



LAURO KONDARZEWSKI

Probabilidade Geométrica e Grandes Números

Santo André, 2013



Universidade Federal do ABC

Centro de Matemática, Computação e Cognição

Lauro Kondarzewski

Probabilidade Geométrica e Grandes Números

Orientador: Prof. Dr. Rafael de Mattos Grisi

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

ESTE EXEMPLAR CORRESPONDE A VERSÃO FINAL DA DISSERTAÇÃO
DEFENDIDA PELO ALUNO LAURO KONDARZEWSKI,
E ORIENTADA PELO PROF. DR. RAFAEL DE MATTOS GRISI.

Santo André, 2013

AGRADECIMENTOS

Agradeço a Deus, a meus pais, a meus irmãos e a meus colegas de trabalho por terem me suportado nestes dias, ao longo do curso PROFMAT, nos quais ficaram suscetíveis a toda sorte de mudança de humor por que passei, e pela compreensão e paciência que tiveram comigo dando me todo o apoio necessário.

Agradeço a meu Orientador, Prof. Dr. Rafael de Mattos Grisi, pelos ensinamentos, pela boa vontade, pela paciência e pela dedicação, pois sem o seu cuidado este trabalho jamais teria acontecido.

RESUMO

Este trabalho visa mostrar Probabilidades em quatro visões: a frequentista, por contagem, a axiomática e a geométrica.

Mostra-se que as vertentes frequentista, por contagem e geométrica atendem às demandas axiomáticas, e através de alguns exemplos, culmina-se por se fazer uma ligação entre todas estas frentes e a Lei dos Grandes Números, cuja demonstração se verifica no Apêndice.

Palavras-chave: Probabilidade geométrica, Grandes números

ABSTRACT

This paper targets to show Probabilities in four ways of description: frequentist, by counting, axiomatics and geometric.

It shows that frequentist, by counting and geometric ways attend to axiomatics vision demands, and through some examples, it finishes by linking all these ways of seeing probabilities with Large Numbers Law, which demonstration is in the Appendix.

Keywords: Geometric probability, Large Numbers

CONTEÚDO

Introdução	1
1 NOÇÕES DE PROBABILIDADE	3
1.1 Probabilidade como frequência de repetições	6
1.2 Probabilidade via Contagem	8
2 CONCEITOS TEÓRICOS	11
2.1 Axiomas da Probabilidade	12
2.1.1 Probabilidade Condicional e Independência	15
2.2 Lei dos Grande Números	17
2.2.1 Erro nas aproximações	22
2.3 Atividades Práticas	24
3 PROBABILIDADE GEOMÉTRICA	27
3.1 Probabilidade em Espaços Contínuos	27
3.2 Problemas em Probabilidade Geométrica	29
3.2.1 Dividindo um Segmento e Construindo Triângulos	29
3.2.2 A Agulha de Buffon	31
3.3 Atividades Práticas	35
A CÁLCULO DA ÁREA EM BAIXO DA CURVA DO SENO	39

INTRODUÇÃO

No primeiro capítulo, noções básicas de probabilidade são apresentadas, bem como termos e conceitos que são utilizados em todo o texto. A noção frequentista de probabilidades está neste capítulo. Em seguida, apresenta-se a noção de probabilidades via contagem, bem como suas propriedades.

No segundo capítulo, introduz-se os axiomas de Kolmogorov e suas principais propriedades, culminando na Lei dos Grandes Números de Bernoulli.

No terceiro e último capítulo, apresenta-se a noção de probabilidade geométrica, além do famoso problema da Agulha de Buffon.

No apêndice A é mostrado um cálculo da área sob a curva do seno sem a utilização de cálculo integral.

No apêndice B, tem-se uma demonstração da Lei dos Grandes Números.

NOÇÕES DE PROBABILIDADE

No minidicionário Aurélio, o substantivo feminino probabilidade se refere à qualidade de provável. O adjetivo provável já figura com significados diversos como o que se pode provar; que parece verdadeiro; verossímil; com base em indícios convincentes, mas sem prova definitiva ou que se pode, com motivos razoáveis; esperar que aconteça (fato, acontecimento).

Surge então uma pergunta: como modelar matematicamente tal noção. Em outras palavras, como mensurar tal conceito e traduzir em número. Nas seções que seguem vamos apresentar algumas tentativas de fazer isso, algumas menos bem sucedidas que outras, mas todas com alguma influência na probabilidade moderna

Para tal, precisamos antes definir alguns conceitos importantes, e comuns a todas os modelos que apresentaremos.

Nosso principal objetivo é mensurar a probabilidade de observarmos certos fatos quando da realização de um *experimento*. Um experimento é a realização (controlada ou não) de qualquer fenômeno reproduzível. Existem dois tipos de experimentos a saber:

- **Experimento Determinístico ou Determinista:** são experimentos realizados acerca de resultados já esperados, mas que visam a compreensão global de fenômenos. Por exemplo, as leis físicas são modelos matemáticos que reproduzem com grande exatidão diversos fenômenos físicos, e foram obtidas através de experimentação laboratorial isolando-se alguns fatores de modo a se compreender minuciosamente tais fenômenos, e assim podem ser aplicadas na vida. Várias aplicações são notáveis no dia a dia, como nos produtos das Engenharias.

- **Experimentos Aleatórios:** um experimento é considerado aleatório quando não temos nenhum controle sobre o resultado. Por exemplo, ao se atirar uma moeda temos certeza que ela irá cair, mas poderá cair com uma das faces voltada para cima, mas não temos nenhuma expectativa sobre qual face observaremos.

O nosso alvo é “medir” a probabilidade de um certo fato A que se percebe repetir em um dado experimento. Deste ponto em diante, o chamaremos Evento A . Todos os fatos percebidos, iguais a A ou não, ocorridos durante todo o experimento serão chamados daqui por diante de Espaço Amostral. Em outras palavras, os eventos são subconjuntos do espaço amostral.

Definição 1.1. Chamaremos de **Espaço Amostral** o conjunto formado por todos os resultados possíveis de um experimento, e de **evento** qualquer subconjunto observável do espaço amostral.

Vamos tentar esclarecer um pouco tais conceitos.

Exemplo 1.1. Considere o experimento de lançar uma moeda e verificar se ocorreu cara ou coroa. O espaço amostral neste caso é o conjunto $\Omega = \{cara, coroa\}$ e alguns eventos possíveis são $A = \{cara\}$ e $B = \{coroa\}$. Se o experimento for o lançamento de um dado, o espaço amostral é $1, 2, 3, 4, 5, 6$ e o evento pode ser a ocorrência de pontuação maior que 4, ou seja, o conjunto $5, 6$.

Exemplo 1.2. Se o experimento for o lançamento de um dado, teremos então o espaço amostral $\Omega = \{1, 2, 3, 4, 5, 6\}$. Neste caso, o evento $A = \{\text{ocorrência de pontuação maior que } 4\}$ é representado pelo conjunto $A = \{5, 6\}$.

É interessante observar que em qualquer espaço amostral Ω os conjuntos \emptyset e Ω são eventos. Estes são conhecidos como eventos triviais, pois representam as ideias básicas de *não aconteceu nada* e *aconteceu qualquer coisa*, respectivamente.

Como estaremos trabalhando com conjuntos e subconjuntos é importante que saibamos interpretar as operações entre conjuntos com base neste contexto. Para isso considere um espaço amostral Ω e dois eventos $A, B \subseteq \Omega$.

UNIÃO DE EVENTOS: A união dos eventos A e B nos dá um novo evento $A \cup B$ que pode ser descrito por

$$A \cup B = \{\text{foi observado o evento } A \text{ ou o evento } B\}.$$

O mesmo conceito se estende para qualquer coleção de eventos. Assim, para eventos $A_1, A_2, A_3, \dots \subseteq \Omega$ temos

$$\bigcup_{k=1}^n A_k = \{\text{foi observado o evento } A_k \text{ para algum } k \text{ entre } 1 \text{ e } n\}.$$

INTERSECÇÃO DE EVENTOS: A intersecção dos eventos A e B nos dá um novo evento $A \cap B$ que pode ser descrito por

$$A \cap B = \{\text{foi observado o evento } A \text{ e o evento } B\}.$$

Estendendo para qualquer coleção de eventos temos, para eventos $A_1, A_2, A_3, \dots \subseteq \Omega$ que

$$\bigcap_{k=1}^n A_k = \{\text{foi observado o evento } A_k \text{ para todo } k \text{ entre } 1 \text{ e } n\}.$$

COMPLEMENTO: Dizer que ocorreu o complemento de um evento A é o mesmo que dizer que o evento A não foi observado, e isso é representado pelo conjunto A^c . Assim

$$A^c = \{\text{o evento } A \text{ não foi observado}\}.$$

DIFERENÇA DE EVENTOS: A diferença dos eventos A e B nos dá um novo evento $A \setminus B$ que pode ser descrito por

$$A \setminus B = \{\text{foi observado o evento } A \text{ mas não o evento } B\}.$$

Quando dois eventos são disjuntos ($A \cap B = \emptyset$) diremos que estes eventos não podem ser observados simultaneamente, e portanto são **mutuamente exclusivos**. No

lançamento de um dado, por exemplo, os eventos $A = \{\text{observar um valor par}\}$ e $B = \{\text{observar um valor ímpar}\}$ são mutuamente exclusivos.

De modo mais geral vale

Definição 1.2. Diremos que eventos $A_1, \dots, A_n \subseteq \Omega$ são **mutuamente exclusivos** se

$$A_i \cap A_j = \emptyset, \quad \text{para todo } i \neq j, \quad i, j = 1, \dots, n.$$

Ou seja, se quaisquer dois eventos são disjuntos.

Com estes conceitos na cabeça, podemos seguir para as primeiras definições de probabilidade.

1.1 PROBABILIDADE COMO FREQUÊNCIA DE REPETIÇÕES

A palavra frequência significa, pelo dicionário Aurélio, o ato ou efeito de frequentar, repetição amiudada de fatos ou acontecimentos, as pessoas que frequentam um lugar ou o número de ciclos que um sistema com movimento periódico efetua na unidade de tempo.

A noção Frequentista ou Frequencial de Probabilidades nos diz que, se em uma dada experiência repetida n vezes, onde n deve ser grande, um certo evento A foi observado n_A vezes, então a Probabilidade de ocorrer este evento é sua frequência relativa, ou seja, o quociente

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

Esta definição de probabilidade apareceu no início do século XX como tentativa de corrigir problemas no modelo de contagem, que se baseava fortemente em simetrias do problema para definir as probabilidades, e portanto falhava na presença de falta de simetrias.

O modelo acima claramente possui algumas vantagens. É um modelo simples, com poucas hipóteses, e fácil de ser compreendido. Ele supostamente capturaria as probabilidades “naturais” de seu experimento sem que fosse necessário supor nada

sobre o problema em questão. Mas é exatamente neste ponto que surgem diversas dúvidas quanto à esta definição:

- Até que ponto esse número é confiável?
- Quando podemos nos dar por satisfeitos quanto ao número de experiências a serem realizadas de modo que o quociente retrate fielmente ao valor esperado? O que é um número satisfatório de experiências? Seria a Probabilidade apenas uma razão?
- Como seria modelar um experimento frequentista? Quais seriam as dificuldades?

A principal crítica a este modelo é que a noção de frequência relativa está baseada na hipótese simples, mas muito forte, de que a frequência relativa se aproxima de um dado valor quando n é grande, e que tal valor não depende dos resultados obtidos nas diversas repetição, sendo assim intrínseco ao experimento.

Mesmo que aceitássemos esta hipótese, temos ainda o problema de que a aproximação da frequência a um valor é apenas teórica. Na prática estaríamos lidando com uma quantidade fixa de lançamentos, e neste caso a definição passaria a depender deste total. Podemos dizer então que uma probabilidade de um evento será confiável à medida que ele consiga traduzir aproximadamente a frequência de repetição do evento A estudado. Precisariamos saber então o que significa n grande.

Observe também que só saberemos as probabilidades de dados eventos **a posteriori**, ou seja, depois que o experimento foi realizado. Sem um modelo **a priori** é muito difícil testar a validade dos valores encontrados.

Sua facilidade de definir e o fato de ser natural para boa parte das pessoas, torna este um bom modelo para se introduzir Probabilidades a alunos de Ensino Fundamental. Deve-se obviamente levar em conta o tipo de experimento a se apresentar, que de certo modo, tem que priorizar fácil assimilação de conceitos sem ser cansativo, e de simples sistematização na execução, observação e anotação de resultados.

O trabalho com as criança fica ainda mais rico quando confrontamos tal definição com a também intuitiva definição por contagem, que introduziremos a seguir.

1.2 PROBABILIDADE VIA CONTAGEM

Como dissemos anteriormente, precisamos um modelo a priori para que possamos validar as frequências encontradas no modelo frequentista. Tal noção, no entanto, veio historicamente antes da noção frequentista apresentada anteriormente. De fato, o conceito frequentista foi apresentado como modelo a posteriori, com a intenção de comparar e validar os modelos a priori existentes.

A primeira noção de probabilidade apareceu com Pascal e Fermat no século XVII, com a definição clássica sendo enunciada por Laplace no século XVIII.

A definição usada por Pascal e enunciada por Laplace era baseada na contagem de casos favoráveis na realização de um experimento.

Definição 1.3 (Definição Clássica de Probabilidade). Probabilidade é um número entre 0 e 1 que mede as chances de um dado evento A qualquer ocorrer em um dado espaço amostral Ω , e é medido pela fração entre o número de elementos em A , representado por $n(A)$, e o número de elementos de Ω , representado por $n(\Omega)$.

Assim, se representarmos a probabilidade de ocorrer o evento A por $\mathbb{P}(A)$, temos

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

Observe que este modelo não depende da realização do experimento em questão, de tal forma que as probabilidades atribuídas são baseadas em hipóteses feitas sobre o experimento em questão. Como queríamos, este é um modelo a priori de probabilidade. Mais a frente comentaremos mais sobre as vantagens e desvantagens e tal modelo, mas antes vamos ver alguns exemplos e propriedades deste modelo.

Exemplo 1.3. Se jogarmos uma moeda, o espaço amostral seria portanto $\Omega = \{cara, coroa\}$ e teríamos, por exemplo

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

Exemplo 1.4. Se jogarmos um dado, o espaço amostral seria $\Omega = \{1, 2, 3, 4, 5, 6\}$. Considerando

$$A = \{\text{o valor lançado é par}\} = \{2, 4, 6\}$$

e

$$B = \{\text{o valor lançado é maior que 4}\} = \{5, 6\}$$

teríamos

$$\begin{aligned}\mathbb{P}(A) &= \frac{1}{2} \\ \mathbb{P}(B) &= \frac{1}{3} \\ \mathbb{P}(A \cap B) &= \frac{1}{6} \\ \mathbb{P}(A \cup B) &= \frac{2}{3}.\end{aligned}$$

Vamos explorar algumas das propriedades da definição acima.

(i) $\mathbb{P}(\emptyset) = 0$

Isso segue diretamente do fato de $n(\emptyset) = 0$.

(ii) $\mathbb{P}(\Omega) = 1$, que segue de

$$\mathbb{P}(\Omega) = \frac{n(\Omega)}{n(\Omega)} = 1$$

(iii) $0 \leq \mathbb{P}(A) \leq 1$

De fato, como $A \subseteq \Omega$, então $0 \leq n(A) \leq n(\Omega)$ e a propriedade segue.

(iv) $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$

Se A é um evento de Ω e A^c é seu evento complementar, então temos que $A \cup A^c = \Omega$, e $A \cap A^c = \emptyset$ deste modo $n(A) + n(A^c) = n(\Omega)$. Assim

$$\mathbb{P}(A) + \mathbb{P}(A^c) = \frac{n(A)}{n(\Omega)} + \frac{n(A^c)}{n(\Omega)} = \frac{n(\Omega)}{n(\Omega)} = 1.$$

(v) Se $A, B \subseteq \Omega$ são mutuamente exclusivos ($A \cap B = \emptyset$) então $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$.

Basta perceber que $n(A \cup B) = n(A) + n(B)$.

(vi) Se A_1, A_2, \dots, A_n são mutuamente exclusivos então

$$\mathbb{P}\left(\bigcup_{k=1}^n A_k\right) = \mathbb{P}(A_1) + \mathbb{P}(A_2) + \dots + \mathbb{P}(A_n).$$

Basta notar que se A_1, A_2, \dots, A_n são mutuamente exclusivos então

$$n\left(\bigcup_{k=1}^n A_k\right) = n(A_1) + n(A_2) + \dots + n(A_n).$$

(vii) Para quaisquer eventos $A, B \subseteq \Omega$ vale $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$

Para quaisquer subconjuntos A, B de Ω sabemos que $n(A \cup B) = n(A) + n(B) - n(A \cap B)$. O resultado segue dividindo por $n(\Omega)$.

(viii) $\mathbb{P}(A \setminus B) = \mathbb{P}(A) - \mathbb{P}(A \cap B)$

Segue da definição de diferença de conjuntos que $n(A \setminus B) = n(A) - n(A \cap B)$, de onde segue o resultado.

É interessante observar que as propriedades (iv), (v), (vii) e (viii) são todas particularizações ou consequências diretas das propriedades (i), (ii), (iii) e (vi). Este fato será útil no próximo capítulo, quando definirmos os axiomas de probabilidade.

O modelo de probabilidades via contagem é o modelo ensinado hoje no ensino médio, e realmente tem a imensa vantagem de ser fácil de se compreender. Mas tal facilidade nasce de um conjunto de hipóteses extremamente fortes, que limitam em muito suas aplicações. Tais hipóteses são o espaço amostral finito e resultados equiprováveis.

Exploraremos estes problemas no próximo capítulo, onde trataremos da definição axiomática de probabilidade colocada por Kolmogorov.

2

CONCEITOS TEÓRICOS

Considere o experimento de se lançar uma moeda até que se observe a ocorrência de cara pela primeira vez. Para estudar este experimento precisamos antes delimitar seu espaço amostral. Neste caso, como cada resultado do experimento é uma sequência de *coroas* terminadas em uma *cara*, temos que o espaço amostral é dado por

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

que é claramente um espaço com infinitos elementos.

Imagine agora um jogador trapaceiro em um cassino. Dispondo de meios para isso ele constrói um certo dado de modo que a metade inferior do cubo que o compõe seja de acrílico, cuja densidade é de $1,19g/cm^3$ e a outra seja de polietileno de baixa densidade, $0,94g/cm^3$. Por esses motivos, a face que estiver feita com acrílico tenderá a ficar embaixo mais vezes, a face de cima será menos provável, enquanto as faces laterais teriam as mesmas chances de ocorrência. Deste modo, no lançamento deste dado, apesar de ainda termos o espaço amostral $\Omega = \{1, 2, 3, 4, 5, 6\}$, não podemos mais dizer que a probabilidade de observar cada valor é igual a $1/6$.

Os exemplos acima ilustram as maiores limitações encontradas na definição de probabilidade pela contagem de casos favoráveis. Precisamos então de uma nova definição, mas sem perder de vista o que já conseguimos. Precisamos então de uma definição que seja consistente com a noção frequentista e com a definição de Laplace.

2.1 AXIOMAS DA PROBABILIDADE

A noção atual de probabilidade, responsável por colocar esta área novamente no centro da atenção de matemáticos ao redor do mundo, foi proposta em por Kolmogorov (1903-1987). A ideia central é propor um conjunto minimal de axiomas que descrevam as propriedades que esperamos encontrar em uma probabilidade.

As principais propriedades esperadas são exatamente aquelas descritas no modelo de Laplace. Seria possível listar todas como axiomais, mas como foi observado, parte delas é consequência das demais, permitindo que o conjunto seja reduzido a um total mais razoável

Segundo Kolmogorov, uma probabilidade ou medida de probabilidade \mathbb{P} em um espaço amostral Ω deve satisfazer aos seguintes critérios ou axiomas:

- (i) $\mathbb{P}(A) \geq 0$ para todo evento $A \subseteq \Omega$;
- (ii) $\mathbb{P}(\Omega) = 1$;
- (iii) Se A_1, A_2, \dots, A_n são mutuamente exclusivos então

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

Vamos voltar ao exemplo do dado viciado que colocamos acima.

Exemplo 2.1. Como já colocado, no lançamento de um dado (viciado ou não) o espaço amostral que estamos interessados é $\Omega = \{1, 2, 3, 4, 5, 6\}$. Para definir uma probabilidade em Ω podemos começar atribuindo um valor de probabilidade para cada elemento.

Pelo axioma (iii) é necessário então que a probabilidade de um evento seja dado pela soma da probabilidade de seus elementos, uma vez que se $A = \{x_1, \dots, x_i\}$ podemos escrever

$$A = \{x_1\} \cup \dots \cup \{x_i\},$$

onde os eventos $\{x_1\}, \dots, \{x_i\}$ são mutuamente exclusivos.

Da mesma forma, pelos axiomas (ii) e (iii) precisamos que

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

Imagine, então, que se criou um dado que por construção, metade do hexaedro seja constituído por uma substância pouco densa e a outra por substância muito densa que nos dê probabilidade p para a face mais leve e probabilidade q para a mais densa, restando para as laterais a mesma probabilidade r . Assim, teríamos a expressão $p + q + 4r = 1$. Suponha que $p = 1/3$ e $q = 1/6$. Temos assim que

$$\frac{1}{3} + \frac{1}{6} + 4r = 1.$$

E portando $4r = 1/2$, resultando na probabilidade $r = 1/8$ para as outras 4 faces.

Se o lado mais denso for o da face 1 e o oposto for o da face 6 teremos

$$\mathbb{P}(\{1\}) = \frac{1}{6}, \quad \mathbb{P}(\{6\}) = \frac{1}{3} \quad \text{e} \quad \mathbb{P}(\{2\}) = \mathbb{P}(\{3\}) = \mathbb{P}(\{4\}) = \mathbb{P}(\{5\}) = \frac{1}{8}.$$

Procedendo como no exemplo acima, se o espaço amostral é formado por um número finito de elementos, a probabilidade de um evento é dada pela soma das probabilidades dos elementos daquele evento (cada elemento representa um conjunto unitário). E portanto, para definir uma probabilidade em um espaço amostral finito, basta atribuímos valores positivos a cada elemento, de modo que a soma de todos seja igual a 1

Vale observar o modelo de Laplace e quanto o frequentista (com o número de lançamentos fixo) atendem aos axiomas de Kolmogorov. Para o modelo de contagem verificamos isso nas propriedades que elencamos. Para o modelo frequentista a verificação é similar, uma vez que a probabilidade de um evento é também obtida através da contagem de ocorrências.

Vamos agora colocar algumas propriedades importantes de uma probabilidade, verificando a veracidade de algumas delas. As demais ficarão como exercício para o leitor curioso.

P₁ $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$ para todo evento A ;

Demonstração. Demonstração: Considerando que $\Omega = A \cup A^c$, segue do axioma (iii) que $\mathbb{P}(\Omega) = \mathbb{P}(A) + \mathbb{P}(A^c)$. O axioma (ii) nos informa que $\mathbb{P}(\Omega) = 1$, e portanto $1 = \mathbb{P}(A) + \mathbb{P}(A^c)$ e o resultado segue. \square

CONCEITOS TEÓRICOS

P2 $\mathbb{P}(\emptyset) = 0$

Demonstração. Observe que $\emptyset = \Omega^c$, de modo que por (P1) e pelo axioma (ii), $\mathbb{P}(\emptyset) = 1 - \mathbb{P}(\Omega) = 0$ □

P3 $0 \leq \mathbb{P}(A) \leq 1$

P4 Se $A \subseteq B$ então $\mathbb{P}(A) \leq \mathbb{P}(B)$ e $\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A)$

Demonstração. Observe que se $A \subseteq B$ então $B = (B \setminus A) \cup A$ e que $(B \setminus A) \cap A = \emptyset$. Segue do axioma (iii) que

$$\mathbb{P}(B) = \mathbb{P}(B \setminus A) + \mathbb{P}(A),$$

mostrando uma das afirmativas. A segunda sai do axioma (i), que garante que $\mathbb{P}(B \setminus A) \geq 0$, e portanto

$$\mathbb{P}(A) \leq \mathbb{P}(A) + \mathbb{P}(B \setminus A) = \mathbb{P}(B).$$

□

P5 Para quaisquer eventos $A, B \subseteq \Omega$, vale que $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.

Demonstração. Observe que

$$A = (A \cap B) \cup (A \cap B^c)$$

$$B = (A \cap B) \cup (A^c \cap B)$$

e portanto

$$A \cup B = (A \cap B) \cup (A \cap B^c) \cup (A^c \cap B).$$

Segue que

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

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

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

Temos assim que

$$\mathbb{P}(A) + \mathbb{P}(B) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap B) + \mathbb{P}(A \cap B^c) + \mathbb{P}(A^c \cap B) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cup B),$$

e o resultado segue. □

P6 Para quaisquer eventos A, B em Ω , vale que $\mathbb{P}(A \cup B) \leq \mathbb{P}(A) + \mathbb{P}(B)$.

2.1.1 Probabilidade Condicional e Independência

Uma das questões teóricas centrais desenvolvidas por Kolmogorov é o estudo da correlação entre eventos, em outras palavras, como o acontecimento de um evento trás informações sobre um outro evento. Considere, por exemplo, uma companhia seguradora que precisa avaliar o risco que um determinado segurado trás para a companhia. Este risco é medido pela probabilidade de um segurado qualquer sofrer um sinistro no período de um ano. Tendo necessidade de personalizar seu produto, a empresa determina o perfil do usuário através do formulário preenchido no momento da cotação. É comum, por exemplo, que uma mulher com 29 anos pague menos que um homem de mesma idade, pois os fatores sexo e idade estão relacionados a riscos que são levantados durante anos através dos diversos tipos de sinistros e de seus respectivos condutores.

Nesses casos, a probabilidade do evento {o usuário sofrerá um sinistro} é modificada em face da ocorrência prévia de outro, como {o segurado é homem}.

Este cálculo, que descreveremos a seguir, é chamado de Probabilidade Condicionada ou Probabilidade Condicional.

Dados dois eventos A, B com probabilidades positivas, queremos calculemos a probabilidade de ocorrer o evento A sabendo que o evento B já tenha acontecido.

Com o intuito de simplificar e motivar a definição, vamos considerar o modelo clássico de Laplace, via contagem de casos favoráveis.

A primeira coisa a perceber é que para qualquer definição que fizermos, sabendo que B ocorreu, todo e qualquer elemento em $A \setminus B$ passa a ter probabilidade 0. Deste modo estamos interessados apenas nos elementos de $A \cap B$.

O segundo passo é perceber que a nova probabilidade de B deve ser 1, pois a probabilidade de B ocorrer dado que B ocorreu é claramente 1.

O terceiro passo é mais complicado, mas passa por supor que se cada elemento em B tinha a mesma probabilidade antes de B ocorrer, isso deveria permanecer assim após a ocorrência de B .

Com isso o número de casos favoráveis à ocorrência do evento A são iguais a $n(A \cap B)$.

E como agora só nos interessam os elementos em B , o total de resultados possíveis passa a ser $n(B)$. Assim, representando por $\mathbb{P}(A|B)$ a probabilidade de ocorrer o evento A dado que o evento B ocorreu, temos

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

Se, na fórmula acima obtida, nós dividirmos o numerador e o denominador por $n(\Omega)$, teremos

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

Esta expressão, que agora tem significado em um modelo de probabilidade qualquer, pode ser usado para definir probabilidade condicional de forma mais geral. Desta forma, temos:

Definição 2.1. Seja \mathbb{P} uma probabilidade em um espaço amostral Ω . Tome A, B eventos em Ω tais que $\mathbb{P}(B) > 0$. Definimos a Probabilidade de do evento A dado que o evento B ocorra por

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

Para ilustrar considere o seguinte exemplo.

Exemplo 2.2. Considere o modelo do dado viciado descrito no Exemplo 2.1. Temos então $\Omega = \{1, 2, 3, 4, 5, 6\}$ e

$$\mathbb{P}(\{1\}) = \frac{1}{6}, \quad \mathbb{P}(\{6\}) = \frac{1}{3} \quad \text{e} \quad \mathbb{P}(\{2\}) = \mathbb{P}(\{3\}) = \mathbb{P}(\{4\}) = \mathbb{P}(\{5\}) = \frac{1}{8}.$$

Segue que se $A = \{1, 3, 5\}$ e $B = \{1, 2, 3\}$ então

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{P(B)} = \frac{\mathbb{P}(\{1, 3\})}{P(B)} = \frac{1/6 + 1/8}{1/6 + 1/8 + 1/8} = \frac{7}{10}.$$

É interessante notar que, no exemplo acima temos

$$\mathbb{P}(A) = \frac{5}{12} \neq \frac{7}{10} = \mathbb{P}(A|B),$$

de modo que a ocorrência do evento B trás informações sobre a ocorrência de A .

Mas existem eventos tais que a ocorrência de um deles não influencia na probabilidade dos demais. Neste caso diremos que os eventos Independentes.

Nesses moldes, se um evento A e um evento B são independentes, então

$$\mathbb{P}(A|B) = \mathbb{P}(A) \quad \text{e} \quad \mathbb{P}(B|A) = \mathbb{P}(B).$$

Como consequência imediata, temos

$$P(A) = P(A \cap B)/P(B),$$

e portanto

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

Definição 2.2. Diremos que dois eventos A e B em Ω são independentes se

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

caso contrário diremos que os eventos são dependentes.

2.2 LEI DOS GRANDE NÚMEROS

Já construímos uma definição mais geral de probabilidade que não depende da realização do experimento. Para fechar esta teoria de modo consistente precisamos que toda ela “converse” com o modelo a posteriori definido pelas frequências relativas de eventos. Para isso precisamos que, de algum modo, a noção de frequência relativa de um evento seja coerente com a probabilidade deste evento.

Para entender com mais exatidão o que pretendemos mostrar, lembre-se que pela hipótese frequentista a frequência relativa de ocorrência de um evento A em n repetições de um experimento se aproxima de $\mathbb{P}(A)$ quando n cresce. É este portanto o nosso objetivo.

Isso significa que quando n cresce a diferença entre a frequência relativa de A e $\mathbb{P}(A)$ se aproxima de 0. De maneira mais formal, chamando de S_n o número de vezes que o evento A foi observado em n repetições de um experimento, e fazendo $p = \mathbb{P}(A)$, queremos mostrar que quando $n \rightarrow \infty$

$$\left| \frac{S_n}{n} - p \right| \rightarrow 0.$$

Assim se L é o evento

$$L = \left\{ \left| \frac{S_n}{n} - p \right| \rightarrow 0, \text{ quando } n \rightarrow \infty \right\},$$

então uma possibilidade seria mostrar que

$$\mathbb{P}(L) = 1.$$

Este é um resultado que, apesar de verdadeiro, possui um prova conceitualmente pesada, que foge dos objetivos deste trabalho.

Vamos então buscar uma outra maneira de abordar o problema. Para isso podemos pensar que, para qualquer m que considerarmos, a diferença $\left| \frac{S_n}{n} - p \right|$ vai eventualmente ficar menor que $1/m$. Mas note que a diferença considerada é de fato um valor aleatório, e que mesmo que cheguemos muito próximo do valor de p , sempre existe a probabilidade de observarmos o evento A muitas vezes seguidas, e acabarmos nos distanciando novamente do valor que queremos.

Para deixar esta ideia mais clara, pense no lançamento de uma moeda não viciada. Suponha que depois de 1000 jogadas tenhamos observado cara 501 vezes. Isso significa que a diferença da frequência relativa e a probabilidade real é de $1/1000$, mas se observamos cara muitas vezes seguidas nos próximos lançamentos, a diferença vai aumentar, passando de $1/1000$!

O que vamos fazer então é mostrar que a probabilidade disto ocorrer diminui quando n cresce.

Ou seja, definindo o evento

$$\left\{ \text{a diferença entre } \frac{S_n}{n} \text{ e } p \text{ é maior que } \frac{1}{m} \right\} = \left\{ \left| \frac{S_n}{n} - p \right| > \frac{1}{m} \right\},$$

o que queremos é mostrar que

$$\mathbb{P} \left(\left| \frac{S_n}{n} - p \right| > \frac{1}{m} \right) \rightarrow 0$$

quando $n \rightarrow \infty$.

Esta é a chamada *Lei Fraca dos Grandes Números de Bernoulli* que demonstraremos a seguir.

Parte da importância deste resultado reside no fato de nos permitir a comparação de modelos. Ou seja, primeiro fazemos hipóteses sobre como deve ser o modelo a priori, e com este modelo em mãos podemos confirmar o resultado, repetindo o experimento diversas vezes.

Teorema 2.3 (Lei Fraca dos Grandes Números de Bernoulli). *Seja S_n o total de vezes que um evento A foi observado em n repetições independentes de um experimento aleatório e $p = \mathbb{P}(A)$ a probabilidade do evento A . Nestas condições vale que*

$$\mathbb{P}\left(\left|\frac{S_n}{n} - p\right| \geq \frac{1}{m}\right) \leq \frac{m^2}{4n}, \quad (2.1)$$

e portanto

$$\mathbb{P}\left(\left|\frac{S_n}{n} - p\right| > \frac{1}{m}\right) \rightarrow 0$$

quando $n \rightarrow \infty$.

Demonstração. Primeiro note que S_n é um valor aleatório no conjunto $\{0, 1, 2, \dots, n\}$. Note também que uma sequência de resultados do experimento onde A ocorreu k vezes tem probabilidade $p^k(1-p)^{n-k}$. Como existem um total de $\binom{n}{k}$ sequências onde isso acontece, segue que

$$\mathbb{P}(S_n = k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Observe agora que para todo $m > 0$ vale que

$$\left(\frac{k}{n} - p\right)^2 < \frac{1}{m^2} \Leftrightarrow \left|\frac{k}{n} - p\right| < \frac{1}{m} \Leftrightarrow -\frac{1}{m} < \frac{k}{n} - p < \frac{1}{m} \Leftrightarrow np - \frac{n}{m} < k < \frac{n}{m} + np,$$

e assim

$$\left(\frac{k}{n} - p\right)^2 \geq \frac{1}{m^2} \Leftrightarrow \left|\frac{k}{n} - p\right| \geq \frac{1}{m} \Leftrightarrow k \leq np - \frac{n}{m} \text{ ou } k \geq np + \frac{n}{m}.$$

Com isso, se

$$K_n = \left\{0, 1, \dots, np - \frac{n}{m}\right\} \cup \left\{np + \frac{n}{m}, \dots, n - 1, n\right\} = \left\{k : \left(\frac{k}{n} - p\right)^2 \geq \frac{1}{m^2}\right\}$$

não é difícil de ver que

$$\mathbb{P}\left(\left|\frac{S_n}{n} - p\right| \geq \frac{1}{m}\right) = \sum_{k \in K_n} \mathbb{P}(S_n = k).$$

Assim

$$\begin{aligned} \sum_{k=0}^n \left(\frac{k}{n} - p\right)^2 \mathbb{P}(S_n = k) &\geq \sum_{k=np - \frac{n}{m}}^{np + \frac{n}{m}} \left(\frac{k}{n} - p\right)^2 \mathbb{P}(S_n = k) \\ &\geq \sum_{k=np - \frac{n}{m}}^{np + \frac{n}{m}} \frac{1}{m^2} \mathbb{P}(S_n = k) \\ &= \frac{1}{m^2} \sum_{k=np - \frac{n}{m}}^{np + \frac{n}{m}} \mathbb{P}(S_n = k) \\ &= \frac{1}{m^2} \mathbb{P}\left(\left|\frac{S_n}{n} - p\right| \geq \frac{1}{m}\right) \end{aligned}$$

E portanto

$$\mathbb{P}\left(\left|\frac{S_n}{n} - p\right| \geq \frac{1}{m}\right) \leq m^2 \sum_{k=0}^n \left(\frac{k}{n} - p\right)^2 \mathbb{P}(S_n = k). \quad (2.2)$$

Precisamos agora calcular a soma do lado direito da desigualdade (2.2) acima. Para isso vamos primeiro lembrar algumas igualdades combinatórias cuja verificação deixamos a cargo do leitor interessado. São elas

$$k \binom{n}{k} = n \binom{n-1}{k-1}$$

e

$$k(k-1) \binom{n}{k} = n(n-1) \binom{n-2}{k-2}.$$

De fato a segunda igualdade pode ser obtida através da primeira fazendo

$$k(k-1) \binom{n}{k} = n(k-1) \binom{n-1}{k-1} = n(n-1) \binom{n-2}{k-2}.$$

Voltando observe primeiro que

$$\sum_{k=0}^n k \mathbb{P}(S_n = k) = \sum_{k=0}^n k \binom{n}{k} p^k (1-p)^{n-k},$$

e como para $k = 0$ o primeiro termo da soma é 0, temos

$$\begin{aligned}
 \sum_{k=0}^n k \mathbb{P}(S_n = k) &= \sum_{k=1}^n k \binom{n}{k} p^k (1-p)^{n-k} \\
 &= \sum_{k=1}^n n \binom{n-1}{k-1} p^k (1-p)^{n-k} \\
 &= np \sum_{k=1}^n \binom{n-1}{k-1} p^{k-1} (1-p)^{n-k} \\
 &= np \sum_{j=0}^{n-1} \binom{n-1}{j} p^j (1-p)^{(n-1)-j} \quad \text{fazendo } j = k-1 \\
 &= np(p+1-p)^{n-1} \quad \text{binômio de Newton} \\
 &= np
 \end{aligned}$$

Do mesmo modo mostramos (e deixamos como desafio para o leitor) que

$$\sum_{k=0}^n k(k-1) \mathbb{P}(S_n = k) = n(n-1)p^2.$$

Somando as duas igualdades acima encontramos que

$$\sum_{k=0}^n k^2 \mathbb{P}(S_n = k) = n(n-1)p^2 = np + n(n-1)p^2.$$

Voltando a soma original temos que

$$\begin{aligned}
 \sum_{k=0}^n \left(\frac{k}{n} - p \right)^2 \mathbb{P}(S_n = k) &= \sum_{k=0}^n \left(\frac{k^2}{n^2} - 2\frac{k}{n}p + p^2 \right) \mathbb{P}(S_n = k) \\
 &= \frac{np + n(n-1)p^2}{n^2} - 2p^2 + p^2 \\
 &= \frac{p + (n-1)p^2}{n} - p^2 \\
 &= \frac{p + (n-1)p^2 - np^2}{n} \\
 &= \frac{p(1-p)}{n}
 \end{aligned}$$

Segue assim que

$$\mathbb{P} \left(\left| \frac{S_n}{n} - p \right| \geq \frac{1}{m} \right) \leq \frac{m^2 p(1-p)}{n}, \quad (2.3)$$

e com isso mostramos que para todo $m > 0$

$$\mathbb{P} \left(\left| \frac{S_n}{n} - p \right| > \frac{1}{m} \right) \rightarrow 0,$$

quando n cresce. □

2.2.1 Erro nas aproximações

Além de mostrar que podemos aproximar probabilidades por frequências de repetições, o resultado que apresentamos acima permite que estimemos uma espécie de erro associado à aproximação feita. O segredo para isso está na equação (2.1).

Primeiro note que o erro cometido ao aproximarmos uma probabilidade p pela proporção $\frac{S_n}{n}$ é dado exatamente pela diferença

$$\left| \frac{S_n}{n} - p \right|.$$

Deste modo, o que a equação (2.1) nos diz é que a probabilidade de cometermos um erro maior que $\frac{1}{m}$ em n repetições de um experimento, é menor que $\frac{m^2}{4n}$.

Agora, se perguntarmos qual a probabilidade de cometermos um erro **menor** que $\frac{1}{m}$, então teríamos

$$\mathbb{P} \left(\left| \frac{S_n}{n} - p \right| < \frac{1}{m} \right) = 1 - \mathbb{P} \left(\left| \frac{S_n}{n} - p \right| \geq \frac{1}{m} \right) \geq 1 - \frac{m^2}{4n}.$$

Ou seja, com probabilidade maior $1 - \frac{m^2}{4n}$, o erro cometido em n repetições é menor que $\frac{1}{m}$.

Exemplo 2.3 (Aproximação da área de um círculo). Considere um quadrado de lado 1 e um círculo de diâmetro 1 inscrito neste quadrado. Sabemos que, escolhendo um ponto ao acaso dentro do quadrado, a probabilidade de escolhermos um ponto dentro do círculo é dada por

$$p = \frac{\pi}{4},$$

que é precisamente a área do círculo considerado. Queremos então saber qual a probabilidade de cometermos um erro maior que 0,1 quando aproximamos p pela frequência relativa de pontos dentro do círculo em 1000 repetições do experimento.

Solução: Queremos calcular

$$\mathbb{P} \left(\left| \frac{S_{1000}}{1000} - \frac{\pi}{4} \right| \geq \frac{1}{10} \right).$$

Sabemos de (2.1) que

$$\mathbb{P} \left(\left| \frac{S_{1000}}{1000} - \frac{\pi}{4} \right| \geq \frac{1}{10} \right) \leq \frac{m^2}{4n} = \frac{10^2}{4 \cdot 1000} = 0,025.$$

Concluimos que com probabilidade maior que 97,5%, cometemos um erro menor que 0,1 na aproximação.

Exemplo 2.4 (Aproximação da área de um círculo - Parte II). Voltemos ao problema anterior, agora com uma questão diferente. Já sabemos que com uma confiança de 97,5% podemos afirmar que a frequência relativa de pontos “dentro” do círculo, em 1000 repetições, difere de $\pi/4$ em no máximo 0,1. Suponha agora que estejamos satisfeitos com uma confiança de 90%, mas que queremos um erro de no máximo 0,05. Quantas repetições do experimento precisaríamos realizar?

Solução: Queremos encontrar n tal que

$$\mathbb{P} \left(\left| \frac{S_n}{n} - \frac{\pi}{4} \right| \geq \frac{1}{20} \right) \leq 0,1.$$

Sabemos de (2.1) que

$$\mathbb{P} \left(\left| \frac{S_n}{n} - \frac{\pi}{4} \right| \geq \frac{1}{20} \right) \leq \frac{m^2}{4n} = \frac{20^2}{4 \cdot n} = \frac{100}{n}.$$

Basta então que

$$\frac{100}{n} = \frac{1}{10},$$

ou seja, $n = 1000$.

Concluimos assim que com as mesmas 1000 repetições conseguimos um erro menor, mas com menos confiança.

Dos exemplos acima pode parecer que o total de repetições necessárias para que possamos usar a frequência relativa como aproximação para proporção real é muito grande, pelo menos para que tenhamos confiança na nossa aproximação. Mas o

que acontece na verdade é que a estimativa da probabilidade dada em (2.1) é muito grosseira, e a probabilidade real é bem menor que aquela.

Na prática o resultado que usamos para cálculo destes erros é o chamado *Teorema Central do Limite*, que afirma que

$$\mathbb{P} \left(\left| \frac{S_n - np}{\sqrt{np(1-p)}} \right| \leq a \right) \sim \Phi(a),$$

onde $\Phi(a)$ é a área definida em baixo da curva da função

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

Infelizmente a demonstração deste resultado está muito além do escopo deste trabalho.

2.3 ATIVIDADES PRÁTICAS

Propomos algumas atividades baseadas na Lei dos Grandes Números, com a finalidade de familiarizar o estudante de ensinos fundamental e médio com o conceito intuitivo de probabilidades.

- **Atividade 1 - Verificação de conceitos** A ideia desta atividade é tentar solidificar os conceitos envolvidos na Lei dos Grandes Números através da repetição de experimentos simples, com espaços amostrais também simples.

Podemos começar com o lançamento de uma moeda. Divida a turma em duplas e entregue uma moeda para cada par. Peça que cada dupla lance a moeda 50 vezes, anotando o resultado de cada lançamento. Em seguida peça que calculem a proporção de caras conseguidas.

Anote o resultado de cada dupla no quadro, analisando os resultados com a turma. É importante salientar que boa parte dos resultados devem ser diferentes de $1/2$, mas os resultados muito distantes de $1/2$ devem ser raros.

Repita esta mesma atividade com outros tipos de experimento, como dados ou cartas. Deste caso explore eventos mais complexos. Com cartas, por exemplo, peça que verifiquem o total de vezes que foi sorteado uma carta de valor par e vermelha, ou uma figura preta.

Peça que calculem a probabilidade dos eventos para comparar com os resultados dos experimentos.

- **Atividade 2 - Encontrando valores desconhecidos** Prepare urnas com 5 bolas pretas e 10 bolas brancas de igual tamanho e peso. As bolas podem ser feitas, por exemplo, de isopor, mas se não tiver bolas use qualquer tipo de objeto, e diferencie-os de alguma forma. O importante é que todos tenham a mesma dimensão e peso.

Separe os alunos em grupos de 2 ou 3 alunos. Informe aos alunos que dentro de cada urna existem 15 bolas, mas não o total de bolas brancas ou pretas.

Peça que cada grupo realize um total de 30 sorteios, com reposição da bola sorteada. A partir dos resultados dos sorteios os alunos devem então estimar o total de bolas brancas e pretas dentro da urna.

Verifique quantos grupos acertaram e quantos erraram. Discuta os resultados com a turma.

Esta atividade pode ser repetida com diversas configurações de bolas nas urnas, incluindo mais de duas cores.

É interessante chamar a atenção dos alunos para a semelhança entre este experimento e uma pesquisa eleitoral.

3

PROBABILIDADE GEOMÉTRICA

Um dos problemas que levantamos no caso do modelo clássico de Laplace foi sua incapacidade de incluir experimentos onde o espaço amostral é infinito. Existem várias maneiras de se definir este tipo de modelos, mas a teoria envolvida é demasiadamente complexa e técnica para os propósitos deste texto.

Escolhemos então explorar um modelo particular que se passa em espaços contínuos. Para o leitor que não conhece o conceito de espaço contínuo basta pensar em um intervalo da reta, ou mesmo uma região no plano ou no espaço, como um quadrado ou uma esfera.

3.1 PROBABILIDADE EM ESPAÇOS CONTÍNUOS

Qual é a probabilidade de se acertar, utilizando arco e flecha, no meio exato do alvo? Ou de forma mais geral: qual é a chance de se acertar em um ponto escolhido no alvo?

Se considerarmos que as probabilidades são igualmente espelhadas ao longo do alvo, então a probabilidade de acertarmos um ponto qualquer deveria ser zero, uma vez que existem infinitos pontos no alvo e estamos buscando apenas 1.

Mas, se é nula a possibilidade de se acertar um ponto, então, como podemos proceder na determinação das chances de um atirador de acertar no alvo, dado que sua superfície está repleta de pontos?

A ideia é simples: a probabilidade de acertar um ponto é nula, mas a probabilidade de acertar uma região do alvo, não é! Pensando nesta ideia Gnedenko (1912-1995)

propôs, em seu livro *The Theory of Probability*, o modelo de probabilidade geométrica, onde tomando Ω como uma do plano, a probabilidade de se escolher um ponto qualquer em uma certa subregião A é dada pela razão entre a área de A e a área de Ω . Ou seja,

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

Este conceito também se aplica a pontos localizados em segmentos de reta. Suponha que se deseja escolher um ponto qualquer no segmento CD contido em um segmento de reta AB . Desse modo, podemos dizer que a probabilidade do segmento CD é igual a razão entre os comprimentos de CD e de AB .

$$\mathbb{P}(CD) = \frac{\text{comprimento } CD}{\text{comprimento } AB}.$$

De modo análogo, podemos estender tal conceito para volumes, isto é, se Ω é uma região tridimensional,

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

Exemplo 3.1. Considere Ω como um quadrado de lado 2, e tome A como o círculo de raio 1 inscrito no quadrado. Pergunta-se, escolhendo-se um ponto ao acaso dentro do quadrado Ω , qual a probabilidade de este esteja dentro do círculo A .

Neste caso temos

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

Deste modo, independente da dimensão que estivermos trabalhando denotaremos o comprimento, área ou volume de uma região A por $|A|$.

Antes de mais nada é importante observar que este modelo proposto por Gnedenko também atende aos de Kolmogorov, pois

- Como $|A| \geq 0$, então $\mathbb{P}(A) \geq 0$;
- $\mathbb{P}(\Omega) = |\Omega|/|\Omega| = 1$;
- Se A_1, A_2, \dots, A_n são regiões em Ω com $A_i \cap A_j = \emptyset$, para $i \neq j$, então

$$|A_1 \cup \dots \cup A_n| = |A_1| + \dots + |A_n|,$$

e portanto

$$\mathbb{P}(A_1 \cup \dots \cup A_n) = \mathbb{P}(A_1) + \dots + \mathbb{P}(A_n).$$

Observe que, escondido nesta definição está a hipótese de que as probabilidades estão uniformemente distribuídas na região Ω . Isso porque a probabilidade de uma dada região depende apenas da sua área, e não da posição desta dentro de Ω . Em outras palavras, a probabilidade geométrica nos fornece espaço equiprovável infinito, pois se $|A| = |B|$, então $\mathbb{P}(A) = \mathbb{P}(B)$.

3.2 PROBLEMAS EM PROBABILIDADE GEOMÉTRICA

Nesta seção trataremos de dois problemas clássicos da probabilidade geométrica. O primeiro trata da divisão aleatório de um intervalo em três partes, e o segundo é o famoso problema da Agulha de Buffon.

3.2.1 *Dividindo um Segmento e Construindo Triângulos*

O primeiro problema que trataremos pode ser enunciado da seguinte forma. Dado um segmento de reta de segmento qualquer, divida o segmento aleatoriamente em três pedaços. Qual a probabilidade de que estes pedaços formem um triângulo?

De fato, como é comum em problemas envolvendo probabilidade, existem diversas maneiras distintas de resolver este problema, cada uma delas levando a um resultado distinto. A principal dificuldade, que é o que diferencia as diversas soluções, está em definir matematicamente o que significa “dividir o segmento aleatoriamente em três pedaços”. Podemos, por exemplo, escolher um ponto aleatório no intervalo $[0, L]$, onde L é o comprimento do segmento, e depois dividir o maior dos segmentos em dois da mesma forma.

Aqui vamos dividir o segmento da seguinte forma: Escolha independentemente dois pontos U_1 e U_2 uniformemente distribuídos no intervalo $[0, L]$. Tais pontos dividem o intervalo em três segmentos consecutivos, de comprimentos l_1 , l_2 e l_3 , respectivamente.

Para que os três segmentos formem um triângulo, é necessário e suficiente que seja satisfeita a desigualdade triangular. Ou seja,

$$l_1 + l_2 \geq l_3, \quad l_2 + l_3 \geq l_1 \quad \text{e} \quad l_1 + l_3 \geq l_2.$$

Chamemos de $V_M = \max\{U_1, U_2\}$ e $V_m = \min\{U_1, U_2\}$. Deste modo temos que

$$l_1 = V_m$$

$$l_2 = |U_2 - U_1| = V_M - V_m$$

$$l_3 = L - V_M$$

Agora, estudemos os três casos dados pela condição de existência de triângulos.

A primeira inequação $l_1 + l_2 \geq l_3$ nos fornece que

$$V_m + V_M - V_m \geq L - V_M$$

e portanto

$$V_M \geq \frac{L}{2}.$$

Segue que

$$\max\{U_1, U_2\} \geq \frac{L}{2}.$$

A segunda inequação $l_2 + l_3 \geq l_1$ nos diz que $V_M - V_m + L - V_M \geq V_m$, de onde concluímos que

$$V_m \leq \frac{L}{2}.$$

Segue assim que

$$\min\{U_1, U_2\} \leq \frac{L}{2}.$$

Da terceira inequação $l_1 + l_3 \geq l_2$ obtemos $V_m + L - V_M \geq V_M - V_m$, o que nos leva a

$$(V_M - V_m) \leq \frac{L}{2},$$

ou seja,

$$|U_2 - U_1| \leq \frac{L}{2}.$$

Deste modo, o evento $T = \{\text{os três segmentos formam um triângulo}\}$ pode ser escrito por

$$T = \left\{ \max\{U_1, U_2\} \geq \frac{L}{2} \right\} \cap \left\{ \min\{U_1, U_2\} \leq \frac{L}{2} \right\} \cap \left\{ |U_2 - U_1| \leq \frac{L}{2} \right\}.$$

Para estudar melhor este evento, vamos subir uma dimensão no nosso problema. Para isso faça $W = (U_1, U_2)$ um ponto aleatório no intervalo retângulo $R = [0, L] \times [0, L]$. Queremos nos convencer que W está uniformemente distribuído em R . Para isso considere dois subconjuntos $A, B \subset [0, L]$, de modo que $A \times B \subseteq R$. Observe agora que

$$W \in A \times B \iff U_1 \in A \text{ e } U_2 \in B,$$

de modo que

$$\mathbb{P}(W \in A \times B) = \mathbb{P}(\{U_1 \in A\} \cap \{U_2 \in B\}).$$

que por hipótese são eventos independentes. Segue assim que

$$\begin{aligned} \mathbb{P}(W \in A \times B) &= \mathbb{P}(U_1 \in A) \cdot \mathbb{P}(U_2 \in B) \\ &= \frac{|A|}{L} \cdot \frac{|B|}{L} \\ &= \frac{|A| \cdot |B|}{L^2} \end{aligned}$$

E portanto,

$$\mathbb{P}(W \in A \times B) = \frac{|A \times B|}{|R|}.$$

Como nem todo subconjunto de R pode ser escrito como um produto cartesiano de subconjuntos de $[0, L]$, a princípio isso não seria o suficiente para mostrar que W é um ponto aleatório uniforme em R , mas é o suficiente para nos convencer que isso é realmente verdade.

Sendo assim, vamos identificar os três eventos $\{\max\{U_1, U_2\} \geq \frac{L}{2}\}$, $\{\min\{U_1, U_2\} \leq \frac{L}{2}\}$ e $\{|U_2 - U_1| \leq \frac{L}{2}\}$ como regiões de R . Para identificar melhor estas regiões observe a Figura 1.

Com isso fica fácil ver que $\mathbb{P}(T) = \frac{1}{4}$.

3.2.2 A Agulha de Buffon

Conta a história que o conde de Buffon (Georges Louis Leclerc, 1707-1788) jogava o escafandador de seu cachimbo no chão da enfermaria durante uma internação. O escafandador tinha o tamanho exatamente igual ao da distância entre as bordas do soalho. Curioso, ele anotou a quantidade de vezes em que o escafandador tocava uma

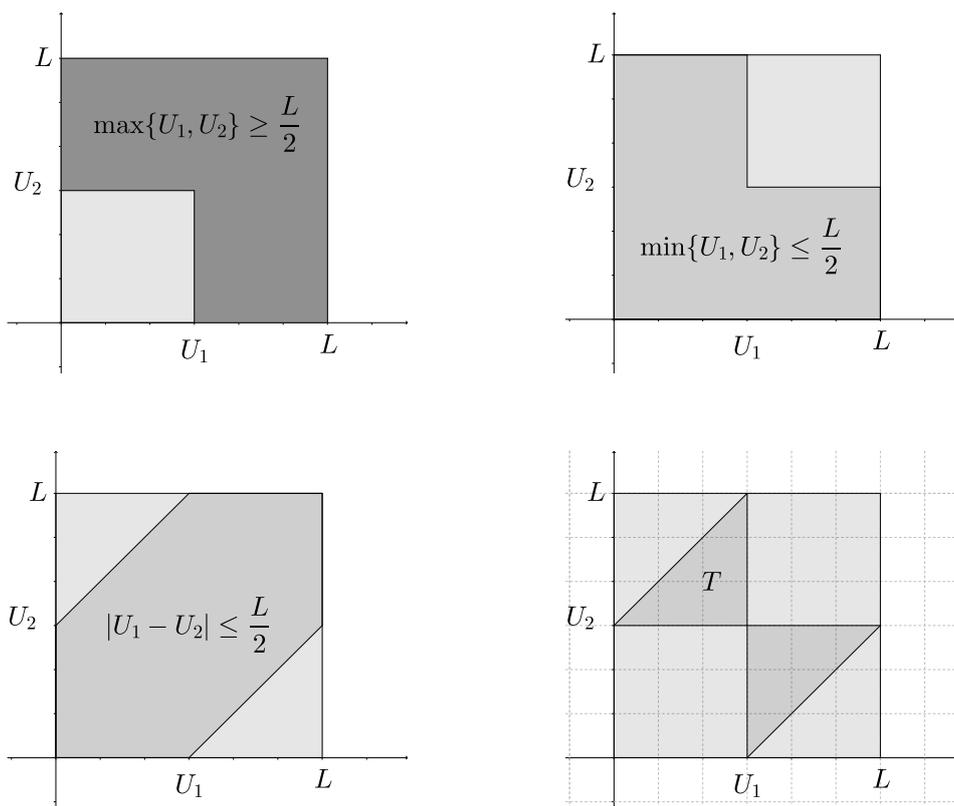


Figura 1: Descrição geométrica dos eventos que compõe o evento T

das bordas e o total de lançamentos realizados. Calculando a frequência relativa, chegou em uma razão próxima de $\frac{2}{\pi}$.

Estudemos o problema do conde.

Seja uma agulha de comprimento M e um plano α que contenha um feixe de retas paralelas que distam L entre si. Lançamos uma agulha aleatoriamente neste plano, e queremos saber qual a probabilidade da agulha tocar uma das retas paralelas.

Para entender melhor este problema, vamos definir algumas variáveis. Chame de θ o menor ângulo formado entre a agulha e a direção paralela às linhas, de modo que $0 \leq \theta \leq \pi/2$.

Defina por U a distância do centro da agulha até a linha mais próxima. Deste modo temos $U \in [0, L/2]$.

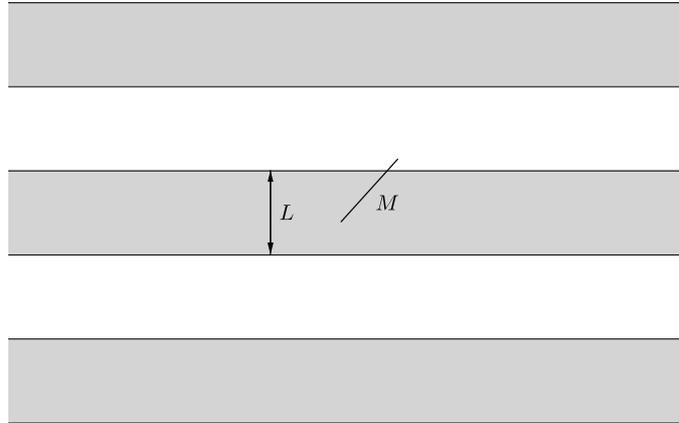


Figura 2: A Agulha de Buffon

É razoável supor que θ está uniformemente distribuído em $[0, \pi/2]$ e U uniformemente distribuído em $[0, L/2]$.

Observe agora que, para que a agulha cruze uma das linhas do feixe precisamos que (ver Figura 3)

$$U \leq \frac{M}{2} \sin \theta.$$

Ou seja,

$$\frac{2U}{M} \leq \sin \theta$$

Segue assim que

$$A = \{\text{agulha cruza uma linha}\} = \left\{ \frac{2U}{M} \leq \sin \theta \right\}.$$

Fazendo $2U/M = V$, é fácil se convencer que V está uniformemente distribuído em $[0, L/M]$. Assim, procedendo da mesma forma que no exemplo anterior, defina $W = (\theta, V)$ e observe que W é uniformemente distribuído em $[0, \pi/2] \times [0, L/M]$. Deste modo, para calcular a probabilidade do evento A precisamos calcular a área da região

$$R = \{(\theta, v) \in [0, \pi/2] \times [0, L/M] : v \leq \sin \theta\}.$$

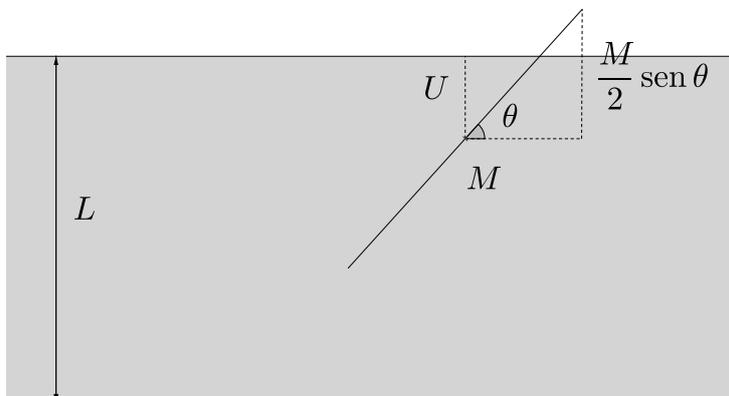


Figura 3: A Agulha de Buffon

Observe que a região acima depende diretamente dos valores de L e M . Como mostra a Figura 4, temos fundamentalmente duas descrições distintas. A primeira quando $L < M$ (agulha mais longa que os soalhos) e $L \geq M$ (agulha mais curta ou de mesmo tamanho que os soalhos).

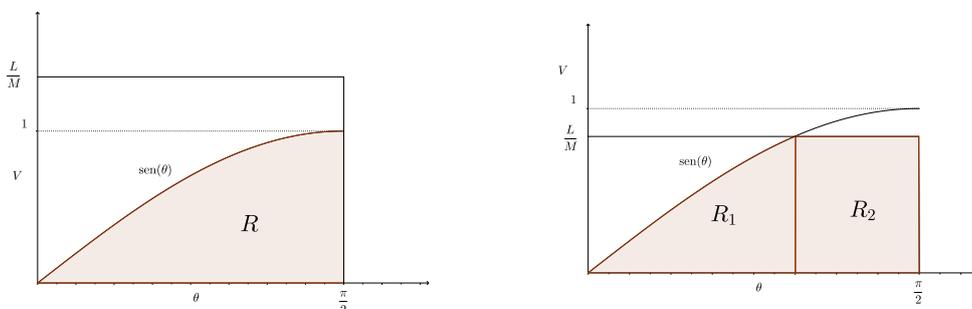


Figura 4: Descrição geométrica da região R no problema da Agulha de Buffon. A figura à esquerda corresponde ao caso $L \geq M$, e a outra ao caso $L < M$

A área destas regiões não são exatamente simples de se calcular apenas com ferramentas do ensino médio, e por isso passamos os cálculos para o apêndice. No

apêndice mostramos que no caso em que $L \geq M$ a área de R é 1. No caso $L < M$ não apresentamos os cálculos de $|R_1|$, deixando deste modo ao longo do texto.

Tomemos então os casos possíveis.

- $L = M$

Neste caso, temos

$$\mathbb{P}(A) = \frac{|R|}{\pi/2 \cdot (L/M)} = \frac{2}{\pi}.$$

- $L > M$

Agora, o que obtemos é

$$\mathbb{P}(A) = \frac{|R|}{\pi/2 \cdot (L/M)} = \frac{2M}{\pi L}.$$

- $L < M$

Neste caso teremos a área sob a curva dividida em duas áreas, já que $y = L/M$ corta o gráfico do $\sin \theta$. Temos

$$\mathbb{P}(A) = \frac{|R_1| + |R_2|}{(\pi/2) \cdot (L/M)} = \frac{|R_1| + (\pi/2 - x)(L/M)}{(\pi/2) \cdot (L/M)},$$

com $x = \arcsen(L/M)$.

O caso $L = M$ corresponde exatamente ao caso descrito na história, de modo que confirmamos os cálculos do conde!

3.3 ATIVIDADES PRÁTICAS

- **Atividade 1 - Jogando discos** Monte uma malha quadriculada de tamanho conhecido. Se quiser, aproveite um chão com azulejos quadrados. O objetivo aqui é jogar um objeto redondo e contar quantas vezes o disco corta um lado do quadrado.

Você pode utilizar qualquer objeto circular, como um CD ou uma moeda, desde que o diâmetro seja menor que o lado dos quadrados da malha. Repita este experimento diversas vezes e calcule com que frequência o objeto redondo cruzou a borda de algum quadrado.

Meça o raio do disco e o lado dos quadrados, e calcule a probabilidade teórica. Para isso note que, se chamarmos o lado dos quadrados de L e o raio do círculo de r , então a distância do centro do disco para alguma lateral do quadrado deve ser menor que o raio r do disco. Ou seja, para que o disco cruze algum dos quadrados, o centro deve cair em um quadrado menor cujo lado é $L - 2r$ (por quê?).

Compare o resultado teórico com as estimativas, discutindo com os alunos.

Repita o experimento com um disco de raio diferente e desconhecido, e peça que os alunos estimem o raio do disco usando a frequência relativa.

- **Atividade 2 - Estimando constantes desconhecidas** Repita o jogo do conde de Buffon para estimar o valor de π . Para isso será necessário alguma superfície com linhas paralelas equidistantes.

Uma ideia é desenhar linhas verticais em uma grande folha de cartolina. Desenhe as linhas com o espaçamento dado pelo comprimento de um palito de dente. Pegue uma caixa de palitos e jogue um a um em cima da cartolina. É importante que os palitos sejam jogados individualmente para evitar algum tipo de correlação entre as posições dos palitos na cartolina. Quanto maior a cartolina melhor, pois assim podemos assumir a hipótese de uniformidade sem grandes problemas.

Seguindo assim o experimento de Buffon, calcule a proporção de palitos que cruzaram alguma linha e use esta proporção para encontrar π . Repita este experimento várias vezes, discutindo os resultados.

Usando π e suas estimativas com 4 casas decimais, calcule o maior erro cometido nas diversas estimativas. Verifique também quantas estimativas estavam fora por uma diferença de 0,1 (ou qualquer outro valor que achar conveniente).

- **Atividade 3 - Estimando área de figuras** Nesta atividade, nosso objetivo é estimar a área de uma figura que não sabemos calcular. Para isso monte uma malha quadriculada, como na Atividade 1, e dentro de cada quadrado desenhe a mesma figura. É importante que cada figura fique igual às demais e na

mesma posição relativa dentro do seu quadrado. Atire um objeto pequeno (como um grão de arroz) repetidamente na malha e calcule a frequência com que o grão caiu dentro de uma das figuras desejadas. Use a Lei dos Grande Números para estimar a área desejada.

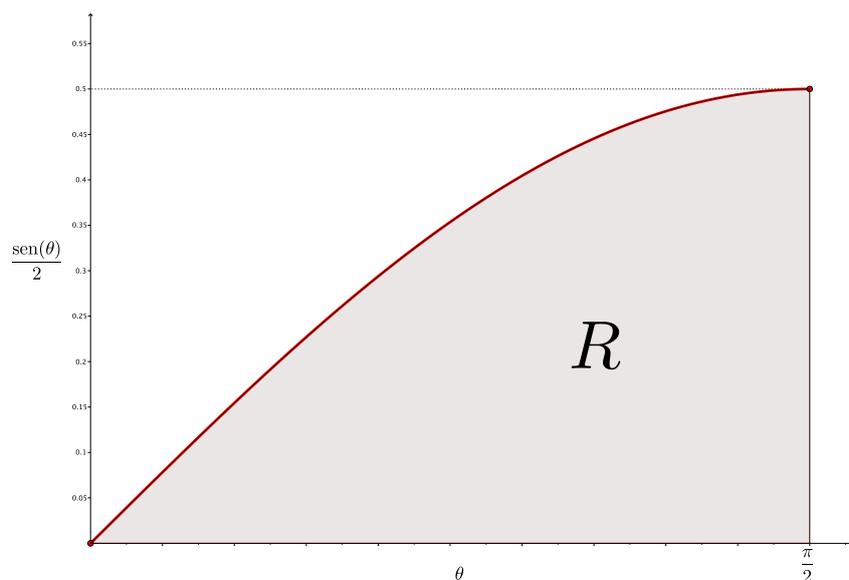
Divida os alunos em grupos e peça que cada grupo repita o mesmo experimento. Compare e discuta os resultados obtidos pelos diversos grupos.

Peça que os alunos desenvolvam um experimento baseado nesta atividade para estimar o valor de π .

A

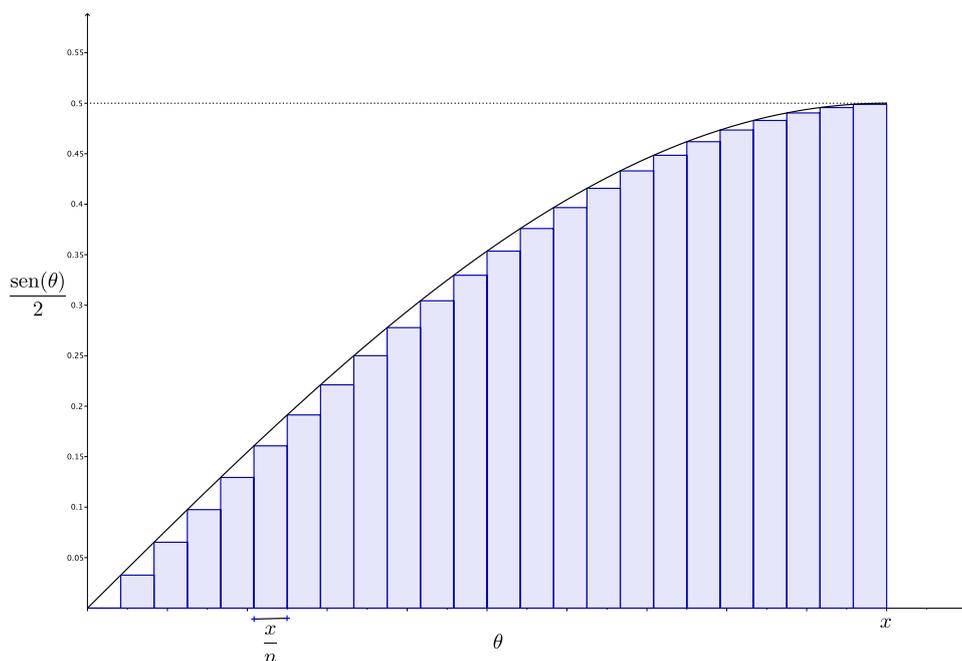
CÁLCULO DA ÁREA EM BAIXO DA CURVA DO SENO

Nosso objetivo neste capítulo é mostrar que a área embaixo do gráfico da função $\frac{\sin(x)}{2}$ entre 0 e $\pi/2$ é igual a $1/2$, como indicado na figura abaixo.



Para isso vamos aproximar a região estudada por uma sequência de retângulos cuja área sabemos calcular. Para ilustrar o que pretendemos fazer, considere a seguinte figura.

CÁLCULO DA ÁREA EM BAIXO DA CURVA DO SENO



Para desenhar os retângulos representados acima, primeiro divide o intervalo $[0, \pi/2]$ em n subintervalos $\pi/2n$. Os extremos destes intervalos estarão então nos pontos

$$0, \frac{\pi}{2n}, 2\frac{\pi}{2n}, 3\frac{\pi}{2n}, \dots, (n-1)\frac{\pi}{2n}, \frac{\pi}{2}.$$

Tomando cada um destes subintervalos como base, construímos retângulos de altura dada pelo valor de $f(x) = \sin(x)/2$ no extremo esquerdo do intervalo. Ou seja, o primeiro retângulo terá base em $[0, \pi/2n]$ e altura $\sin(0)/2 = 0$, o segundo terá base em $[\pi/2n, \pi/n]$ e altura $\sin(\pi/2n)/2$, e assim em diante. Deste modo, para cada $k = 0, 1, \dots, n-1$ teremos retângulos com base em

$$\left[k\frac{\pi}{2n}, (k+1)\frac{\pi}{2n}\right],$$

e altura

$$\sin\left(k\frac{\pi}{2n}\right).$$

Observe que

1. Os retângulos assim determinados ficam embaixo da curva;

2. Quanto maior o valor de n , melhor é a estimativa que teremos da área. De fato, a medida que n cresce a soma da área dos retângulos se aproxima da área que queremos encontrar.

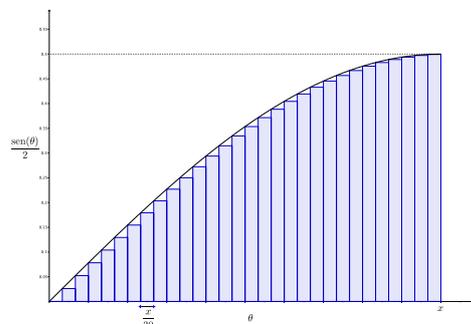
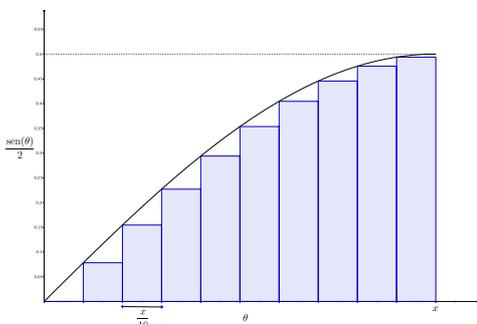


Figura 5: Aproximação por retângulos - $n = 10$ Figura 6: Aproximação por retângulos - $n = 30$

A área $A(n)$ da região determinada pelos retângulos é dada por

$$A(n) = \frac{\pi}{2n} \cdot \frac{\text{sen}(0)}{2} + \frac{\pi}{2n} \cdot \frac{\text{sen}(\pi/2n)}{2} + \dots + \frac{\pi}{2n} \cdot \frac{\text{sen}(k\pi/2n)}{2} + \dots + \frac{\pi}{2n} \cdot \frac{\text{sen}((n-1)\pi/2n)}{2}.$$

Ou seja,

$$A(n) = \frac{\pi}{4n} \sum_{k=0}^{n-1} \text{sen}(k\pi/2n). \quad (\text{A.1})$$

Como comentamos a medida que n cresce, o valor de $A(n)$ se aproxima da área R que queremos calcular. Em outras palavras, se $n \rightarrow \infty$ então $A(n) \rightarrow R$.

Nas próximas seções mostraremos que

$$\sum_{k=0}^{n-1} \text{sen}(k\theta) = \frac{1 - \cos(n\theta)}{2} \frac{\cos(\theta/2)}{\sin(\theta/2)} - \frac{\text{sen}(n\theta)}{2}, \quad (\text{A.2})$$

de onde segue, fazendo $\theta = \pi/2n$, que

$$\begin{aligned} 2A(n) &= \frac{\pi}{2n} \sum_{k=0}^{n-1} \text{sen}(k\pi/2n) \\ &= (1 - \cos(\pi/2)) \frac{\pi/4n}{\sin(\pi/4n)} \cos\left(\frac{\pi}{4n}\right) - \text{sen}(\pi/2) \frac{\pi}{4n}, \end{aligned}$$

e portanto

$$A(n) = \frac{1}{2} \left[\frac{\pi/4n}{\sin(\pi/4n)} \cos\left(\frac{\pi}{4n}\right) - \frac{\pi}{4n} \right] \quad (\text{A.3})$$

Para concluir precisamos descobrir o que acontece com $\frac{\pi/4n}{\sin(\pi/4n)} \cos\left(\frac{\pi}{4n}\right)$ quando n cresce. Para isso, volte sua atenção para a Figura 7, onde temos representado uma circunferência de raio 1 (círculo trigonométrico) e alguns pontos chave.

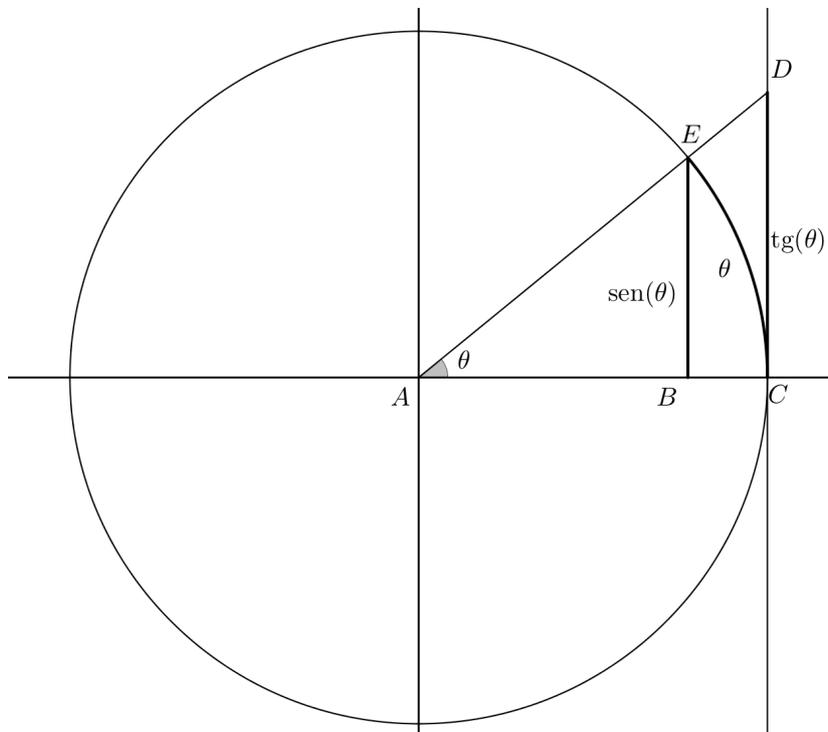


Figura 7: Relação entre θ , $\text{sen}(\theta)$ e $\text{tg}(\theta)$

Observe que

- $EB = \text{sen}(\theta)$;
- $DC = \text{tg}(\theta)$;
- O arco $\widehat{EC} = \theta r = \theta$,

e portanto

$$\text{sen}(\theta) \leq \theta \leq \text{tg}(\theta),$$

ou ainda

$$1 \leq \frac{\theta}{\text{sen}(\theta)} \leq \frac{1}{\cos(\theta)}$$

Assim, se θ se aproxima de 0 ($\theta \rightarrow 0$) temos $\cos(\theta) \rightarrow 1$ e

$$\frac{\theta}{\text{sen}(\theta)} \rightarrow 1.$$

Voltando agora para a equação (A.3), como $n \rightarrow \infty$ implica que $\frac{\pi}{4n} \rightarrow 0$, temos que $\cos\left(\frac{\pi}{4n}\right) \rightarrow 1$, $\frac{\pi/4n}{\sin(\pi/4n)} \rightarrow 1$ e

$$A(n) \rightarrow \frac{1}{2},$$

de onde concluímos que $R = \frac{1}{2}$.

CÁLCULO DE $A(n)$

Primeiro, considere uma P.G. de razão r e primeiro termo 1, dada por

$$1, r, r^2, r^3, r^4, \dots$$

Fazendo

$$S = 1 + r + \dots + r^{n-1},$$

temos

$$rS = r + r^2 + \dots + r^n.$$

Segue assim que

$$(1 - r)S = S - rS = (1 + r + \dots + r^{n-1}) - (r + r^2 + \dots + r^n) = 1 - r^n,$$

e portanto

$$S = \frac{1 - r^n}{1 - r}.$$

Ou seja,

$$\sum_{k=0}^{n-1} r^k = \frac{1 - r^n}{1 - r}.$$

Observe que os argumentos continuam valendo mesmo que $r = a + bi$ seja um número complexo.

CÁLCULO DA ÁREA EM BAIXO DA CURVA DO SENO

Lembrando que $(a + bi)^n$ é ainda um número complexo, e fazendo $(a + bi)^n = a_n + b_n i$ temos

$$\begin{aligned}
 \sum_{k=0}^{n-1} (a + bi)^k &= \frac{1 - (a + bi)^n}{1 - (a + bi)} \\
 &= \frac{1 - (a_n + b_n i)}{1 - (a + bi)} \\
 &= \frac{((1 - a_n) - b_n i)((1 - a) + bi)}{((1 - a) - bi)((1 - a) + bi)} \\
 &= \frac{(1 - a_n)(1 - a) + b_n b + ((1 - a_n)b - (1 - a)b_n)i}{(1 - a)^2 + b^2}
 \end{aligned}$$

E assim

$$\sum_{k=0}^{n-1} (a + bi)^k = \frac{(1 - a_n)(1 - a) + b_n b}{(1 - a)^2 + b^2} + \frac{((1 - a_n)b - (1 - a)b_n)}{(1 - a)^2 + b^2} i \quad (\text{A.4})$$

FORMA TRIGONOMÉTRICA DE UM NÚMERO COMPLEXO

Nosso próximo passo é relembrar como calcular a potência de um número complexo do tipo $(\cos(\theta) + i \sin(\theta))$. Neste intuito observe primeiro que

$$\begin{aligned}
 (\cos(k\theta) + i \sin(k\theta))(\cos(m\theta) + i \sin(m\theta)) &= \\
 &= \cos(k\theta) \cos(m\theta) - \sin(k\theta) \sin(m\theta) + i(\sin(k\theta) \cos(m\theta) + \sin(m\theta) \cos(k\theta)) \\
 &= \cos((k + m)\theta) + i \sin((k + m)\theta).
 \end{aligned}$$

Segue que

$$(\cos(\theta) + i \sin(\theta))^2 = (\cos(\theta) + i \sin(\theta))(\cos(\theta) + i \sin(\theta)) = \cos(2\theta) + i \sin(2\theta)$$

$$(\cos(\theta) + i \sin(\theta))^3 = (\cos(2\theta) + i \sin(2\theta))(\cos(\theta) + i \sin(\theta)) = \cos(3\theta) + i \sin(3\theta)$$

$$(\cos(\theta) + i \sin(\theta))^4 = (\cos(3\theta) + i \sin(3\theta))(\cos(\theta) + i \sin(\theta)) = \cos(4\theta) + i \sin(4\theta)$$

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

De onde segue que $(\cos(\theta) + i \sin(\theta))^n = \cos(n\theta) + i \sin(n\theta)$

Assim, como $(\cos(\theta) + i \sin(\theta))^n = \cos(n\theta) + i \sin(n\theta)$ segue de (A.4) que

$$\begin{aligned} \sum_{k=0}^{n-1} \cos(k\theta) + i \sum_{k=0}^{n-1} \sin(k\theta) &= \sum_{k=0}^{n-1} (\cos(k\theta) + i \sin(k\theta)) \\ &= \sum_{k=0}^{n-1} (\cos(\theta) + i \sin(\theta))^k \\ &= \frac{(1 - \cos(n\theta))(1 - \cos(\theta)) + \text{sen}(n\theta)\text{sen}(\theta)}{(1 - \cos(\theta))^2 + \text{sen}(\theta)^2} \\ &\quad + \frac{(1 - \cos(n\theta))\text{sen}(\theta) - (1 - \cos(\theta))\text{sen}(n\theta)}{(1 - \cos(\theta))^2 + \text{sen}(\theta)^2} i \end{aligned}$$

E portanto

$$\begin{aligned} \sum_{k=0}^{n-1} \sin(k\theta) &= \frac{(1 - \cos(n\theta))\text{sen}(\theta) - (1 - \cos(\theta))\text{sen}(n\theta)}{(1 - \cos(\theta))^2 + \text{sen}(\theta)^2} \\ &= \frac{(1 - \cos(n\theta))\text{sen}(\theta) - (1 - \cos(\theta))\text{sen}(n\theta)}{2 - 2 \cos(\theta)} \\ &= (1 - \cos(n\theta)) \frac{\text{sen}(\theta)}{2(1 - \cos(\theta))} - \frac{\text{sen}(n\theta)}{2} \end{aligned}$$

Para simplificar a análise que faremos a seguir, lembre-se agora que

$$\text{sen}(\theta) = 2\text{sen}(\theta/2) \cos(\theta/2)$$

e

$$\cos(\theta) = 1 - 2 \sin^2(\theta/2).$$

Segue assim que

$$\sum_{k=0}^{n-1} \sin(k\theta) = (1 - \cos(n\theta)) \frac{\sin(\theta/2) \cos(\theta/2)}{2 \sin^2(\theta/2)} - \frac{\sin(n\theta)}{2}$$

ou ainda

$$\sum_{k=0}^{n-1} \sin(k\theta) = \frac{1 - \cos(n\theta)}{2} \frac{\cos(\theta/2)}{\sin(\theta/2)} - \frac{\sin(n\theta)}{2},$$

mostrando a equação (A.2).

BIBLIOGRAFIA

- [1] KATAOKA, V. Y.; RODRIGUES, A.; DE OLIVEIRA, M. S. Utilização do conceito de probabilidade geométrica como recurso didático no ensino de estatística. Notas para minicurso.
- [2] KLAIN, D. A.; ROTA, G.-C. *Introduction to geometric probability*. Cambridge University Press, 1997.
- [3] FELLER, W. *An introduction to probability theory and its applications*. Wiley, January 1968. v. 1.
- [4] ROSS, S.; PERTENCE, A.; DE CONTI, A. *Probabilidade: Um curso moderno com aplicações*. Bookman Companhia ed, 2010.
- [5] LIMA, E.; CARVALHO, P.; WAGNER, E.; MORGADO, A. *A matemática do ensino medio*. Coleção do Professor de Matemática. Sociedade Brasileira de Matemática, 2004.
- [6] ROSS, S. *Introduction to probability models*. Elsevier Science, 2006.
- [7] MAGALHÃES, M. *Probabilidade e variáveis aleatórias*. Edusp, 2006.