguma linha vertical se } x \leq \frac{L}{2} \cos \theta, \\ \text{alguma linha horizontal se } y \leq \frac{L}{2} \sin \theta. \end{cases}$$

Na figura [12](#), são apresentadas três situações: uma em que a agulha não toca nenhuma linha, uma em que ela toca uma linha vertical e uma em que ela toca uma horizontal.

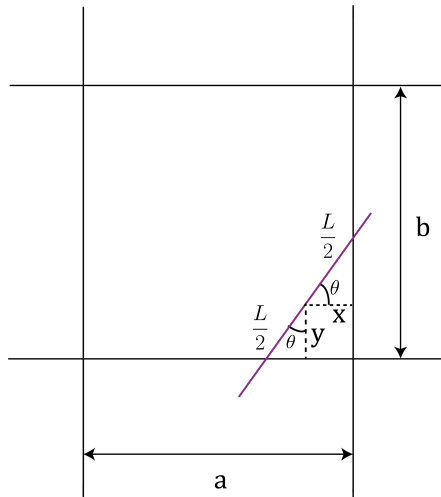
Figura 12 – Não toca, toca vertical, toca horizontal.



Fonte: Elaborada pelo autor (2026).

A figura 13 representa a situação em que a agulha toca em duas retas, simultaneamente.

Figura 13 – Agulha tocando simultaneamente ambas as retas.



Fonte: Elaborada pelo autor (2026).

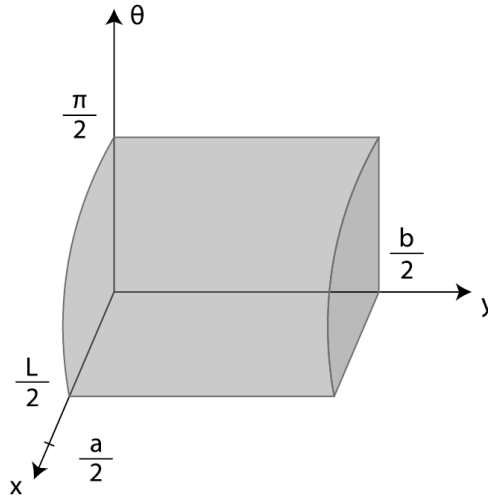
Com isso, temos que o espaço amostral é $\Omega = \{(x, y, \theta); x \in [0, \frac{a}{2}], y \in [0, \frac{b}{2}], \theta \in [0, \frac{\pi}{2}]\}$, configurando um paralelepípedo de volume:

$$\lambda_3(\Omega) = \frac{a}{2} \cdot \frac{b}{2} \cdot \frac{\pi}{2} = \frac{\pi ab}{8}.$$

Agora, devemos analisar cada um dos casos favoráveis. Primeiramente, consideremos o evento A , que é quando a agulha toca uma linha vertical. Para isso, devemos ter $x \leq \frac{L}{2} \cos \theta$, assim como foi no problema 3.9.1. Por outro lado, nesse caso, o valor de y pode assumir qualquer valor em $[0, \frac{b}{2}]$, já que consideramos sempre a menor distância do

centro da agulha ao reticulado. A figura 14 representa geometricamente o espaço amostral A .

Figura 14 – Representação tridimensional da região associada ao evento A .



Fonte: Elaborada pelo autor (2026).

Assim, teremos que o volume de A ($\lambda_3(A)$) será:

$$\lambda_3(A) = \frac{b}{2} \cdot \int_0^{\frac{\pi}{2}} \frac{L}{2} \cdot \cos \theta d\theta = \frac{b \cdot L}{4} \cdot \int_0^{\frac{\pi}{2}} \cos \theta d\theta = \frac{b \cdot L}{4}.$$

Por outro lado, procedendo da mesma forma para as linhas horizontais, evento que chamaremos de B , teremos que $y \leq \frac{L}{2} \cdot \sin \theta$ e o valor de x podendo variar no intervalo $[0, \frac{a}{2}]$ sob o mesmo pretexto do evento anterior. Assim, obteremos o evento descrito pela figura 15 cujo volume será dado por:

$$\lambda_3(B) = \frac{a}{2} \cdot \int_0^{\frac{\pi}{2}} \frac{L}{2} \cdot \sin \theta d\theta = \frac{a \cdot L}{4} \cdot \int_0^{\frac{\pi}{2}} \sin \theta d\theta = \frac{a \cdot L}{4}.$$

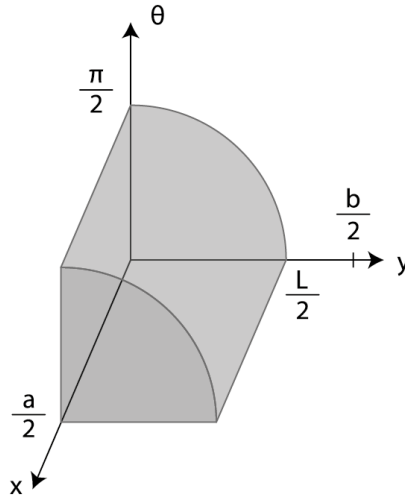
Como $\lambda_3(A \cup B) = \lambda_3(A) + \lambda_3(B) - \lambda_3(A \cap B)$ (essa propriedade fica mais clara ao observar a figura 17), calculemos $\lambda_3(A \cap B)$. Nesse caso, devemos ter ambas as situações se confirmando, a saber, $x \leq \frac{L}{2} \cos \theta$ e $y \leq \frac{L}{2} \sin \theta$. A figura 16 ilustra essa situação:

Nesse caso, para cada valor de θ , teremos uma área que será dada por

$$A(\theta) = \left(\frac{L}{2} \cos \theta \right) \cdot \left(\frac{L}{2} \sin \theta \right) = \frac{L^2}{4} \cdot \cos \theta \cdot \sin \theta.$$

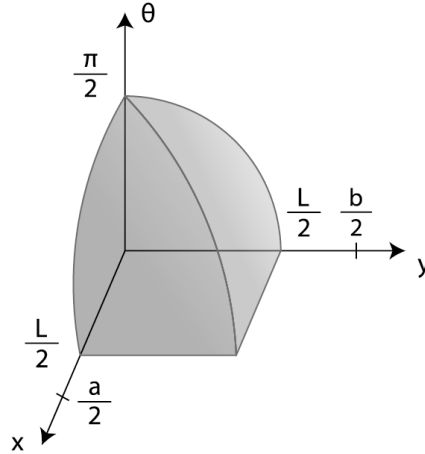
Como θ pertence ao intervalo $[0, \frac{\pi}{2}]$, isso é uma situação clássica do cálculo da área sob a curva. Por isso, basta calcular sua integral definida de zero à $\frac{\pi}{2}$ para encontrar $\lambda_3(A \cap B)$. Daí, obtemos

Figura 15 – Representação tridimensional da região associada ao evento B .



Fonte: Elaborada pelo autor (2026).

Figura 16 – Representação tridimensional da região associada ao evento $A \cap B$.



Fonte: Elaborada pelo autor (2026).

$$\lambda_3(A \cap B) = \int_0^{\frac{\pi}{2}} A(\theta) d\theta = \int_0^{\frac{\pi}{2}} \frac{L^2}{4} \cdot \cos \theta \cdot \sin \theta d\theta = \frac{L^2}{4} \cdot \int_0^{\frac{\pi}{2}} \cos \theta \cdot \sin \theta d\theta.$$

Pelas identidades trigonométricas, sabe-se que $\sin(2\theta) = 2 \sin \theta \cos \theta$. Substituindo, obtemos:

$$\lambda_3(A \cap B) = \frac{L^2}{8} \cdot \int_0^{\frac{\pi}{2}} \sin(2\theta) d\theta = \frac{L^2}{8} \cdot \left[-\frac{\cos 2\theta}{2} \right]_0^{\frac{\pi}{2}} = \frac{L^2}{8} \cdot \left[-\left(-\frac{1}{2} - \frac{1}{2} \right) \right] = \frac{L^2}{8}.$$

Logo,

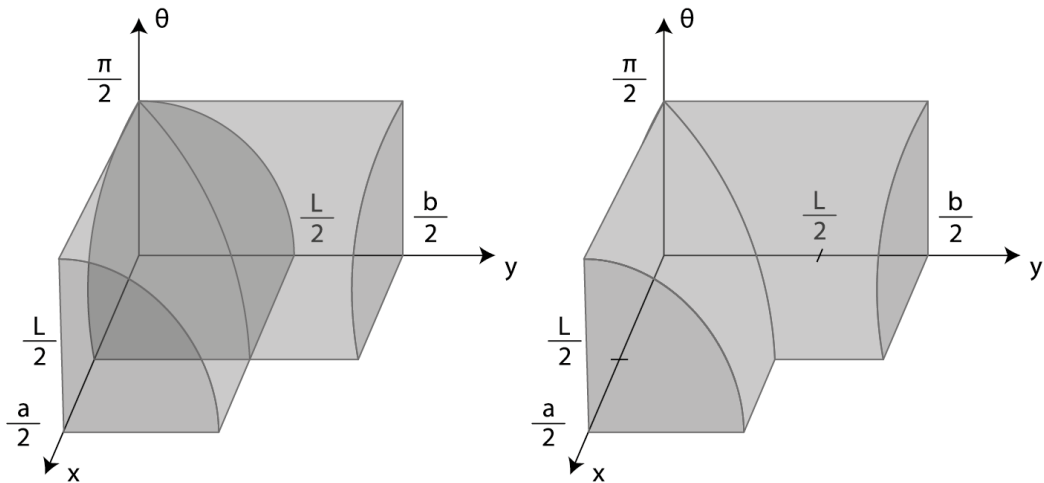
$$P(A \cup B) = \frac{\lambda_3(A) + \lambda_3(B) - \lambda_3(A \cap B)}{\lambda_3(\Omega)}$$

$$P(A \cup B) = \frac{\frac{bL}{4} + \frac{aL}{4} - \frac{L^2}{8}}{\frac{\pi ab}{8}} = \frac{\frac{2L(a+b) - L^2}{8}}{\frac{\pi ab}{8}} = \frac{2L(a+b) - L^2}{\pi ab}$$

$$P(A \cup B) = \frac{2L(a+b) - L^2}{\pi ab}$$

Na figura 17, é ilustrado o evento $A \cup B$, que é um pouco difícil de visualizar mentalmente à primeira vista.

Figura 17 – Representação tridimensional da região associada ao evento $A \cup B$.



Fonte: Elaborada pelo autor (2026).

Uma curiosidade é que, se tivermos $L = a = b$, a probabilidade torna-se:

$$P(A \cup B) = \frac{2L(L + L) - L^2}{\pi L^2} = \frac{3L^2}{\pi L^2}$$

$$P(A \cup B) = \frac{3}{\pi}$$

3.9.2 Paradoxo de Bertrand

O denominado **paradoxo de Bertrand** surge de uma aparente contradição em um problema de probabilidade geométrica proposto por Joseph Bertrand, em 1889. O enunciado clássico é o seguinte:

“Qual é a probabilidade de que uma corda traçada ao acaso em um círculo seja maior que o lado do triângulo equilátero inscrito nesse mesmo círculo?”

À primeira vista, o problema parece bem definido. No entanto, a expressão “corda escolhida ao acaso” é ambígua, e justamente essa ambiguidade conduz a resultados diferentes conforme o método adotado para definir o que significa “escolher aleatoriamente” uma corda. Assim, o chamado paradoxo não decorre de um erro matemático, mas da *indeterminação do espaço amostral* utilizado em cada abordagem. Em outras palavras, cada método define uma maneira distinta de gerar as cordas e, conseqüentemente, um *espaço de probabilidade diferente*, o que conduz a valores de probabilidade incompatíveis entre si, ainda que cada um deles seja internamente coerente.

Bertrand apresentou três formas de interpretar a escolha aleatória:

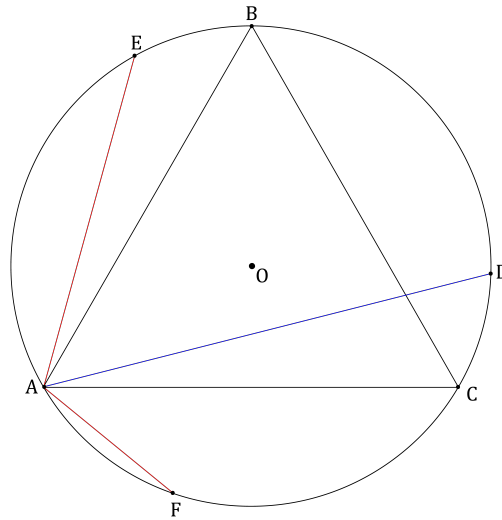
1. Método do arco:

Escolhe-se uma corda de modo que o *menor arco* por ela determinado na circunferência tenha medida angular superior a 120° (uma vez que qualquer lado de um triângulo equilátero inscrito na circunferência terá seu menor arco medindo 120°). Para isso, escolhe-se o primeiro ponto e, depois, é escolhido o segundo de modo que satisfaça o que é pedido. Na figura 18, perceba que a corda AD corresponde com as expectativas, enquanto as cordas AE e AF , não. Isso ocorre pelo fato de que o menor arco formado pela corda AD é maior do que 120° . Como na figura os segmentos estão partindo do ponto A , esse fato é apenas para que a ideia fique mais visível, pois o segmento que queremos não precisa partir de A . Um exemplo disso é o da figura 19, que é escolhido, aleatoriamente, um ponto D e traça-se a corda DE de modo que o menor arco de DE seja maior do que 120° . Com base nessa construção geométrica, conclui-se que a probabilidade de a corda ser maior do que o lado do triângulo equilátero inscrito é igual a:

$$P = \frac{1}{3}.$$

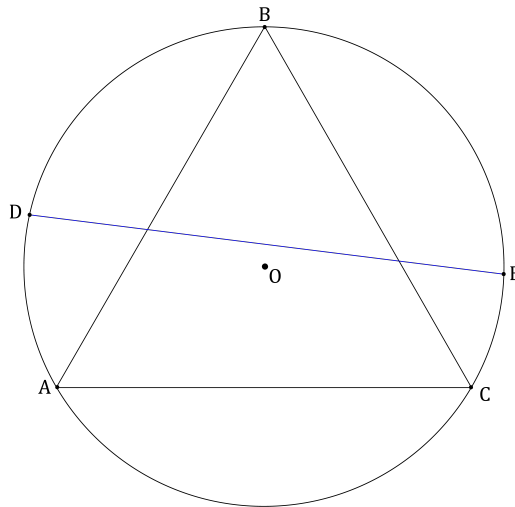
Do ponto de vista formal, a construção descrita no método do arco pode ser modelada por meio de um espaço de probabilidade contínuo bem definido. Observa-se que cada corda considerada pode ser identificada de maneira unívoca com o menor arco por ela determinado na circunferência, o qual pode ser parametrizado por um ângulo central. Assim, tomamos como espaço amostral o conjunto $\Omega = [0, \pi]$, em que cada elemento representa a medida angular do menor arco associado a uma corda escolhida. Sobre esse conjunto, consideramos a σ -álgebra de Borel $\mathcal{F} = \mathcal{B}([0, \pi])$, garantindo que todos os eventos de interesse sejam mensuráveis. A medida de probabilidade P é

Figura 18 – Representação do evento válido para a primeira interpretação do problema.



Fonte: Elaborada pelo autor (2026).

Figura 19 – Representação do evento válido ao selecionar inicialmente um ponto D diferente de algum dos vértices do triângulo.



Fonte: Elaborada pelo autor (2026).

então definida como:

$$P(A) = \frac{\lambda(A)}{\lambda([0, \pi])} = \frac{\lambda(A)}{\pi}, \quad A \in \mathcal{F}.$$

Essa escolha traduz matematicamente a hipótese equiprovável dos arcos possíveis e completa a descrição formal do espaço de probabilidade associado ao método do arco. Nesse modelo, o evento de interesse corresponde às cordas cujo comprimento é estritamente maior que o lado do triângulo equilátero inscrito na circunferência. Como cada corda é identificada com a medida angular do menor arco que ela

determina, e como um lado do triângulo equilátero corresponde a um arco de 120° , isto é, $\frac{2\pi}{3}$, tal evento é descrito pelo subconjunto

$$E = \left(\frac{2\pi}{3}, \pi\right) \subset \Omega,$$

formado por todos os arcos cuja medida angular é superior a 120° e cujo comprimento seria $\lambda(E) = \pi - \frac{2\pi}{3} = \frac{\pi}{3}$. Finalmente, com isso teríamos

$$P(E) = \frac{\frac{\pi}{3}}{\pi} = \frac{1}{3}.$$

Uma forma alternativa de descrever o experimento aleatório consiste em fixar inicialmente um ponto da circunferência, o que é lícito em virtude da invariância do problema por rotações (como o experimento é invariante por rotações, pode-se fixar um ponto da circunferência sem perda de generalidade, considerando apenas a posição relativa do segundo ponto), e escolher em seguida um segundo ponto de maneira uniforme ao longo da circunferência. Essa identificação pode ser interpretada como a abertura da circunferência e sua correspondência com um intervalo real, de modo que tomamos como espaço amostral o conjunto $\Omega = [0, 2\pi r)$, onde r é o raio da circunferência. Sobre esse conjunto, considera-se a σ -álgebra de Borel $\mathcal{F} = \mathcal{B}([0, 2\pi r))$ e a medida de probabilidade P definida por

$$P(A) = \frac{\lambda(A)}{2\pi r}, \quad A \in \mathcal{F}.$$

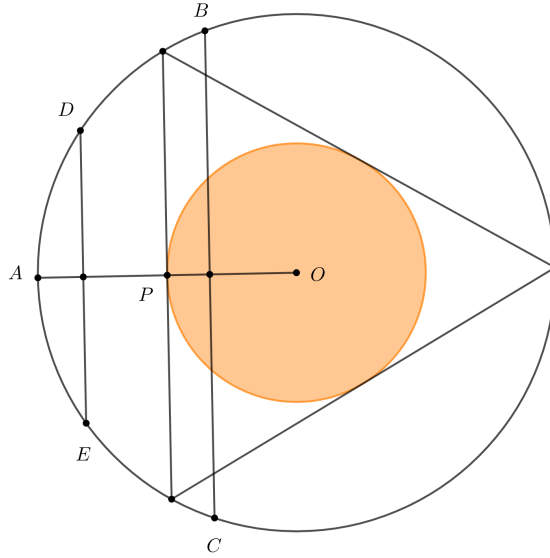
Nesse modelo, o evento pode ser definido como o intervalo $A = \left(\frac{2\pi r}{3}, \frac{4\pi r}{3}\right)$, pois, quando x pertence a esse intervalo, o menor arco determinado pelos dois pontos possui comprimento estritamente maior que $2\pi r/3$ e menor ou igual a πr . Como o intervalo A possui comprimento igual a $2\pi r/3$, conclui-se novamente que a probabilidade de a corda ser maior que o lado do triângulo equilátero inscrito é igual a $1/3$.

2. Método do ponto médio:

Escolhe-se o ponto médio da corda aleatória dentro do círculo. A corda será maior que o lado do triângulo equilátero se, e somente se, o ponto médio estiver dentro do círculo de raio $\frac{r}{2}$ (onde r é o raio do círculo original), como mostra a figura 20. Nela, perceba que a corda BC satisfaz o que se deseja, enquanto a corda DE , não. As cordas foram desenhadas paralelamente para que fosse possível fazer a representação dessa situação em uma única imagem. O triângulo também foi desenhado para facilitar a compreensão. Assim, a probabilidade corresponde à razão entre as áreas dos dois círculos:

$$P = \frac{\pi\left(\frac{r}{2}\right)^2}{\pi r^2} = \frac{1}{4}.$$

Figura 20 – Representação do evento válido para a segunda interpretação do problema.



Fonte: Elaborada pelo autor (2026).

Nesse segundo método, o experimento aleatório consiste em escolher o ponto médio da corda no interior do círculo de raio r . Dessa forma, tomamos como espaço amostral o conjunto

$$\Omega = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq r^2\},$$

isto é, o disco de raio r centrado na origem. Sobre esse conjunto, consideramos a σ -álgebra de Borel $\mathcal{F} = \mathcal{B}(\Omega)$ e definimos a medida de probabilidade P como a normalização da medida de Lebesgue bidimensional, dada por

$$P(A) = \frac{\lambda_2(A)}{\lambda_2(\Omega)}, \quad A \in \mathcal{F},$$

onde λ_2 denota a medida de Lebesgue em \mathbb{R}^2 . Nesse modelo, a corda determinada pelo ponto médio escolhido é maior que o lado do triângulo equilátero inscrito na circunferência se, e somente se, esse ponto médio pertencer ao disco de raio $\frac{r}{2}$ e centro na origem. Assim, o evento de interesse é o subconjunto

$$E = \left\{ (x, y) \in \Omega : x^2 + y^2 < \left(\frac{r}{2}\right)^2 \right\}.$$

Como a razão entre as áreas dos discos de raio $r/2$ e r é igual a $1/4$, conclui-se que a probabilidade de a corda ser maior que o lado do triângulo equilátero inscrito é igual a $\frac{1}{4}$.

Observação 3.9.3. *Um caso curioso é que, em Tijms (2012, p. 237), ele coloca o método do ponto médio como exercício. Segue a descrição do problema em tradução livre:*

“Escolhe-se aleatoriamente um ponto no interior de um círculo de raio r e constrói-se a corda (única) que tem esse ponto como seu ponto médio. Qual é a probabilidade de que essa corda seja maior do que o lado de um triângulo equilátero inscrito no círculo?” (TIJMS, 2012, p. 237)

Ao contrário do enunciado original de Bertrand, neste caso o procedimento aleatório é explicitamente definido, o que fixa a medida de probabilidade associada. Não há, portanto, ambiguidade na escolha da corda, e o problema deixa de apresentar caráter paradoxal.

3. Método do ponto extremo:

Escolhe-se aleatoriamente um ponto sobre a circunferência - no caso da figura 21 de referência, será o ponto A - traça-se o diâmetro AB e, em seguida, seleciona-se um segundo ponto (C) desse diâmetro. Desse ponto C , trace uma corda DE perpendicular ao diâmetro, como mostrado na figura 21. Perceba que, independente do ponto A escolhido, a corda DE sempre será maior do que o lado do triângulo equilátero inscrito se o ponto C escolhido pertencer ao segmento $M'M$, em que M e M' são os pontos médios de OA e OB , respectivamente. Assim, como $|M'M| = \frac{|AB|}{2}$, temos que a probabilidade será:

$$P = \frac{\frac{|AB|}{2}}{|AB|} = \frac{1}{2}.$$

Do ponto de vista formal, o método do ponto extremo pode ser descrito por meio de um espaço de probabilidade unidimensional. Inicialmente, fixa-se um ponto da circunferência, o que é lícito em virtude da invariância do problema por rotações, e considera-se o diâmetro AB por ele determinado. O segundo ponto C é então escolhido ao longo desse diâmetro. Dessa forma, o espaço amostral pode ser identificado com o segmento

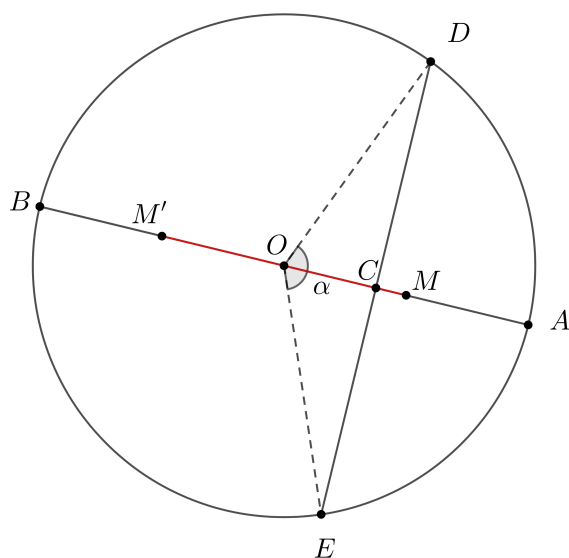
$$\Omega = AB,$$

que representa todas as posições possíveis do ponto C . Sobre Ω , considera-se a σ -álgebra de Borel $\mathcal{F} = \mathcal{B}(AB)$ e define-se a medida de probabilidade P como:

$$P(A) = \frac{\lambda_1(A)}{\lambda_1(AB)}, \quad A \in \mathcal{F},$$

onde λ_1 denota a medida de Lebesgue em \mathbb{R} . Nessa modelagem, a hipótese probabilística adotada é a de que o ponto C é escolhido aleatoriamente ao longo do diâmetro, o que completa a descrição formal do espaço de probabilidade associado ao método do ponto extremo.

Figura 21 – Representação do evento válido para a terceira interpretação do problema.



Fonte: Elaborada pelo autor (2026).

Essas três respostas, $\frac{1}{3}$, $\frac{1}{4}$ e $\frac{1}{2}$, são todas matematicamente corretas dentro de seus respectivos modelos de escolha. Portanto, não há propriamente um paradoxo, mas uma divergência de resultados oriunda da ambiguidade na definição do acaso geométrico. O paradoxo de Bertrand, nesse sentido, ilustra de maneira exemplar a importância da especificação do espaço de probabilidade e da medida associada a ele em qualquer modelo de probabilidade geométrica.

Como observado em [Johnson \(1995\)](#) e [Magalhães \(2024\)](#), o paradoxo não é uma falha da teoria, mas sim um alerta conceitual: a probabilidade de um evento depende essencialmente da forma como se define o processo de escolha aleatória. Perceba que, em experimentos clássicos de natureza discreta, como o lançamento de um dado ou de uma moeda, o espaço amostral é definido de forma clara e única: no caso do dado, por exemplo, não há outra maneira de realizar o experimento senão lançando-o e observando a face superior que resulta da rolagem. Nessas situações, a definição do espaço de probabilidade é direta, e não há ambiguidade na noção de “escolha ao acaso”.

Em contrapartida, no caso do paradoxo de Bertrand, a escolha aleatória de uma corda pode ser formalizada de maneiras distintas, e cada uma delas conduz a um espaço de probabilidade diferente. Assim, a aparente contradição entre os resultados decorre, na verdade, da multiplicidade de interpretações possíveis para o processo aleatório algo que não ocorre nos experimentos discretos usuais. Essa distinção evidencia o papel essencial da definição do espaço amostral e da medida de probabilidade na formulação rigorosa de

problemas de probabilidade geométrica.

Na seção 4 encontram-se problemas semelhantes a esse, nos quais a falta de esclarecimento sobre o processo de escolha resulta em ambiguidades na resolução.

4 Problemas e Soluções

Nesta seção, são apresentados problemas de probabilidade geométrica e em espaços enumeráveis infinitos, selecionados com o objetivo de ilustrar, de forma prática, os conceitos teóricos desenvolvidos nos capítulos anteriores. Cada problema é enunciado de maneira autônoma e, em seguida, é acompanhado de sua respectiva solução detalhada.

As soluções não se limitam à obtenção do resultado final, mas procuram explicitar o raciocínio matemático envolvido, destacando a modelagem probabilística adotada, a escolha do espaço amostral e a identificação dos eventos considerados. Sempre que pertinente, enfatiza-se a relação entre a abordagem intuitiva do problema e o formalismo matemático subjacente. Vez ou outra, serão apresentados problemas que exigirão ideias mirabolantes para serem resolvidos. Nesses casos, vale a pena rever a solução mais de uma vez.

A organização desta seção visa, portanto, cumprir um duplo papel: por um lado, consolidar os fundamentos teóricos da probabilidade geométrica por meio de exemplos resolvidos; por outro, oferecer ao leitor um material que possa ser explorado em contextos didáticos, servindo como apoio tanto para o estudo autônomo quanto para a prática em sala de aula.

Por fim, estando o leitor habituado com a linguagem adotada, denotaremos por λ_n a medida de Lebesgue em \mathbb{R}^n . Em particular, λ_1 , λ_2 e λ_3 correspondem, respectivamente, às noções usuais de comprimento, área e volume.

Problema 4.0.1. *Considere um quadrado de lado 10cm. Um dardo é lançado de modo que sempre atinja um ponto do quadrado ou de sua borda. Associe o impacto do dardo a um ponto do plano e responda:*

1. *É possível calcular a probabilidade de o dardo cair exatamente sobre um dos lados do quadrado?*
2. *Qual é o princípio geral que explica esse tipo de resultado em probabilidade geométrica?*
3. *Qual a probabilidade do dardo acertar um círculo de raio $r = 1\text{cm}$ que está concêntrico em relação ao quadrado?*

Solução:

1. *Sim, é possível calcular essa probabilidade no modelo de probabilidade geométrica, e o valor obtido é*

$$P(\text{o dado cair sobre o lado do quadrado}) = 0.$$

Embora o lado do quadrado tenha comprimento positivo, ele é um subconjunto unidimensional do plano. Ao modelarmos o lançamento do dardo como a escolha aleatória de um ponto no interior do quadrado, a probabilidade de um evento é proporcional à área do conjunto correspondente. Como um segmento de reta tem área nula no plano, a probabilidade associada a esse evento é zero.

2. O princípio geral é que, em probabilidade geométrica, eventos associados a subconjuntos de medida nula (como pontos, curvas ou superfícies sem área ou volume) têm probabilidade zero quando o espaço amostral possui dimensão maior. Assim, um ponto no plano tem probabilidade zero, uma linha no plano tem probabilidade zero, e, analogamente, uma superfície no espaço tem probabilidade zero quando comparada a um volume tridimensional.

3. A área do quadrado (espaço amostral) é $\lambda_2(\Omega) = 100$ e o evento A é $\lambda_2(A) = \pi$. Assim,

$$P(A) = \frac{\pi}{100}$$

Problema 4.0.2. escolhe-se um ponto ao acaso em um disco de raio R . Qual é a probabilidade de esse ponto estar mais próximo do centro do disco do que de sua borda?

Solução: À primeira vista, trata-se de uma questão puramente geométrica. No entanto, assim como ocorre no paradoxo de Bertrand, a resposta depende essencialmente da forma como se modela o processo de escolha aleatória, isto é, da medida de probabilidade adotada no espaço amostral.

Uma primeira interpretação natural consiste em considerar o círculo

$$D = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq R^2\}.$$

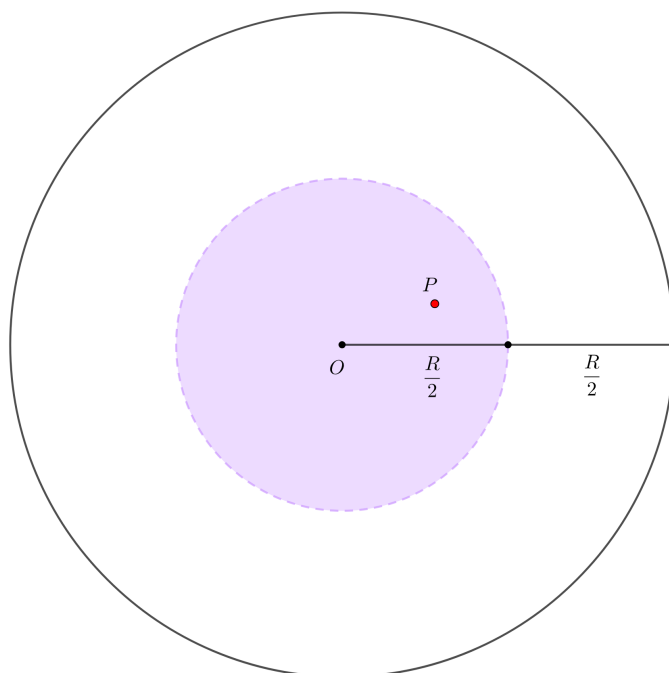
Nesse caso, o evento de interesse corresponde aos pontos cuja distância ao centro é menor que a distância até a borda do disco, o que equivale à condição $r < \frac{R}{2}$. Assim, a probabilidade procurada $P(A)$ é dada pela razão entre a área do disco de raio $\frac{R}{2}$ e a área do disco de raio R (como ilustra a figura 22).

$$P(A) = \frac{\pi(\frac{R}{2})^2}{\pi R^2} = \frac{1}{4}.$$

Considere agora um modelo probabilístico distinto para a escolha do ponto no disco. Em vez de distribuir a probabilidade uniformemente pela área do disco, adotamos o seguinte procedimento:

- escolhe-se um ângulo θ uniformemente no intervalo $[0, 2\pi)$;
- escolhe-se um raio r uniformemente no intervalo $[0, R]$;

Figura 22 – Primeira interpretação do problema, baseada na razão entre áreas.



Fonte: Elaborada pelo autor (2026).

- o ponto sorteado é aquele de coordenadas polares (r, θ) .

Nesse modelo, todos os valores do raio r entre 0 e R têm a mesma chance de serem escolhidos, independentemente da área do evento correspondente no círculo. Ou seja, a aleatoriedade é distribuída uniformemente ao longo do segmento $[0, R]$, e não sobre a área do disco.

A figura 23 ilustra o modelo de escolha aleatória baseado em coordenadas polares. O ponto P é determinado pela escolha de um ângulo θ e de um raio r , ambos independentes. O segmento OM representa a distância $\frac{R}{2}$, que separa os pontos mais próximos do centro daqueles mais próximos da borda do disco. Observa-se que, para qualquer direção fixada por θ , a verificação do evento depende apenas do valor do raio r .

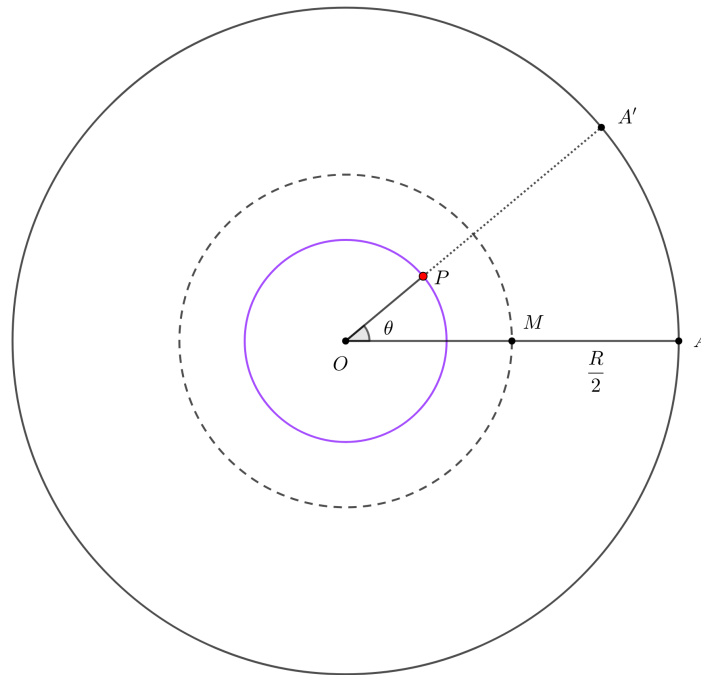
O evento de interesse ocorre quando o ponto está mais próximo do centro do disco do que de sua borda. A distância do ponto ao centro é r , enquanto a distância até a borda é $R - r$. Assim, a condição desejada é $r < \frac{R}{2}$.

Como o raio r foi escolhido uniformemente no intervalo $[0, R]$, a probabilidade desse evento $P(A)$ é dada pela razão entre o comprimento do intervalo favorável $[0, \frac{R}{2}]$ e o comprimento do intervalo total $[0, R]$:

$$P(A) = \frac{\frac{R}{2}}{R} = \frac{1}{2}.$$

Observa-se, portanto, que esse valor difere do obtido no modelo baseado na área.

Figura 23 – Segunda interpretação do problema, baseada na razão entre comprimentos.



Fonte: Elaborada pelo autor (2026).

Isso ocorre porque os dois procedimentos de escolha aleatória não são equivalentes: no primeiro, a probabilidade é distribuída uniformemente pela área do disco, enquanto, neste segundo modelo, ela é distribuída uniformemente ao longo do raio (é como se a escolha de θ não interferisse no resultado).

Problema 4.0.3. *Um número é escolhido ao acaso no intervalo $[0, 1]$. Calcule a probabilidade de que, na expansão decimal deste número:*

- 1. o primeiro algarismo (após a vírgula) seja 0.*
- 2. o segundo algarismo (após a vírgula) seja ímpar.*
- 3. apareça o algarismo 1 em algum momento.*

Solução:

- 1. Como o número é escolhido uniformemente em $[0, 1]$, cada dígito na primeira casa decimal ocorre com probabilidade $\frac{1}{10}$. Assim,*

$$P = \frac{1}{10}.$$

- 2. Assim como no item anterior, nesse caso, teremos $P = \frac{1}{2}$.*

3. Seja

$$A = \{x \in [0, 1] : \text{a expansão decimal de } x \text{ contém o algarismo } 1\}.$$

Para cada $n \in \mathbb{N}$, considere

$$A_n^c = \{x \in [0, 1] : \text{os primeiros } n \text{ dígitos não contêm o algarismo } 1\}.$$

Então,

$$\lambda_1(A_n^c) = \left(\frac{9}{10}\right)^n.$$

Como

$$A^c = \bigcap_{n=1}^{\infty} A_n^c,$$

segue, pela continuidade da medida,

$$\lambda_1(A^c) = \lim_{n \rightarrow \infty} \lambda_1(A_n^c) = \lim_{n \rightarrow \infty} \left(\frac{9}{10}\right)^n = 0.$$

Portanto,

$$\lambda_1(A) = 1 - \lambda_1(A^c) = 1.$$

Concluimos que a probabilidade de que, na expansão decimal de um número escolhido aleatoriamente em $[0, 1]$, apareça o algarismo 1 em algum momento é igual a 1.

Problema 4.0.4.

Dois números x e y são sorteados, ao acaso e independentemente, no intervalo $[0, 1]$ e calcula-se a sua soma $s = x + y$. A seguir, cada um destes números é arredondado para o inteiro mais próximo, obtendo-se os números inteiros \tilde{x} , \tilde{y} e \tilde{s} . Calcule a probabilidade de que \tilde{s} seja igual a $\tilde{x} + \tilde{y}$.

Solução: A figura correspondente ao quadrado $[0, 1]^2$ pode ser decomposta em 6 regiões disjuntas, de acordo com os valores de \tilde{x} , \tilde{y} e \tilde{s} (antes de ler a solução, é importante que o leitor tente fazer isso sozinho). A seguir, apresentamos exemplos explícitos em cada uma dessas regiões, indicando quando ocorre ou não a igualdade $\tilde{s} = \tilde{x} + \tilde{y}$.

- **Região favorável 1:** $\tilde{x} = \tilde{y} = 0$ e $\tilde{s} = 0$

$$x < \frac{1}{2}, \quad y < \frac{1}{2}, \quad x + y < \frac{1}{2}.$$

Exemplo:

$$x = 0,2, \quad y = 0,2.$$

Nesse caso,

$$\tilde{x} = 0, \quad \tilde{y} = 0, \quad s = 0,4 \Rightarrow \tilde{s} = 0,$$

logo $\tilde{s} = \tilde{x} + \tilde{y}$.

- **Região favorável 2:** $(\tilde{x}, \tilde{y}) = (1, 0)$

$$x \geq \frac{1}{2}, \quad y < \frac{1}{2}.$$

Exemplo:

$$x = 0,7, \quad y = 0,3.$$

Temos

$$\tilde{x} = 1, \quad \tilde{y} = 0, \quad s = 1,0 \Rightarrow \tilde{s} = 1,$$

logo $\tilde{s} = \tilde{x} + \tilde{y}$.

- **Região favorável 3:** $(\tilde{x}, \tilde{y}) = (0, 1)$

$$x < \frac{1}{2}, \quad y \geq \frac{1}{2}.$$

Exemplo:

$$x = 0,4, \quad y = 0,8.$$

Então

$$\tilde{x} = 0, \quad \tilde{y} = 1, \quad s = 1,2 \Rightarrow \tilde{s} = 1,$$

logo $\tilde{s} = \tilde{x} + \tilde{y}$.

- **Região favorável 4:** $\tilde{x} = \tilde{y} = 1$ e $\tilde{s} = 2$

$$x \geq \frac{1}{2}, \quad y \geq \frac{1}{2}, \quad x + y \geq \frac{3}{2}.$$

Exemplo:

$$x = 0,8, \quad y = 0,8.$$

Assim,

$$\tilde{x} = 1, \quad \tilde{y} = 1, \quad s = 1,6 \Rightarrow \tilde{s} = 2,$$

logo $\tilde{s} = \tilde{x} + \tilde{y}$.

- **Região desfavorável 1:** $\tilde{x} = \tilde{y} = 0$ e $\tilde{s} = 1$

$$x < \frac{1}{2}, \quad y < \frac{1}{2}, \quad x + y \geq \frac{1}{2}.$$

Exemplo:

$$x = 0,3, \quad y = 0,3.$$

Nesse caso,

$$\tilde{x} = 0, \quad \tilde{y} = 0, \quad s = 0,6 \Rightarrow \tilde{s} = 1,$$

logo $\tilde{s} \neq \tilde{x} + \tilde{y}$.

- **Região desfavorável 2:** $\tilde{x} = \tilde{y} = 1$ e $\tilde{s} = 1$

$$x \geq \frac{1}{2}, \quad y \geq \frac{1}{2}, \quad x + y < \frac{3}{2}.$$

Exemplo:

$$x = 0,6, \quad y = 0,6.$$

Então

$$\tilde{x} = 1, \quad \tilde{y} = 1, \quad s = 1,2 \Rightarrow \tilde{s} = 1,$$

logo $\tilde{s} \neq \tilde{x} + \tilde{y}$.

Para $x \in [0, 1]$, temos $\tilde{x} = 0$ se $x < \frac{1}{2}$ e $\tilde{x} = 1$ se $x > \frac{1}{2}$ (o caso $x = \frac{1}{2}$ tem probabilidade zero). Analogamente para y . Como $s = x + y \in [0, 2]$, tem-se: $\tilde{s} = 0$ se $s < \frac{1}{2}$, $\tilde{s} = 1$ se $\frac{1}{2} \leq s < \frac{3}{2}$ e $\tilde{s} = 2$ se $s \geq \frac{3}{2}$.

No quadrado $[0, 1]^2$, separamos por casos:

1. *Se $\tilde{x} = \tilde{y} = 0$, então $x < \frac{1}{2}$ e $y < \frac{1}{2}$, logo $\tilde{x} + \tilde{y} = 0$. Para que $\tilde{s} = 0$, precisamos $x + y < \frac{1}{2}$, que é um triângulo de área*

$$\frac{1}{2} \left(\frac{1}{2} \right) \left(\frac{1}{2} \right) = \frac{1}{8}.$$

2. *Se $(\tilde{x}, \tilde{y}) = (1, 0)$ ou $(0, 1)$, então estamos em dois retângulos de área total*

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

Neles, vale $x + y \in \left[\frac{1}{2}, \frac{3}{2} \right)$ (fora de bordas de medida zero), logo $\tilde{s} = 1 = \tilde{x} + \tilde{y}$ sempre.

3. *Se $\tilde{x} = \tilde{y} = 1$, então $x \geq \frac{1}{2}$ e $y \geq \frac{1}{2}$, logo $\tilde{x} + \tilde{y} = 2$. Para que $\tilde{s} = 2$, precisamos $x + y \geq \frac{3}{2}$, que dentro de $\left[\frac{1}{2}, 1 \right]^2$ forma um triângulo de área*

$$\frac{1}{2} \left(\frac{1}{2} \right) \left(\frac{1}{2} \right) = \frac{1}{8}.$$

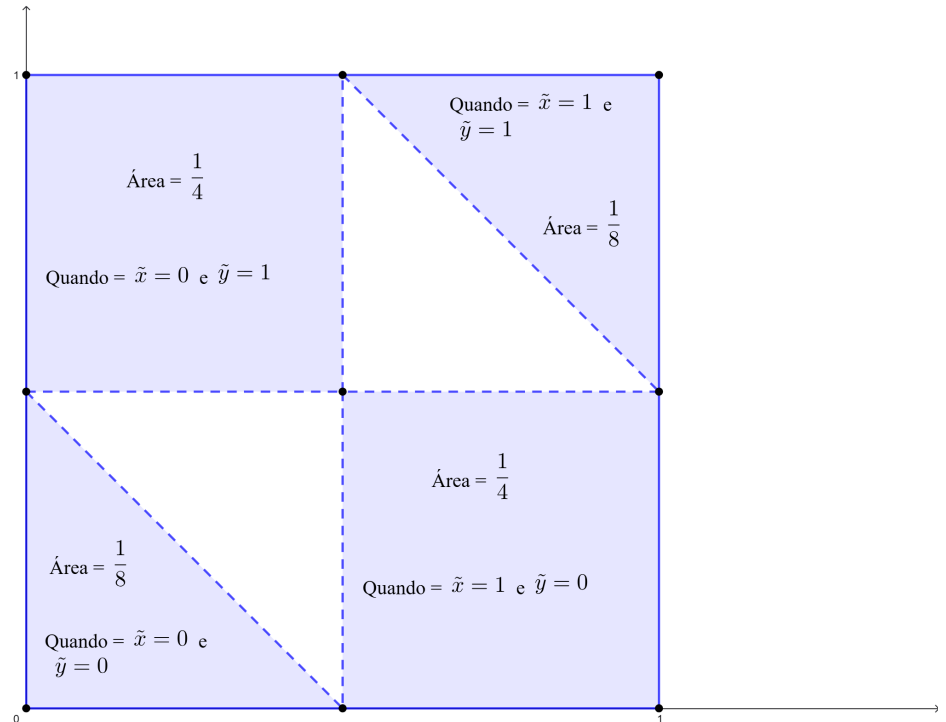
Na figura 24 explicitamos geometricamente o espaço amostral e os eventos.

Somando as áreas favoráveis, a probabilidade P é

$$P = \frac{1}{8} + \frac{1}{2} + \frac{1}{8} = \frac{3}{4}.$$

Os casos em que $x = \frac{1}{2}$ ou $y = \frac{1}{2}$ podem ser ignorados, pois formam subconjuntos de medida bidimensional nula em $[0, 1]^2$. Assim, sua contribuição para a probabilidade é zero. Para evitar problemas na descrição do problema, podemos ignorar quando $s = \frac{1}{2}$, para evitar explicações desnecessárias, já que ela também não altera em nada a probabilidade.

Figura 24 – Espaço amostral do problema em destaque na cor azul. É possível, inclusive, observar onde se encontram as seis regiões exemplificadas inicialmente.



Fonte: Elaborada pelo autor (2026).

Problema 4.0.5 (Olimpíada Paraibana de Matemática - 2025). *Seja R uma região do plano que representa um conjunto de resultados possíveis. Neste contexto de probabilidade geométrica, consideramos que todos os pontos são equiprováveis, ou seja, possuem resultados igualmente possíveis. Se $A \subset R$ é um subconjunto que representa um evento desejado, então a probabilidade desse evento ocorrer é dada por*

$$P(A) = \frac{\text{área de } A}{\text{área de } R}.$$

Um ponto é escolhido ao acaso no quadrado

$$Q = [0, 4] \times [0, 4]$$

do plano cartesiano xy . Qual a probabilidade de que o ponto escolhido satisfaça a condição

$$|x + y - 3| \leq |x - 1|?$$

Solução: *Escolhe-se um ponto (x, y) ao acaso no quadrado*

$$\Omega = Q = [0, 4] \times [0, 4],$$

em que

$$\lambda_2(\Omega) = 16.$$

Deseja-se calcular a probabilidade do evento

$$A = \{(x, y) \in \Omega : |x + y - 3| \leq |x - 1|\}.$$

Por definição de probabilidade geométrica,

$$P(A) = \frac{\lambda_2(A)}{\lambda_2(\Omega)}.$$

Para isso, analisemos a desigualdade. Distinguímos dois casos conforme o sinal de $x - 1$.

- **Caso 1:** $0 \leq x < 1$. Temos $|x - 1| = 1 - x$, logo

$$-(1 - x) \leq x + y - 3 \leq 1 - x.$$

Somando $3 - x$ em todos os termos,

$$2 \leq y \leq 4 - 2x.$$

- **Caso 2:** $1 \leq x \leq 4$. Temos $|x - 1| = x - 1$, e

$$-(x - 1) \leq x + y - 3 \leq x - 1.$$

Somando $3 - x$,

$$4 - 2x \leq y \leq 2.$$

2. Interpretação geométrica. O evento A é a união disjunta das seguintes regiões:

um triângulo com vértices $A(0, 2)$, $B(0, 4)$, $C(1, 2)$; um triângulo com vértices $C(1, 2)$, $D(2, 0)$, $E(2, 2)$; um retângulo com vértices $D(2, 0)$, $E(2, 2)$, $F(4, 2)$, $G(4, 0)$, como mostra a figura 25.

3. Cálculo das áreas.

- Área do triângulo ABC :

$$\lambda_2(ABC) = \frac{1 \cdot 2}{2} = 1.$$

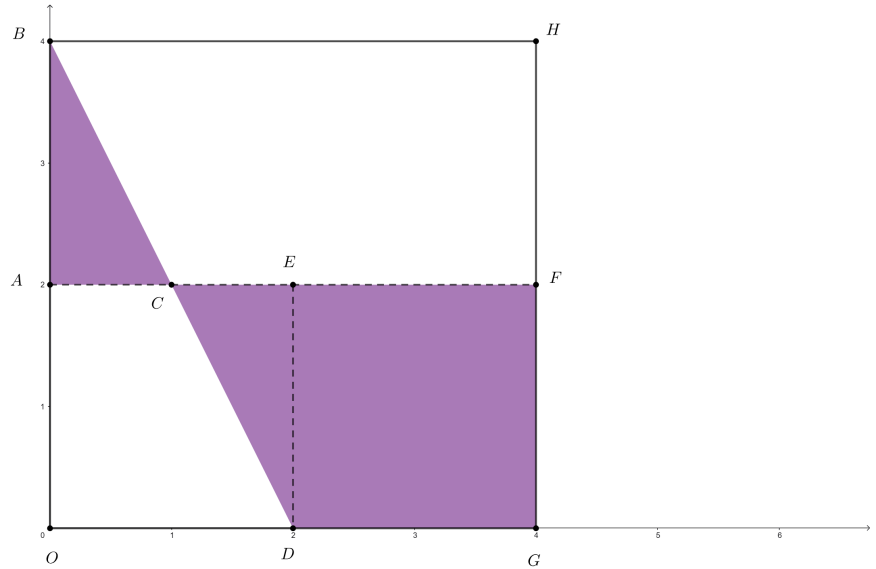
- Área do triângulo CDE :

$$\lambda_2(CDE) = \frac{1 \cdot 2}{2} = 1.$$

- Área do retângulo $DEFG$:

$$\lambda_2(DEFG) = 2 \cdot 2 = 4.$$

Figura 25 – Espaço amostral e evento do problema.



Fonte: Elaborada pelo autor (2026).

Assim,

$$\lambda_2(A) = 1 + 1 + 4 = 6.$$

4. Probabilidade. Portanto,

$$P(A) = \frac{\lambda_2(A)}{\lambda_2(\Omega)} = \frac{6}{16} = \frac{3}{8}.$$

$$P(A) = \frac{3}{8}.$$

Problema 4.0.6. *Cristina e Maria, que são pessoas muito pontuais, marcaram um encontro às 16h00. Se cada uma delas chegará ao encontro em um instante qualquer entre 16 e 17 horas e se dispõe a esperar no máximo 10 minutos pela outra, qual é a probabilidade delas se encontrarem?*

Antes da solução, sugerimos que tente resolvê-lo apenas tendo como parâmetro o problema resolvido na seção 3.7.1.

Solução: A resolução deste problema baseia-se em uma modelagem por meio da geometria analítica. Sejam x e y os instantes (medidos em horas) em que Cristina e Maria chegam ao local do encontro, respectivamente. Como ambas podem chegar em qualquer instante entre 16h00 e 17h00, podemos supor, sem perda de generalidade, que

$$x, y \in [0, 1],$$

onde o instante 0 corresponde a 16h00 e o instante 1 corresponde a 17h00.

Dessa forma, o espaço amostral é dado por

$$\Omega = [0, 1] \times [0, 1],$$

cuja medida é

$$\lambda_2(\Omega) = 1.$$

As duas se encontrarão se, e somente se, a diferença entre os horários de chegada for de no máximo 10 minutos, isto é,

$$|x - y| \leq \frac{1}{6}.$$

Denotemos por

$$A = \{(x, y) \in \Omega : |x - y| \leq \frac{1}{6}\}$$

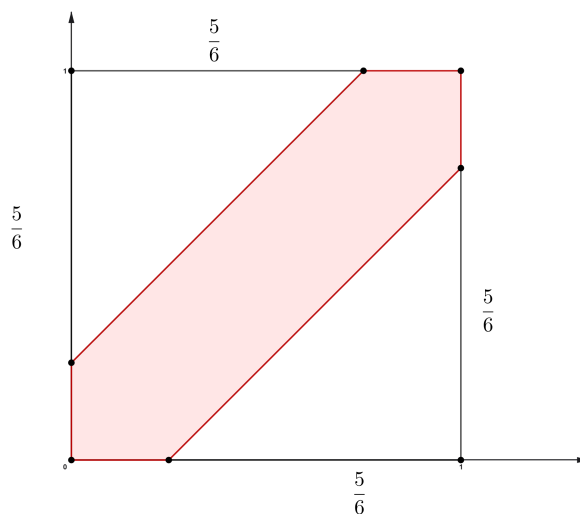
o evento correspondente ao encontro entre Cristina e Maria.

Geometricamente, o conjunto A corresponde à região r compreendida entre as retas

$$y = x + \frac{1}{6} \quad e \quad y = x - \frac{1}{6}.$$

A região complementar A^c é formada por dois triângulos retângulos congruentes, cada um com catetos de comprimento $\frac{5}{6}$, como mostra a figura 26. Assim,

Figura 26 – Espaço amostral e o evento do problema apresentados geometricamente na figura.



Fonte: Elaborada pelo autor (2026).

$$\lambda_2(A^c) = 2 \cdot \frac{1}{2} \left(\frac{5}{6}\right)^2 = \frac{25}{36}.$$

Logo,

$$\lambda_2(A) = \lambda_2(\Omega) - \lambda_2(A^c) = 1 - \frac{25}{36} = \frac{11}{36}.$$

Portanto, a probabilidade de Cristina e Maria se encontrarem é

$$\frac{\lambda_2(A)}{\lambda_2(\Omega)} = \frac{11}{36}.$$

Problema 4.0.7. (OMU - 2025 - 2ª Fase) Um espagete cru caiu e, com o impacto, quebrou-se simultaneamente em alguns pedaços de maneira aleatória. Responda:

1. Se o espagete se quebrou em exatamente três pedaços, qual a probabilidade de as medidas destes pedaços serem as medidas de um triângulo?

Solução: Este é o problema clássico do espagete, basta consultar a resposta do problema na seção 3.7.2.

2. Se o espagete se quebrou em quatro pedaços, então:

- a) Determine as condições (algébricas) que descrevem todas as configurações possíveis para os tamanhos dos quatro pedaços e, usando o GeoGebra, grafique a região do espaço delimitada por essas condições.

Solução: Sejam $x, y, z, w \in \mathbb{R}_+^*$ e $L = x + y + z + w$. Definamos os pontos de corte do espagete com extremidades em $(0, 0)$ e $(L, 0)$ pelas abscissas

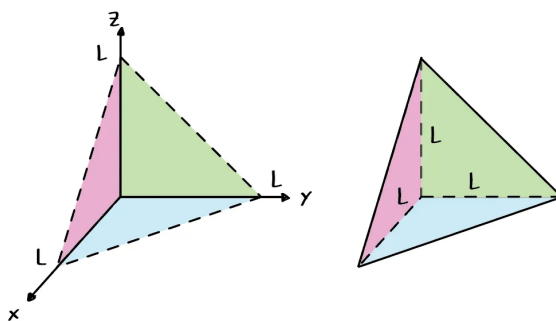
$$x, \quad x + y, \quad x + y + z.$$

Efetuada tais cortes, obtêm-se quatro pedaços de comprimentos x, y, z, w . As condições

$$x > 0, \quad y > 0, \quad z > 0, \quad w > 0, \quad x + y + z + w = L$$

representam geometricamente um tetraedro no espaço \mathbb{R}^3 (o domínio de escolha de (x, y, z) com $w = L - (x + y + z) > 0$). A representação geométrica disso está na figura 27.

Figura 27 – Representação geométrica do tetraedro do problema.



Fonte: Elaborada pelo autor (2026).

O volume do espaço amostral é portanto

$$\Omega = \lambda_3 \left\{ (x, y, z) : x > 0, \quad y > 0, \quad z > 0, \quad x + y + z < L \right\} = \frac{L^3}{6}.$$

b) Determine as condições (algébricas) que quatro números devem satisfazer para que sejam medidas de um quadrilátero.

Solução: Esse é um resultado clássico da geometria.

Teorema: Dados quatro segmentos de comprimentos positivos, eles formam um quadrilátero (simples) se, e somente se, a soma dos comprimentos de quaisquer três deles for maior do que o comprimento do quarto.

Demonstração: (Necessidade \implies). Seja $ABCD$ um quadrilátero com

$$AB = a, \quad BC = b, \quad CD = c, \quad DA = d.$$

Considere a diagonal BD de comprimento m . Nos triângulos ABD e CBD aplica-se a desigualdade triangular:

$$a + b > m, \quad m + c > d.$$

Da segunda desigualdade segue $m > d - c$; substituindo em $a + b > m$ obtemos $a + b > d - c$, isto é $a + b + c > d$. De maneira análoga obtêm-se as demais desigualdades

$$a + b + d > c, \quad a + c + d > b, \quad b + c + d > a,$$

o que prova a necessidade.

(Suficiência \impliedby). Suponha, sem perda de generalidade, que

$$a \geq b \geq c \geq d > 0$$

e que as quatro desigualdades do enunciado são satisfeitas. Seja $r = b + d'$, com d' um número positivo suficientemente próximo de d e menor que d , de modo que ainda valha $r + c > a$. Então a, r, c satisfazem as desigualdades triangulares e formam um triângulo. Analogamente, os comprimentos r, b, d formam outro triângulo. Tomando a união desses dois triângulos que compartilham o lado de comprimento r , obtemos um quadrilátero simples cujos lados, em ordem, são a, b, c, d . Isso conclui a prova da suficiência.

Tendo provado isso, obtemos a resposta do item em questão.

c) Utilizando (i) e (ii), determine as condições (algébricas) que devem ser satisfeitas para que os tamanhos destes quatro pedaços formem um quadrilátero e, usando o GeoGebra, grafique a região do espaço delimitada por essas condições.

Solução: Pelo item b-ii, as condições para que x, y, z, w formem um quadrilátero são

$$x < y + z + w,$$

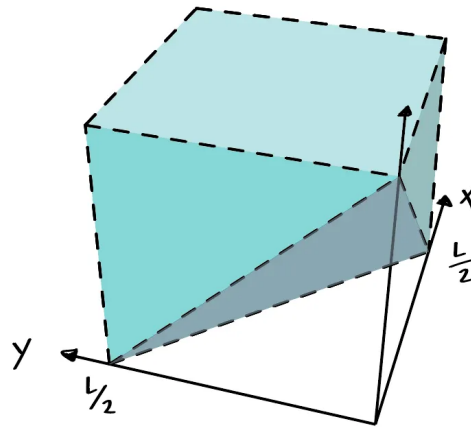
$$y < x + z + w,$$

$$z < x + y + w,$$

$$w < x + y + z.$$

Substituindo $w = L - (x + y + z)$ obtém-se as condições $x, y, z < \frac{L}{2}$ e $x + y + z > \frac{L}{2}$. Juntas, elas representam o volume delimitado pelo cubo de aresta $L/2$ e o plano $z = \frac{L}{2} - x - y$, como ilustrado na figura 28.

Figura 28 – Representação geométrica do sólido geométrico tridimensional resultante do corte de um cubo por um plano de acordo com o item da questão.



Fonte: Elaborada pelo autor (2026).

Tomando apenas a porção que representa comprimentos favoráveis geometricamente, tais condições descrevem a região do cubo de aresta $L/2$ limitada pelos planos $x + y + z = \frac{L}{2}$ e $x + y + z = L$ (A região pode ser vista como a interseção entre o primeiro tetraedro de vértices $(L, 0, 0)$, $(0, L, 0)$, $(0, 0, L)$, $(0, 0, 0)$ e a região do cubo descrita anteriormente.) A representação geométrica é dada na figura 29.

- d) Calcule a probabilidade de os tamanhos dos pedaços serem as medidas de um quadrilátero.

Solução: Para calcular o volume desejado (região descrita em b-iii), observe que o cubo de aresta $L/2$ tem volume $\left(\frac{L}{2}\right)^3 = \frac{L^3}{8}$. Da porção do cubo que excede o tetraedro mencionado, podemos retirar duas pirâmides retas congruentes cujas bases têm área $\frac{L^2}{8}$ e alturas $\frac{L}{2}$. O volume de cada pirâmide é

$$\frac{\text{base} \times \text{altura}}{3} = \frac{\frac{L^2}{8} \cdot \frac{L}{2}}{3} = \frac{L^3}{48}.$$

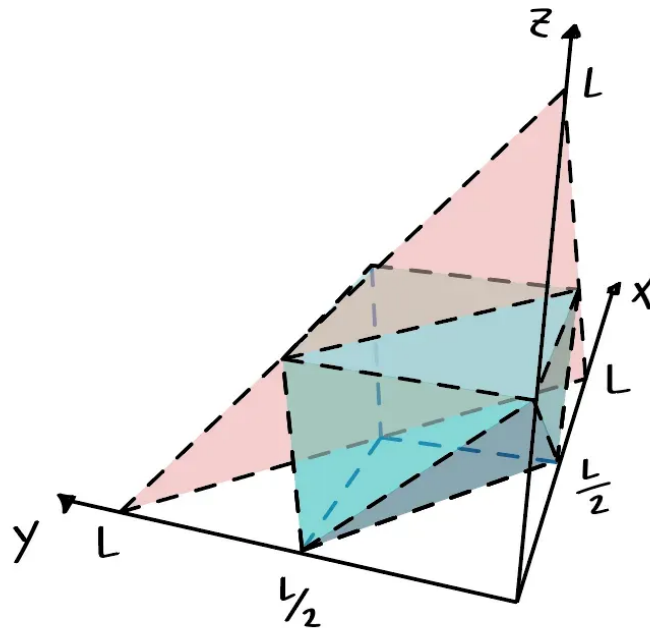
Retirando as duas pirâmides do cubo obtemos

$$V = \frac{L^3}{8} - 2 \cdot \frac{L^3}{48} = \frac{L^3}{8} - \frac{L^3}{24} = \frac{3L^3 - L^3}{24} = \frac{2L^3}{24} = \frac{L^3}{12}.$$

Pelo item b-i, lembrando $\Omega = \frac{L^3}{6}$, tem-se

$$P(V) = \frac{\lambda_3(V)}{\lambda_3(\Omega)} = \frac{\frac{L^3}{12}}{\frac{L^3}{6}} = \frac{1/12}{1/6} = \frac{1}{2}.$$

Figura 29 – Representação gráfica solicitada pelo problema.



Fonte: Elaborada pelo autor (2026).

Problema 4.0.8. (OMPD 2022) Três amigos internautas combinaram de participar de uma call no Discord para discutir a OMPD 2022. Eles combinaram de entrar no Discord em algum horário entre as 0h e 1h da madrugada, mas cada um deles vai entrar na call em algum horário aleatório, vai esperar uma fração x de hora e, se nenhum dos outros amigos estiver online durante esse tempo, ele sairá do aplicativo. Sabendo que a probabilidade de que cada um dos amigos não veja nenhum dos outros dois amigos online é de $\frac{1}{2}$, quanto vale x ?

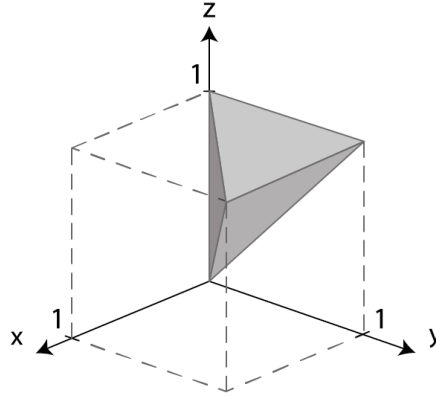
Solução: O método para solucionar esse problema é bem criativo. Segue.

Sejam T_1, T_2, T_3 horários escolhidos ao acaso e de forma independente no intervalo $[0, 1]$ (medindo horas a partir de 0h). Ordenando-os, obtemos: $0 \leq T_1 \leq T_2 \leq T_3 \leq 1$. Daí, percebe-se que o espaço amostral Ω será a interseção do produto cartesiano $[0, 1]^3$ com a condição anterior. Nesse caso, a figura obtida será um tetraedro, como mostrado na figura 30, cujo volume é $\lambda_3(\Omega) = \frac{1}{6}$.

Como queremos que eles não se encontrem, basta perceber que os intervalos $[T_1, T_1 + x]$ e $[T_2, T_2 + x]$ não podem se sobrepor, isto é, $T_2 \notin [T_1, T_1 + x)$ e $T_3 \notin [T_2, T_2 + x)$ (perceba que deixar o intervalo aberto ou fechado não interfere, pois o extremo do intervalo tem volume nulo. Faremos isso para deixar a notação a seguir melhor). Com isso, concluímos que $T_3 - T_2 \geq x$ e $T_2 - T_1 \geq x$.

A ideia curiosa desse problema vem no próximo passo. Ao invés de ficar buscando

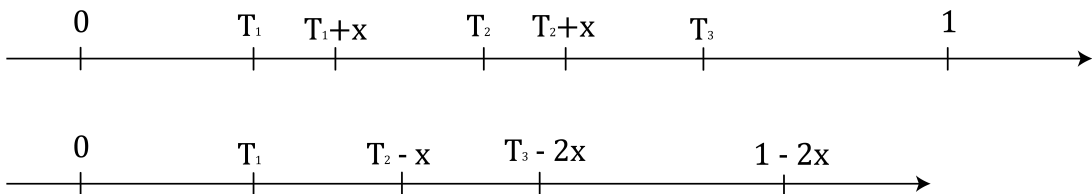
Figura 30 – Tetraedro formado pela condição imposta pelo problema.



Fonte: Elaborada pelo autor (2026).

maneiras de calcular o evento nessas condições, vamos modificar o comprimento do nosso intervalo. Perceba que a distância x deve existir, pois, caso contrário, dois deles sempre se encontrariam. Queremos que eles nunca se encontrem, dois a dois. Por isso, basta retirar o comprimento x duas vezes do intervalo $[0, 1]$ pois, com isso, qualquer ponto escolhido tal que $0 \leq T_1 \leq T_2 - x \leq T_3 - 2x \leq 1 - 2x$ satisfará as condições propostas. Essa mudança no comprimento dos segmentos está representada na figura 31.

Figura 31 – A mudança de variáveis corresponde a uma translação dos instantes de entrada, que elimina os intervalos de segurança de comprimento x entre eles, reduzindo o intervalo original $[0, 1]$ ao intervalo efetivo $[0, 1 - 2x]$.



Fonte: Elaborada pelo autor (2026).

No entanto, como nesse passo há uma mudança de variáveis (pois o T_2 se tornará outro valor nesse novo segmento), como mostra a figura 31 (perceba que o T_2 torna-se $T_2 - x$ agora). façamos a mudança de variável $t_1 = T_1$, $t_2 = T_2 - x$ e $t_3 = T_3 - 2x$. Observe que a relação inversa é $T_1 = t_1$, $T_2 = t_2 + x$, $T_3 = t_3 + 2x$, logo a mudança é uma bijeção. Essa mudança de variável serve apenas para reescrever o problema em um novo sistema de coordenadas, simplificando as desigualdades e evitando expressões como $T_3 - 2x$.

Perceba que, após a mudança de variáveis, o problema se reduz a considerar os

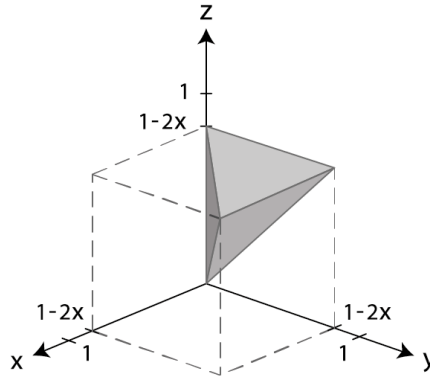
pontos $(t_1, t_2, t_3) \in [0, 1 - 2x]^3$ que satisfazem a condição de ordenação

$$0 \leq t_1 \leq t_2 \leq t_3 \leq 1 - 2x.$$

Geometricamente, esse conjunto corresponde à região do cubo $[0, 1 - 2x]^3$ em que as coordenadas estão ordenadas. Essa região forma um tetraedro, análogo àquele obtido no caso inicial sobre o cubo $[0, 1]^3$.

Em outras palavras, estamos apenas “encolhendo” o cubo original $[0, 1]^3$ para $[0, 1 - 2x]^3$ e, dentro dele, considerando novamente apenas a região ordenada, como mostra a figura 32. Daí, o volume do novo tetraedro será calculado de forma semelhante.

Figura 32 – Novo tetraedro após fazer as modificações necessárias.



Fonte: Elaborada pelo autor (2026).

Assim, o evento A , correspondente aos casos em que nenhum dos amigos se encontra, é representado por um tetraedro de volume

$$\lambda_3(A) = \frac{(1 - 2x)^3}{6}.$$

Sabe-se que a probabilidade de nenhum deles se encontrar é

$$P(A) = \frac{\lambda_3(A)}{\lambda_3(\Omega)} = \frac{\frac{(1-2x)^3}{6}}{\frac{1}{6}} = (1 - 2x)^3.$$

Como $P(A) = \frac{1}{2}$, segue que

$$(1 - 2x)^3 = \frac{1}{2}.$$

Resolvendo essa equação obtemos $x = \frac{2 - \sqrt[3]{4}}{4}$.

Observação 4.0.9. *A seguir, consideramos apenas problemas cujo espaço amostral é infinito enumerável.*

Problema 4.0.10. *Dois jogadores lançam sucessivamente uma moeda honesta. Ganha o primeiro que obtiver um resultado igual ao obtido na jogada anterior. Qual é a probabilidade de que o primeiro a lançar a moeda seja o ganhador?*

Solução: *Considere dois jogadores, A e B, que lançam sucessivamente uma moeda honesta, alternando as jogadas, sendo A o primeiro a lançar. Ganha o jogador que, em sua jogada, obtiver o mesmo resultado da jogada imediatamente anterior. Observemos inicialmente que não é possível haver vencedor na primeira jogada, pois não há jogada anterior para comparar. A partir da segunda jogada, em cada lançamento, o jogador vence se o resultado coincidir com o da jogada anterior.*

Vitórias do jogador A: *O jogador A realiza as jogadas de número ímpar. Para que A vença na jogada $2i + 1$ (como $i \in \mathbb{N}$, ele só pode ganhar na terceira, quinta, sétima, ..., jogadas, já que a primeira não conta), é necessário que:*

- *não tenha ocorrido dois resultados iguais nas $2i$ primeiras jogadas;*
- *o resultado da jogada $2i + 1$ seja igual ao da jogada $2i$.*

Como a moeda é honesta, a probabilidade de não haver duas jogadas consecutivas iguais nas primeiras $2i$ jogadas é igual a $(\frac{1}{2})^{2i-1}$, pois a comparação só começa a partir da segunda jogada. Logo, a probabilidade de A vencer na jogada $2i + 1$ é

$$\left(\frac{1}{2}\right)^{2i-1} \cdot \frac{1}{2} = \left(\frac{1}{2}\right)^{2i} = \left(\frac{1}{4}\right)^i.$$

Assim, pode-se calcular a probabilidade de A ganhar como

$$\frac{1}{4} + \left(\frac{1}{4}\right)^2 + \left(\frac{1}{4}\right)^3 + \cdots + \left(\frac{1}{4}\right)^i + \cdots = \sum_{i=1}^{\infty} \left(\frac{1}{4}\right)^i$$

Como se trata de uma progressão geométrica de razão $\frac{1}{4}$, temos que

$$\sum_{i=1}^{\infty} \left(\frac{1}{4}\right)^i = \frac{\frac{1}{4}}{1 - \frac{1}{4}} = \frac{1}{3}$$

Portanto, a probabilidade de A vencer é dada por

$$P(A \text{ vence}) = \frac{1}{3}.$$

Vitórias do jogador B: *O jogador B realiza as jogadas de número par. Para que B vença na jogada $2i$ (com $i \in \mathbb{N}$), isto é, na segunda, quarta, sexta, ... jogadas, é necessário que:*

- não tenha ocorrido dois resultados iguais nas $2i - 1$ primeiras jogadas;
- o resultado da jogada $2i$ seja igual ao da jogada $2i - 1$.

Como a primeira jogada não possui resultado anterior para comparação, apenas as $2i - 2$ jogadas seguintes precisam ser diferentes da jogada imediatamente anterior. Como a moeda é honesta, a probabilidade de uma jogada diferir da anterior é $\frac{1}{2}$. Logo, a probabilidade de não ocorrer coincidência nas primeiras $2i - 1$ jogadas é

$$\left(\frac{1}{2}\right)^{2i-2}.$$

Multiplicando pela probabilidade de coincidência na jogada $2i$, obtemos que a probabilidade de B vencer nessa jogada é

$$\left(\frac{1}{2}\right)^{2i-2} \cdot \frac{1}{2} = \left(\frac{1}{2}\right)^{2i-1}.$$

Assim, pode-se calcular a probabilidade de B ganhar como

$$\frac{1}{2} + \left(\frac{1}{2}\right)^3 + \left(\frac{1}{2}\right)^5 + \dots = \sum_{i=1}^{\infty} \left(\frac{1}{2}\right)^{2i-1} = 2 \cdot \sum_{i=1}^{\infty} \left(\frac{1}{2}\right)^{2i} = 2 \cdot \sum_{i=1}^{\infty} \left(\frac{1}{4}\right)^i.$$

Como se trata de uma progressão geométrica de razão $\frac{1}{4}$, assim como na situação do jogador A , temos que

$$P(B \text{ vence}) = 2 \cdot \frac{1}{3} = \frac{2}{3}.$$

Além disso, a probabilidade deles jogarem, infinitamente, sem nenhum deles ganhar é dada por $\lim_{i \rightarrow \infty} \left(\frac{1}{2}\right)^i = 0$.

Problema 4.0.11. *Arlindo tem um real e Bernardo tem dois. Eles combinam de apostar um real em sucessivas disputas de cara ou coroa, até que um deles perca todo o seu dinheiro. Qual é a probabilidade de que Arlindo seja o ganhador?*

Solução: Representemos o estado do jogo por um par ordenado (a, b) , em que a denota a quantia de dinheiro de Arlindo e b a de Bernardo. O estado inicial é $(1, 2)$, e o jogo termina quando se atinge um dos pares $(3, 0)$ (vitória de Arlindo) ou $(0, 3)$ (vitória de Bernardo). Para que Arlindo seja o vencedor, o processo deve eventualmente atingir o estado $(3, 0)$. Observemos que isso ocorre quando o estado $(2, 1)$ é seguido por uma vitória de Arlindo. Assim, as sequências possíveis que levam à vitória de Arlindo são da forma:

$$(1, 2) \rightarrow (2, 1) \rightarrow (3, 0),$$

ou

$$(1, 2) \rightarrow (2, 1) \rightarrow (1, 2) \rightarrow (2, 1) \rightarrow (3, 0),$$

ou ainda

$$(1, 2) \rightarrow (2, 1) \rightarrow (1, 2) \rightarrow (2, 1) \rightarrow (1, 2) \rightarrow (2, 1) \rightarrow (3, 0),$$

e assim sucessivamente.

Em cada disputa, a probabilidade de vitória de qualquer jogador é $\frac{1}{2}$. Logo, a probabilidade de Arlindo vencer na primeira sequência é $\left(\frac{1}{2}\right)^2$, na segunda é $\left(\frac{1}{2}\right)^4$, na terceira é $\left(\frac{1}{2}\right)^6$, e assim por diante. Portanto, a probabilidade de vitória de Arlindo é dada pela soma

$$P(\text{Arlindo vence}) = \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^4 + \left(\frac{1}{2}\right)^6 + \dots = \sum_{k=1}^{\infty} \left(\frac{1}{2}\right)^{2k}.$$

Trata-se de uma progressão geométrica de razão $\frac{1}{4}$, cuja soma é

$$P(\text{Arlindo vence}) = \frac{\frac{1}{4}}{1 - \frac{1}{4}} = \frac{1}{3}.$$

Problema 4.0.12. *A e B lançam sucessivamente um par de dados não enviesados até que um deles obtenha soma de pontos 7, caso em que a disputa termina e o vencedor é o jogador que obteve soma 7. Se A é o primeiro a jogar, qual a probabilidade de A ser o vencedor?*

Solução: Começando por A, ou A ganha na primeira rodada, ou na terceira, ou na quinta, e assim sucessivamente. Assim, a probabilidade de A ganhar é a soma das probabilidades correspondentes a cada uma dessas situações.

Para que A vença na i -ésima jogada, é necessário que nenhuma das $i - 1$ jogadas anteriores resulte em soma 7 e que a i -ésima jogada resulte em soma 7. Como a probabilidade de não sair 7 em uma jogada é $\frac{5}{6}$ e a de sair 7 é $\frac{1}{6}$, a probabilidade de A vencer na jogada $2i - 1$ é

$$P = \left(\frac{5}{6}\right)^{2(i-1)} \cdot \frac{1}{6}.$$

Logo,

$$P(A) = \sum_{i=1}^{\infty} \left(\frac{5}{6}\right)^{2(i-1)} \frac{1}{6}.$$

Fatorando:

$$P(A) = \frac{1}{6} \sum_{k=0}^{\infty} \left(\frac{25}{36}\right)^k.$$

Soma geométrica:

$$P(A) = \frac{1}{6} \cdot \frac{1}{1 - \frac{25}{36}} = \frac{1}{6} \cdot \frac{36}{11} = \frac{6}{11}.$$

De modo análogo, a probabilidade de B ganhar, utilizando o mesmo raciocínio, é

$$P(B) = \sum_{i=1}^{\infty} \left(\frac{5}{6}\right)^{2i-1} \frac{1}{6} = \frac{1}{6} \cdot \sum_{i=1}^{\infty} \left(\frac{5}{6}\right)^{2i} \cdot \frac{6}{5} = \frac{1}{5} \cdot \sum_{i=1}^{\infty} \left(\frac{25}{36}\right)^i$$

que também é uma progressão geométrica de razão $\frac{25}{36}$, com termo inicial sendo $\frac{25}{36}$. Calculando a soma pela fórmula da soma dos termos da progressão geométrica, obtemos $\sum_{i=1}^{\infty} \left(\frac{25}{36}\right)^i = \frac{\frac{25}{36}}{1 - \frac{25}{36}} = \frac{25}{11}$. Daí, concluímos que

$$P(B) = \frac{1}{5} \cdot \frac{25}{11} = \frac{5}{11}.$$

Observação. A probabilidade de que nenhum dos jogadores obtenha soma 7 após $2x$ jogadas é $\left(\frac{5}{6}\right)^{2x}$, a qual tende a zero quando $x \rightarrow \infty$. Assim, a probabilidade de que o jogo nunca termine tende a zero.

Referências

- ARAÚJO, E. A. d. *Probabilidade Geométrica no Ensino Médio: Uma Experiência Usando o Geogebra*. Dissertação (Dissertação (Mestrado Profissional em Matemática)) — Universidade Estadual da Paraíba, Campina Grande, 2017. PROFMAT – Centro de Ciências e Tecnologia. Citado na página [30](#).
- BANACH, S.; TARSKI, A. Sur la décomposition des ensembles de points en parties respectivement congruentes. *Fundamenta Mathematicae*, v. 6, p. 244–277, 1924. Citado na página [55](#).
- BOYER, C. B.; MERZBACH, U. C. História da matemática. São Paulo: Blucher, 2012. Citado 3 vezes nas páginas [6](#), [7](#) e [62](#).
- FERNANDEZ, P. J. *Medida e Integração*. 2. ed. Rio de Janeiro: Associação Instituto Nacional de Matemática Pura e Aplicada, 2002. Projeto Euclides. Citado na página [36](#).
- FOLLAND, G. B. *Real Analysis: Modern Techniques and Their Applications*. 2. ed. New York: John Wiley & Sons, 1999. Citado na página [41](#).
- GRINSTEAD, C. M.; SNELL, J. L. Introduction to probability. 2. ed. *Providence: American Mathematical Society*, 2006. Citado na página [59](#).
- HAZZAN, S. Fundamentos de matemática elementar: combinatória e probabilidade. São Paulo: Atual, 2013. Citado na página [7](#).
- JAMES, B. R. Probabilidade: um curso em nível intermediário. Instituto de Matemática Pura e Aplicada - IMPA, 2010. Citado 8 vezes nas páginas [15](#), [16](#), [17](#), [23](#), [33](#), [51](#), [52](#) e [53](#).
- JOHNSON, A. Geometric probability. *Lexington, MA: COMAP, Inc.*, 1995. Citado 4 vezes nas páginas [6](#), [30](#), [60](#) e [72](#).
- LIMA, E. L. Curso de análise vol. 1. *Rio de Janeiro: Instituto Nacional de Matemática Pura e Aplicada*, 2013. Citado na página [13](#).
- MAGALHÃES, M. N. Probabilidade e variáveis aleatórias. *São Paulo: Editora da Universidade de São Paulo*, 2024. Citado 3 vezes nas páginas [15](#), [22](#) e [72](#).
- MARGINEAN, D. The lebesgue measure on \mathbb{R}^n . *Procedia Technology*, v. 19, p. 789–792, 2015. Citado na página [41](#).
- MORGADO, A. C.; CARVALHO, P. C. P. Matemática discreta. Rio de Janeiro: SBM, 2015. Citado na página [57](#).
- ROLLA, L. T.; LIMA, B. N. B. d. *Probabilidade*. [S.l.: s.n.], 2025. Citado 2 vezes nas páginas [35](#) e [38](#).
- ROYDEN, H. L.; FITZPATRICK, P. M. *Real Analysis*. 4. ed. Boston: Pearson Education, 2010. Citado 5 vezes nas páginas [36](#), [38](#), [41](#), [42](#) e [45](#).
- SOLOMON, H. Geometric probability. Philadelphia: Society for Industrial and Applied Mathematics (SIAM), 1978. Citado na página [62](#).

STONE, H. P. Introdução à teoria da probabilidade. *Rio de Janeiro: Editora Interciência Ltda*, 1978. Citado na página [14](#).

TIJMS, H. C. *Understanding Probability*. 3. ed. Cambridge: Cambridge University Press, 2012. Citado 2 vezes nas páginas [70](#) e [71](#).

TOMKOWICZ, G.; WAGON, S. *The Banach–Tarski Paradox*. 2. ed. [S.l.]: Cambridge University Press, 2016. v. 163. (Encyclopedia of Mathematics and its Applications, v. 163). Citado na página [55](#).

TUNALA, N. Determinação de probabilidades por método geométrico. *Revista do Professor de Matemática (RPM)*, v. 20, p. 16–22, 1992. Citado na página [55](#).

WAGNER, E. Probabilidade geométrica. *REVISTA DO PROFESSOR DE MATEMÁTICA (RPM)*, s.d. Citado na página [57](#).

ANEXO A – Demonstrações dos Teoremas

A.1 O Teorema de Vitali

Definição A.1.1 (Relação de equivalência por diferença racional). *No intervalo $[0, 1]$ definimos a relação \sim por*

$$x \sim y \iff x - y \in \mathbb{Q}.$$

É imediato verificar que \sim é uma relação de equivalência em $[0, 1]$.

Definição A.1.2 (Conjunto de Vitali). *Um conjunto de Vitali $V \subset [0, 1]$ é um conjunto que contém exatamente um representante de cada classe de equivalência de $[0, 1]$ para a relação \sim acima.*

Definição A.1.3 (Axioma da Escolha). *Se \mathcal{F} é uma família de conjuntos não vazios, então existe uma função de seleção (ou função de escolha) f definida em \mathcal{F} tal que $f(A) \in A$ para todo $A \in \mathcal{F}$.*

Teorema A.1.4 (Vitali). *Existe um subconjunto $V \subset [0, 1]$ tal que não existe nenhuma função $m : \mathcal{P}(\mathbb{R}) \rightarrow [0, +\infty]$ satisfazendo simultaneamente:*

(1) $m(\emptyset) = 0$;

(2) m é σ -aditiva;

(3) $m(I) = \ell(I)$ para todo intervalo $I \subset \mathbb{R}$;

(4) $m(E + x) = m(E)$ para todo $E \subset \mathbb{R}$ e todo $x \in \mathbb{R}$ (invariância por translação).

Em particular, tal V não é mensurável no sentido de Lebesgue.

Demonstração: *Consideremos a relação de equivalência \sim em $[0, 1]$. Cada classe de equivalência é um subconjunto não vazio de $[0, 1]$. Aplicando o Axioma da Escolha à família formada por todas as classes de equivalência, obtemos uma função de escolha que seleciona exatamente um elemento de cada classe. O conjunto V formado por esses representantes é um conjunto de Vitali contido em $[0, 1]$. Seja, então,*

$$Q := \mathbb{Q} \cap [-1, 1],$$

que é enumerável. Para cada $q \in Q$ considere a translação

$$V + q = \{v + q : v \in V\}.$$

Mostremos que as coleções $\{V + q : q \in Q\}$ são mutuamente disjuntas. De fato, suponha que existam $q_1 \neq q_2$ em Q e pontos $v_1, v_2 \in V$ tais que $v_1 + q_1 = v_2 + q_2$. Então $v_1 - v_2 = q_2 - q_1 \in \mathbb{Q}$, logo $v_1 \sim v_2$. Pela construção de V segue-se $v_1 = v_2$, o que implica $q_1 = q_2$, contradizendo a hipótese. Assim, para $q_1 \neq q_2$ tem-se $(V + q_1) \cap (V + q_2) = \emptyset$.

Além disso, a união dessas traduções contém $[0, 1]$. De fato, para cada $x \in [0, 1]$ existe $v \in V$ com $x \sim v$, isto é, $x - v \in \mathbb{Q}$; como $x, v \in [0, 1]$ tem-se $x - v \in [-1, 1]$, portanto $x \in V + q$ com $q = x - v \in Q$. Consequentemente,

$$[0, 1] \subseteq \bigcup_{q \in Q} (V + q) \subseteq [-1, 2].$$

Suponha que exista $m : \mathcal{P}(\mathbb{R}) \rightarrow [0, +\infty]$ satisfazendo todas as quatro condições enunciadas pelo teorema. Pela σ -aditividade e pela disjunção das traduções,

$$m\left(\bigcup_{q \in Q} (V + q)\right) = \sum_{q \in Q} m(V + q).$$

Pela invariância por translação tem-se $m(V + q) = m(V)$ para todo $q \in Q$. Como Q é infinita enumerável, a soma $\sum_{q \in Q} m(V)$ vale:

$$\sum_{q \in Q} m(V) = \begin{cases} 0, & \text{se } m(V) = 0, \\ +\infty, & \text{se } m(V) > 0. \end{cases}$$

Por outro lado, da inclusão $[0, 1] \subseteq \bigcup_{q \in Q} (V + q) \subseteq [-1, 2]$ e da monotonicidade de m (seguida da σ -aditividade) obtemos

$$1 = m([0, 1]) \leq m\left(\bigcup_{q \in Q} (V + q)\right) \leq m([-1, 2]) = \ell([-1, 2]) = 3.$$

Assim $m\left(\bigcup_{q \in Q} (V + q)\right)$ é finito e ≥ 1 , o que contradiz as duas possibilidades anteriores (ou é 0 ou é $+\infty$). Portanto não pode existir medida m definida em toda $\mathcal{P}(\mathbb{R})$ satisfazendo simultaneamente tais propriedades.

□

A.2 O Teorema de Carathéodory

Teorema A.2.1. *Seja μ^* uma medida exterior em Ω e \mathcal{M} a coleção de todos os subconjuntos de Ω que são mensuráveis no sentido de Carathéodory. Então:*

1. \mathcal{M} é uma σ -álgebra;
2. a restrição $\mu := \mu^*|_{\mathcal{M}}$ é uma medida em (Ω, \mathcal{M}) .

Demonstração: Primeiro, note que para todo $E \subset \Omega$ e todo $A \subset \Omega$ sempre vale a desigualdade

$$\mu^*(A) \leq \mu^*(A \cap E) + \mu^*(A \cap E^c),$$

pois $A = (A \cap E) \cup (A \cap E^c)$ e a subaditividade aplicada aos dois conjuntos dá a desigualdade. A propriedade de Carathéodory exige a desigualdade reversa para todos os conjuntos A .

(i) \mathcal{M} é uma σ -álgebra.

(fechamento por complemento) Se $E \in \mathcal{M}$, então para todo A ,

$$\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E^c).$$

Isso é simétrico em E e E^c , então $E^c \in \mathcal{M}$.

(contém o vazio) Colocando $E = \emptyset$ vemos facilmente que $\emptyset \in \mathcal{M}$ (já que $\mu^*(A) = \mu^*(A \cap \emptyset) + \mu^*(A \cap \Omega) = 0 + \mu^*(A)$).

(fechamento por uniões finitas) Sendo $A, B \in \mathcal{M}$, temos

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c).$$

Porém, perceba que $E \cap A^c = (E \cap A^c \cap B) \cup (E \cap A^c \cap B^c)$, logo

$$\mu^*(E \cap A^c) = \mu^*(E \cap A^c \cap B) + \mu^*(E \cap A^c \cap B^c).$$

Somando ambas as equações e manipulando os termos, ficamos com

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c \cap B) + \mu^*(E \cap A^c \cap B^c).$$

Pela subaditividade de μ^* ,

$$\mu^*(E \cap A) + \mu^*(E \cap A^c \cap B) \geq \mu^*((E \cap A) \cup (E \cap A^c \cap B)).$$

Porém, observe que $(E \cap A) \cup (E \cap A^c \cap B) = E \cap (A \cup B)$, portanto

$$\mu^*(E) \geq \mu^*(E \cap (A \cup B)) + \mu^*(E \cap (A \cup B)^c).$$

Como a desigualdade oposta é obténível pela subaditividade, obtemos

$$\mu^*(E) = \mu^*(E \cap (A \cup B)) + \mu^*(E \cap (A \cup B)^c),$$

o que implica $A \cup B \in \mathcal{M}$.

(fechamento por uniões contáveis) Sejam $E \subset \Omega$, $B_n = \bigcup_{k=1}^n A_k$ e $B = \bigcup_{k=1}^{\infty} A_k$, com $\{A_k\}_{k=1}^{\infty} \in \mathcal{M}$. Logo,

$$\mu^*(E \cap B_n) = \mu^*((E \cap B_n) \cap A_n) + \mu^*((E \cap B_n) \cap A_n^c).$$

Como os $\{A_k\}_{k=1}^{\infty}$ são disjuntos dois a dois, $B_n \cap A_n = A_n$ e, para todo $n \geq 2$, temos $B_n \cap A_n^c = B_{n-1}$. Portanto

$$\mu^*(E \cap B_n) = \mu^*(E \cap A_n) + \mu^*(E \cap B_{n-1}).$$

Iterando tal processo (ou aplicando o caso finito recursivamente), obtemos que

$$\mu^*(E \cap B_n) = \sum_{k=1}^n \mu^*(E \cap A_k).$$

Como \mathcal{M} é fechado para uniões contáveis, $B_n \in \mathcal{M}$ para todo n inteiro positivo. Portanto

$$\mu^*(E) = \mu^*(E \cap B_n) + \mu^*(E \cap B_n^c),$$

donde

$$\mu^*(E) = \sum_{k=1}^n \mu^*(E \cap A_k) + \mu^*(E \cap B_n^c).$$

Pela sub-aditividade de μ^* ,

$$\sum_{k=1}^n \mu^*(E \cap A_k) \geq \mu^*\left(\bigcup_{k=1}^n (E \cap A_k)\right) = \mu^*\left(E \cap \left(\bigcup_{k=1}^n A_k\right)\right).$$

Além disso, como $B_n \subset B$, temos $B^c \subset B_n^c$ e portanto $(E \cap B^c) \subset (E \cap B_n^c)$. Pela monotonicidade de μ^* , $\mu^*(E \cap B^c) \leq \mu^*(E \cap B_n^c)$. Note, então, que

$$\mu^*(E) \geq \mu^*\left(E \cap \left(\bigcup_{k=1}^n A_k\right)\right) + \mu^*(E \cap B^c),$$

e fazendo $n \rightarrow \infty$,

$$\mu^*(E) \geq \mu^*(E \cap B) + \mu^*(E \cap B^c).$$

Pela sub-aditividade de μ^* , obtemos a desigualdade contrária, logo

$$\mu^*(E) = \mu^*(E \cap B) + \mu^*(E \cap B^c) \implies B \in \mathcal{M}.$$

Isso demonstra que \mathcal{M} é uma σ -álgebra, como queríamos.

(ii) $\mu^*|_{\mathcal{M}}$ é uma medida.

Basta mostrar que μ é aditiva para uniões contáveis de conjuntos disjuntos.

Novamente, sejam $E \subset \Omega$, $B_n = \bigcup_{k=1}^n A_k$ e $B = \bigcup_{k=1}^{\infty} A_k$, com $\{A_k\}_{k=1}^{\infty} \in \mathcal{M}$ disjuntos dois a dois. Perceba que

$$\mu^*(E) \geq \sum_{k=1}^{\infty} \mu^*(E \cap A_k) + \mu^*(E \cap B^c).$$

Como E é arbitrário,

$$\mu^*(B) \geq \sum_{k=1}^{\infty} \mu^*(B \cap A_k) + \mu^*(B \cap B^c),$$

mas $\mu^*(B \cap B^c) = \mu^*(\emptyset) = 0$ e, para cada k inteiro positivo, $B \cap A_k = A_k$, donde

$$\mu^*(B) \geq \sum_{k=1}^{\infty} \mu^*(A_k).$$

Pela sub-aditividade de μ^* , obtemos a desigualdade contrária, logo $\mu^*(B) = \sum_{k=1}^{\infty} \mu^*(A_k)$ e a restrição $\mu^*|_{\mathcal{M}}$ é uma medida.

□

A.3 O Teorema da Extensão de Carathéodory

Teorema A.3.1. *Seja \mathcal{A} uma álgebra de subconjuntos de Ω e $\mu_0 : \mathcal{A} \rightarrow [0, +\infty]$ uma pré-medida. Então existe uma medida exterior μ^* em Ω tal que:*

- (1) *todo conjunto de \mathcal{A} é mensurável no sentido de Carathéodory (isto é, $\mathcal{A} \subset \mathcal{M}$);*
- (2) *a restrição de μ^* à σ -álgebra \mathcal{M} dos conjuntos mensuráveis é uma medida μ que estende μ_0 .*

Além disso, se μ_0 é σ -finita (isto é, existe $(B_n) \subset \mathcal{A}$ com $\Omega = \bigcup_n B_n$ e $\mu_0(B_n) < \infty$ para todo n), então essa extensão é única na σ -álgebra gerada por \mathcal{A} .

Demonstração: Antes de demonstrar tal teorema, definamos $\mu^* : \mathcal{P}(\Omega) \rightarrow [0, +\infty]$ como

$$\mu^*(E) = \inf \left\{ \sum_{k=1}^{\infty} \mu_0(A_k) : \{A_k\}_{k=1}^{\infty} \in \mathcal{A}, E \subset \bigcup_{k=1}^{\infty} A_k \right\}.$$

Mostremos que μ^* é uma medida exterior.

Tomando $A_k = \emptyset$ para todo k inteiro positivo, pelo ínfimo na definição de μ^* ,

$$\mu^*(\emptyset) \leq \sum_{k=1}^{\infty} \mu_0(A_k) = \sum_{k=1}^{\infty} \mu_0(\emptyset) = 0.$$

Mas de $\mu^* : \Omega \rightarrow [0, +\infty]$, temos $0 \leq \mu^*(\emptyset)$. Portanto $\mu^*(\emptyset) = 0$.

Para mostrar a monotonicidade de μ^* , veja que, se $A \subset B \subset \Omega$, então toda cobertura de B também cobre A , e pelo ínfimo na definição de μ^* , devemos ter que $\mu^*(A) \leq \mu^*(B)$.

Agora, passaremos a demonstrar a sub-aditividade de μ^* .

Seja $\delta > 0$. Pelo ínfimo na definição de μ^* , existe uma cobertura $\bigcup_{i=1}^{\infty} E_{i,j}$ de A_j , com $A_j, \{E_{i,j}\}_{i=1}^{\infty} \in \mathcal{P}(\Omega)$ para cada j inteiro positivo, tal que

$$\sum_{i=1}^{\infty} \mu_0(E_{i,j}) \leq \mu^*(A_j) + \delta 2^{-j}.$$

Como $A_j \subset \bigcup_{i=1}^{\infty} E_{i,j}$ para cada j inteiro positivo, $\bigcup_{j=1}^{\infty} A_j \subset \bigcup_{i,j=1}^{\infty} E_{i,j}$, donde

$$\sum_{i,j=1}^{\infty} \mu_0(E_{i,j}) \leq \sum_{j=1}^{\infty} \mu^*(A_j) + \sum_{j=1}^{\infty} \delta 2^{-j}.$$

Veja que $\sum_{j=1}^{\infty} \delta 2^{-j} = \delta \sum_{j=1}^{\infty} 2^{-j} = \delta$, portanto

$$\sum_{i,j=1}^{\infty} \mu_0(E_{i,j}) \leq \sum_{j=1}^{\infty} \mu^*(A_j) + \delta.$$

Agora, tomando o ínfimo na definição de μ^* ,

$$\mu^*\left(\bigcup_{j=1}^{\infty} A_j\right) \leq \sum_{i,j=1}^{\infty} \mu_0(E_{i,j}),$$

e por transitividade,

$$\mu^*\left(\bigcup_{j=1}^{\infty} A_j\right) \leq \sum_{j=1}^{\infty} \mu^*(A_j) + \delta.$$

Como δ é arbitrário, basta que $\delta \rightarrow 0$. Isso mostra que μ^* é, de fato, uma medida exterior.

Agora, partiremos para a demonstração do teorema.

(i) $\mathcal{A} \subset \mathcal{M}$.

Basta mostrar que, para todo $A \in \mathcal{A}$ e $E \subset \Omega$, temos

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c).$$

Seja $\{B_k\}_{k=1}^{\infty} \subset \mathcal{A}$ uma cobertura de E , isto é, $E \subset \bigcup_{k=1}^{\infty} B_k$. Para cada k ,

$$B_k = (B_k \cap A) \cup (B_k \cap A^c).$$

Como μ_0 é aditiva em famílias contáveis de conjuntos dois a dois disjuntos e $(B_k \cap A), (B_k \cap A^c) \in \mathcal{A}$,

$$\mu_0(B_k) = \mu_0(B_k \cap A) + \mu_0(B_k \cap A^c),$$

e somando em k ,

$$\sum_{k=1}^{\infty} \mu_0(B_k) = \sum_{k=1}^{\infty} \mu_0(B_k \cap A) + \sum_{k=1}^{\infty} \mu_0(B_k \cap A^c).$$

Note que $\{B_k \cap A\}_{k=1}^{\infty}$ é uma cobertura de $E \cap A$ por elementos de \mathcal{A} , e pelo ínfimo na definição de μ^* , $\sum_{k=1}^{\infty} \mu_0(B_k \cap A) \geq \mu^*(E \cap A)$.

Analogamente para $\{B_k \cap A^c\}_{k=1}^{\infty}$,

$$\sum_{k=1}^{\infty} \mu_0(B_k) \geq \mu^*(E \cap A) + \mu^*(E \cap A^c).$$

Como $\{B_k\}_{k=1}^{\infty}$ é uma cobertura arbitrária de E , tomando o ínfimo em μ^* sobre todas as coberturas de E ,

$$\mu^*(E) \geq \mu^*(E \cap A) + \mu^*(E \cap A^c).$$

Pela sub-aditividade de μ^* , obtemos a desigualdade contrária, logo

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c).$$

Como A é arbitrário, μ^* é tal que todo $A \in \mathcal{A}$ é mensurável.

(ii) μ é uma medida que estende μ_0 .

Pelo anterior, $\mu = \mu^* \upharpoonright_{\mathcal{M}}$ é uma medida.

Como A é uma cobertura de si mesmo e $\mu^*(A)$ é o ínfimo sobre todas as coberturas, por definição de μ^* , temos $\mu(A) = \mu^*(A) \leq \mu_0(A)$.

Agora, seja $\{B_k\}_{k=1}^{\infty} \subset \mathcal{A}$ uma cobertura de A e seja $C_k = (A \cap B_k) \setminus \bigcup_{i=1}^{k-1} B_i \in \mathcal{A}$.

Note que $C_i \cap C_j = \emptyset$ para todo $i \neq j$ e $\bigcup_{i=1}^k C_i = A \cap \bigcup_{i=1}^k B_i$, donde

$$\bigcup_{k=1}^{\infty} C_k = A \cap \bigcup_{k=1}^{\infty} B_k = A.$$

Como C_k e $B_k \setminus C_k$ são disjuntos, pela aditividade contável de μ_0 ,

$$\mu_0(B_k) = \mu_0(C_k) + \mu_0(B_k \setminus C_k),$$

e somando em k ,

$$\sum_{k=1}^{\infty} \mu_0(B_k) = \sum_{k=1}^{\infty} \mu_0(C_k) + \sum_{k=1}^{\infty} \mu_0(B_k \setminus C_k).$$

Perceba que

$$\sum_{k=1}^{\infty} \mu_0(B_k) \geq \sum_{k=1}^{\infty} \mu_0(C_k) = \mu_0(A).$$

Como $\{B_k\}_{k=1}^{\infty}$ é uma cobertura arbitrária de A , pelo ínfimo na definição de μ^* ,

$$\mu^*(A) \geq \sum_{k=1}^{\infty} \mu_0(B_k) \geq \mu_0(A) \implies \mu(A) \geq \mu_0(A),$$

e como já mostramos que $\mu(A) = \mu^*(A) \leq \mu_0(A)$, $\mu(A) = \mu_0(A)$ e μ estende μ_0 .

Caso μ_0 é σ -finita.

Definamos:

- \mathcal{P} é um π -sistema se é fechado por interseções finitas.
- \mathcal{L} é um λ -sistema (Sistema de Dynkin) se $\Omega \in \mathcal{L}$, é fechado por complementos relativos a Ω e fechado por uniões contáveis de conjuntos disjuntos.

Utilizaremos o seguinte teorema, cuja demonstração foge do escopo do presente trabalho:

Teorema $\pi - \lambda$ (Dynkin): se \mathcal{P} é π -sistema e \mathcal{L} é λ -sistema com $\mathcal{P} \subset \mathcal{L}$, então $\sigma(\mathcal{P}) \subset \mathcal{L}$. Em particular, se uma λ -sistema contém um π -sistema gerador, então contém a σ -álgebra gerada por ele.

Note que toda álgebra \mathcal{A} é um π -sistema por ser fechada por interseção finita.

Seja ν outra medida definida em $\sigma(\mathcal{A})$ tal que $\nu|_{\mathcal{A}} = \mu_0$. Queremos provar $\nu = \mu$ em $\sigma(\mathcal{A})$.

Para qualquer $E \subset \Omega$ e qualquer cobertura $(A_k) \subset \mathcal{A}$ de E ,

$$\nu(E) \leq \nu\left(\bigcup_k A_k\right) \leq \sum_k \nu(A_k) = \sum_k \mu_0(A_k).$$

Tomando ínfimo sobre todas as coberturas obtemos $\nu(E) \leq \mu^*(E)$. Em particular, na σ -álgebra onde ν está definida, $\nu \leq \mu$.

Analisaremos, então, o caso finito global (isto é, $\mu(\Omega) < \infty$).

Se $\mu(\Omega) < \infty$ (e então $\nu(\Omega) = \mu_0(\Omega) = \mu(\Omega) < \infty$), considere a classe

$$\mathcal{C} := \{E \in \sigma(\mathcal{A}) : \nu(E) = \mu(E)\}.$$

Mostramos que \mathcal{C} é um λ -sistema e contém \mathcal{A} . Então, por $\pi - \lambda$ (já que \mathcal{A} é π -sistema), $\sigma(\mathcal{A}) \subset \mathcal{C}$ e portanto $\nu = \mu$ em $\sigma(\mathcal{A})$.

Observe que:

- $\Omega \in \mathcal{C}$ porque $\nu(\Omega) = \mu(\Omega)$.
- Se $E \in \mathcal{C}$ então $\nu(E^c) = \nu(\Omega) - \nu(E) = \mu(\Omega) - \mu(E) = \mu(E^c)$, então $E^c \in \mathcal{C}$.
- Se $E_1, E_2, \dots \in \mathcal{C}$ são disjuntos então

$$\nu\left(\bigcup_k E_k\right) = \sum_k \nu(E_k) = \sum_k \mu(E_k) = \mu\left(\bigcup_k E_k\right),$$

portanto $\bigcup_k E_k \in \mathcal{C}$.

Assim \mathcal{C} é λ -sistema. Como $\mathcal{A} \subset \mathcal{C}$ (porque ν e μ coincidem em \mathcal{A} pela hipótese), pela $\pi - \lambda$ concluímos $\sigma(\mathcal{A}) \subset \mathcal{C}$. Logo $\nu = \mu$ em $\sigma(\mathcal{A})$.

Agora, analisemos o σ -finito geral.

Suponha apenas σ -finitude: exista $(B_n)_{n \geq 1} \subset \mathcal{A}$ com $\Omega = \bigcup_n B_n$ e $\mu_0(B_n) < \infty$. Para cada n definimos medidas finitas nas partes B_n por

$$\mu^{(n)}(E) := \mu(E \cap B_n), \quad \nu^{(n)}(E) := \nu(E \cap B_n),$$

definidas para $E \in \sigma(\mathcal{A})$. Note que $\mu^{(n)}(\Omega) = \mu(B_n) = \mu_0(B_n) < \infty$ e $\nu^{(n)}(\Omega) = \nu(B_n) = \mu_0(B_n) < \infty$. Além disso $\mu^{(n)}$ e $\nu^{(n)}$ coincidem em \mathcal{A} intersectada com B_n . Aplicando o caso finito global à σ -álgebra $\sigma(\mathcal{A})$ restrita a B_n (ou, de forma equivalente, aplicando o caso finito global às medidas $\mu^{(n)}, \nu^{(n)}$ no mesmo espaço), deduzimos

$$\mu(E \cap B_n) = \nu(E \cap B_n) \quad \text{para todo } E \in \sigma(\mathcal{A}), \forall n.$$

Finalmente, para qualquer $E \in \sigma(\mathcal{A})$ temos

$$\mu(E) = \sum_{n=1}^{\infty} \mu(E \cap B_n) = \sum_{n=1}^{\infty} \nu(E \cap B_n) = \nu(E),$$

onde a troca de ordens e a soma são válidas porque as interseções $E \cap B_n$ partem E em pedaços disjuntos e ambos os lados são não negativos. Assim $\mu(E) = \nu(E)$ para todo $E \in \sigma(\mathcal{A})$. Portanto a extensão é única na σ -álgebra gerada por \mathcal{A} .

□

Observação A.3.2. A condição de σ -finitude é essencial para a unicidade na σ -álgebra gerada por \mathcal{A} . Sem σ -finitude artefatos podem surgir e não podemos garantir unicidade.

A.4 Compatibilidade com a Sigma-Álgebra de Borel

Proposição A.4.1. A σ -álgebra de Borel $\mathcal{B}(\mathbb{R}^n)$ está contida na σ -álgebra de Lebesgue $\mathcal{L}(\mathbb{R}^n)$. Além disso, a medida de Lebesgue $\mu = \mu^*|_{\mathcal{L}(\mathbb{R}^n)}$ estende a pré-medida geométrica μ_0 , isto é,

$$\mu(R) = \mu_0(R), \quad \forall R \in \mathcal{R}.$$

Demonstração: Seja μ_0 a pré-medida geométrica definida na família \mathcal{R} dos retângulos semiabertos de \mathbb{R}^n e seja μ^* a medida exterior associada a μ_0 . Considere a coleção

$$\mathcal{L}(\mathbb{R}^n) = \{E \subset \mathbb{R}^n : \mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E^c), \forall A \subset \mathbb{R}^n\}.$$

Pelo Teorema de Extensão de Carathéodory, $\mathcal{L}(\mathbb{R}^n)$ é uma σ -álgebra e a restrição $\mu = \mu^*|_{\mathcal{L}(\mathbb{R}^n)}$ é uma medida.

Passo 1: Retângulos semiabertos são mensuráveis e preservam a pré-medida. Pelo próprio teorema de Carathéodory, todo conjunto pertencente ao domínio original da pré-medida μ_0 é mensurável no sentido de Carathéodory. Em particular,

$$\mathcal{R} \subset \mathcal{L}(\mathbb{R}^n) \quad e \quad \mu(R) = \mu_0(R), \quad \forall R \in \mathcal{R}.$$

Como $\mathcal{L}(\mathbb{R}^n)$ é uma σ -álgebra, segue que

$$\sigma(\mathcal{R}) \subset \mathcal{L}(\mathbb{R}^n).$$

Passo 2: A σ -álgebra gerada por \mathcal{R} coincide com a de Borel. Seja \mathcal{Q} a família dos retângulos abertos de \mathbb{R}^n da forma

$$Q = \prod_{i=1}^n (a_i, b_i), \quad a_i, b_i \in \mathbb{Q}, \quad a_i < b_i.$$

A família \mathcal{Q} é enumerável, pois \mathbb{Q} é enumerável e produtos finitos de conjuntos enumeráveis permanecem enumeráveis.

Mostremos que todo conjunto aberto $G \subset \mathbb{R}^n$ pode ser escrito como união enumerável de elementos de \mathcal{Q} . Seja $x \in G$. Pela definição de conjunto aberto, existe $\varepsilon > 0$ tal que

$$\prod_{i=1}^n (x_i - \varepsilon, x_i + \varepsilon) \subset G.$$

Como os números racionais são densos em \mathbb{R} , é possível escolher $a_i, b_i \in \mathbb{Q}$ satisfazendo

$$x_i - \varepsilon < a_i < x_i < b_i < x_i + \varepsilon,$$

para todo $i = 1, \dots, n$. Assim, obtemos um retângulo $Q \in \mathcal{Q}$ tal que $x \in Q \subset G$. Repetindo esse procedimento para cada ponto de G e tomando a união de todos os retângulos obtidos, conclui-se que G é união de elementos de \mathcal{Q} . Como \mathcal{Q} é enumerável, essa união pode ser tomada enumerável. Logo,

$$\mathcal{B}(\mathbb{R}^n) = \sigma(\mathcal{Q}).$$

Por outro lado, cada retângulo $Q \in \mathcal{Q}$ pode ser escrito como união enumerável de retângulos semiabertos. De fato,

$$\prod_{i=1}^n (a_i, b_i) = \bigcup_{k=1}^{\infty} \prod_{i=1}^n \left[a_i + \frac{1}{k}, b_i \right).$$

Cada conjunto do lado direito pertence a \mathcal{R} , o que implica

$$\mathcal{Q} \subset \sigma(\mathcal{R}) \quad \Rightarrow \quad \sigma(\mathcal{Q}) \subset \sigma(\mathcal{R}).$$

Como também é claro que $\mathcal{R} \subset \mathcal{B}(\mathbb{R}^n)$, obtemos

$$\sigma(\mathcal{R}) \subset \mathcal{B}(\mathbb{R}^n) \quad e \quad \mathcal{B}(\mathbb{R}^n) = \sigma(\mathcal{Q}) \subset \sigma(\mathcal{R}),$$

logo

$$\sigma(\mathcal{R}) = \mathcal{B}(\mathbb{R}^n).$$

Passo 3: Conclusão. Dos passos anteriores, segue que

$$\mathcal{B}(\mathbb{R}^n) = \sigma(\mathcal{R}) \subset \mathcal{L}(\mathbb{R}^n).$$

Além disso, a medida de Lebesgue μ estende a pré-medida geométrica μ_0 , isto é, $\mu(R) = \mu_0(R)$ para todo $R \in \mathcal{R}$.

□