# Problemas contra-intuitivos como motivadores do estudo de conceitos de probabilidade no ensino médio

***Sérgio Luiz Daltoso Junior***

SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito:

Assinatura: ______________________

**Sérgio Luiz Daltoso Junior**

# Problemas contra-intuitivos como motivadores do estudo de conceitos de probabilidade no ensino médio

Dissertação apresentada ao Instituto de Ciências Matemáticas e de Computação – ICMC-USP, como parte dos requisitos para obtenção do título de Mestre – Programa de Mestrado Profissional em Matemática. *VERSÃO REVISADA*

Área de Concentração: Matemática

Orientadora: Profa. Dra. Geraldine Goes Bosco

**USP – São Carlos**
**Junho de 2016**

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi
e Seção Técnica de Informática, ICMC/USP,
com os dados fornecidos pelo(a) autor(a)

D152p

Daltoso Junior, Sérgio Luiz
Problemas contra-intuitivos como motivadores do estudo de conceitos de probabilidade no ensino médio / Sérgio Luiz Daltoso Junior; orientadora Geraldine Goes Bosco. - São Carlos - SP, 2016.
96 p.

Dissertação (Mestrado - Programa de Pós-graduação em Mestrado Profissional em Matemática em Rede Nacional) - Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, 2016.

1. Educação Matemática. 2. Ensino de Probabilidade. 3. Independência. 4. Ensino Médio. I. Bosco, Geraldine Goes, orient. II. Título.

**Sérgio Luiz Daltoso Junior**

# Counterintuitive problems as a motivating study of probabilistic concepts in high school math classes

Master dissertation submitted to the Instituto de Ciências Matemáticas e de Computação – ICMC-USP, in partial fulfillment of the requirements for the degree of the Master – Program in Mathematics Professional Master. *FINAL VERSION*

Concentration Area: Mathematics

Advisor: Profa. Dra. Geraldine Goes Bosco

**USP – São Carlos**
**June 2016**

*Esse trabalho é dedicado àqueles que alguma vez se sentiram diferentes das outras pessoas, como eu já me senti.*

*É possível fazer a diferença no mundo: basta acreditar em si mesmo.*

# AGRADECIMENTOS

Agradeço ao meu esposo, André, pelo incentivo e companhia durante a realização deste trabalho, especialmente por ter acreditado em mim quando eu mesmo apresentava as minhas dúvidas.

À minha orientadora, Geraldine, pela paciência, dedicação e compreensão, especialmente nas madrugadas que me acompanhava online e pelo apoio nos momentos mais difíceis.

Aos meus pais, Sergio e Liliane, e aos meus irmãos, Gabriel e Mariana, pelo incentivo à realização deste curso de mestrado.

A todos os demais familiares, por estarem ao meu lado e terem compreendido a necessidade das minhas faltas em compromissos da família devidos aos compromissos do mestrado.

Aos meus amigos, pelo apoio moral e incentivo durante todo o curso deste trabalho.

Aos meus alunos, que a cada dia me dão um sopro novo de esperança na educação.

Aos meus colegas de trabalho e à direção dos colégios onde trabalhei durante a execução deste trabalho, pelos auxílios oferecidos quando necessário.

Aos colegas e docentes do programa, que estiveram comigo até aqui.

Aos meus professores de matemática, que me inspiraram a escolher essa carreira.

Aos membros da Comissão de Pós-Graduação e aos funcionários pela atenção e dedicação.

*“Não haveria criatividade*
*sem a curiosidade que nos move*
*e que nos põe pacientemente impacientes*
*diante do mundo que não fizemos,*
*acrescentando a ele algo que fazemos.”*
*(Paulo Freire)*

# RESUMO

DALTOSO JUNIOR, S. L.. **Problemas contra-intuitivos como motivadores do estudo de conceitos de probabilidade no ensino médio**. 2016. 96 f. Dissertação (Mestrado – Programa de Mestrado Profissional em Matemática) – Instituto de Ciências Matemáticas e de Computação (ICMC/USP), São Carlos – SP.

Probabilidade é uma das áreas onde mais se encontram aplicações diárias da Matemática em nossas vidas. Muitas vezes, porém, o ensino de probabilidade no Ensino Médio é desestimulante para os alunos. Neste trabalho, estudaram-se problemas cujas respostas podem ser contrárias à intuição para que fosse possível sugerir sua adoção por professores de Matemática do Ensino Médio como forma de motivar o estudo da Probabilidade em seus alunos. Para compreender esses problemas, realizou-se uma revisão teórica sobre a Teoria de Probabilidade, através da qual foi elaborado um texto que aborda tais conceitos em uma linguagem acessível a professores de Matemática do Ensino Médio. Ao final, são disponibilizadas algumas sugestões e orientações para a utilização desses problemas em sala de aula, que podem ser adaptadas ao cotidiano e às turmas de cada professor de Matemática. Espera-se, dessa maneira, que este trabalho sirva como contribuição à prática desses docentes em sala de aula.

**Palavras-chave:** Educação Matemática, Ensino de Probabilidade, Independência, Ensino Médio.

# ABSTRACT

DALTOSO JUNIOR, S. L.. **Problemas contra-intuitivos como motivadores do estudo de conceitos de probabilidade no ensino médio**. 2016. 96 f. Dissertação (Mestrado – Programa de Mestrado Profissional em Matemática) – Instituto de Ciências Matemáticas e de Computação (ICMC/USP), São Carlos – SP.

Probability is one of the areas where we can most find daily applications of Mathematics in our lives. Many times, however, Probability High School classes are not stimulating for the students. In this text, we have studied some probability problems whose answers are not intuitive. It make possible to suggest their adoption by High School's Math teachers as a way to motivate their students to study Probability. To understand these problems, we have made a theoretical review in Probability Theory. With this, we write a text who talks about Probability concepts in an accessible language for High School's Math teachers. At last, we offer some suggestions and guidelines to the use of these problems in classroom, which can be adapted to the daily life of each Math teacher classes. It is expected, therefore, that this text could be a contribution for Math teachers in their classrooms.

**Key-words:** Mathematics Educacion, Teaching of Probability, Independence, High School.

# LISTA DE ILUSTRAÇÕES

# LISTA DE TABELAS

# LISTA DE ABREVIATURAS E SIGLAS

PCNEM . . Parâmetros Curriculares do Ensino Médio
PCNMat . . Parâmetros Curriculares Nacionais de Matemática

# SUMÁRIO

CAPÍTULO

1

# INTRODUÇÃO

## 1.1 Apresentação do problema de pesquisa

No âmbito da Educação Matemática, alguns dos principais conceitos a serem desenvolvidos no Ensino Básico são aqueles relacionados à Probabilidade. Com aplicações nas mais variadas áreas do conhecimento, a Probabilidade é fundamental para que os alunos compreendam o mundo ao seu redor, pois diversas situações da vida diária envolvem o pensamento probabilístico. Um outro fator que nos traz a importância da probabilidade no Ensino Básico está relacionado à visão de mundo da Matemática pelos alunos, pois a Probabilidade é um dos poucos conceitos matemáticos que permitem o desenvolvimento da ideia do incerto, daquilo que se pode prever mas, muitas vezes, não se pode ter certeza.

Apesar do exposto acima, muitas vezes observa-se que a Probabilidade é um assunto do qual muitos alunos do ensino médio acabam demonstrando desinteresse. Esse desinteresse pode ser justificado pela falta de estímulo ao estudo dessa área da Matemática. Muitas vezes, a Probabilidade é exposta aos alunos como um conceito abstrato com aplicações meramente mecânicas da teoria na resolução de exercícios. Outro fator que pode contribuir para tal desinteresse é a formação do professor de matemática, quando a falta de conhecimento mais profundo em Probabilidade pode se refletir em aulas massivamente expositivas e teóricas, tratando esse conceito como um mero jogo de números. O desinteresse pela Probabilidade pode, também, afetar a apropriação e o desenvolvimento de conteúdos em outras áreas do conhecimento, como o estudo da Genética, em Biologia, que desenvolve boa parte de seus conteúdos, habilidades e competências fazendo o uso de conceitos de Probabilidade.

A compreensão da importância da Probabilidade na formação do indivíduo cidadão, como exposto acima, levou o autor dessa pesquisa a escrever essa dissertação. Como não contara com disciplinas de Probabilidade ou Estatística em sua graduação universitária, o autor dessa pesquisa percebeu que as referências teórias que encontrou, ao lidar com o ensino de

Probabilidade em sala de aula, eram apenas aquelas que estudara durante o ensino médio. Durante o desenvolvimento das disciplinas deste programa de mestrado profissional, ao observar a necessidade de aprofundar seus conhecimentos nessa área, o autor viu a oportunidade de se lançar um desafio pessoal ao trabalhar com Probabilidade. Inicialmente, o objetivo do trabalho era tratar de Probabilidade em Genética, buscando aproximar a Matemática da Biologia. Entretanto, após iniciar os estudos teóricos, o autor tomou conhecimento de alguns problemas interessantes que nunca tinha visto serem explorados no Ensino Básico, apesar de conhecer professores que já haviam travado conhecimento com problemas similares. Ao descobrir tais problemas, o objetivo do trabalho foi alterado e autor se propôs a criar alguns textos explicativos que pudessem ser lidos e compreendidos por outros professores de matemática, a fim de que servissem como apoio para a elaboração de aulas envolvendo Probabilidade para o Ensino Médio e também pudessem ser usados, de alguma forma, para motivar os alunos.

Dentro desse contexto, o problema de pesquisa delineou-se: *Como contibuir com professores de Matemática do Ensino Médio para que eles possam ter conhecimento e compreensão de alguns problemas contra-intuitivos a fim de utilizá-los como motivadores para o estudo de Probabilidade em sala de aula?* Com o problema de pesquisa em mãos, traçamos nossos objetivos.

## 1.2 Objetivos

### *Objetivo geral*

- Elaborar um texto explicativo, em linguagem acessível para professores de matemática, que envolva conceitos da Teoria de Probabilidade, problemas contra-intuitivos e sugestões para que seja utilizado como norteador do desenvolvimento de alguns conteúdos ou motivadores para o estudo dessa área da Matemática.

### *Objetivos específicos*

- Desenvolver bases teóricas em probabilidade;
- Elaborar um texto expondo conceitos básicos da Teoria de Probabilidade a nível superior com linguagem acessível a professores do Ensino Médio;
- Estudar alguns problemas contra-intuivos envolvendo Probabilidade;
- Escrever um texto explicativo desses problemas voltado a professores do Ensino Médio;
- Desenvolver sugestões e orientações para professores de Matemática do Ensino Médio utilizarem os problemas estudados em suas aulas.

Para atingir esses objetivos, foi feita uma revisão bibliográfica relativa à maneira como o ensino de Probabilidade na Educação Básica, que apresentamos a seguir na Seção 1.3. Em seguida, realizou-se um estudo aprofundado de conceitos teóricos de Probabilidade. A partir disso, foi elaborado um texto explicativo de tais conceitos com o propósito de ser acessível a professores do Ensino Médio, apresentado no Capítulo 2. Na sequência, são apresentados e analisados alguns problemas contra-intuitivos de probabilidade no Capítulo 3 para, no Capítulo 4, oferecermos algumas sugestões para que professores de Matemática possam utilizar esses conceitos e problemas em suas aulas, adaptando-os a seu cotidiano e suas turmas. Ao final, apresentamos nossas considerações finais, reunidas no Capítulo 5.

## 1.3 O Ensino de Probabilidade nos documentos oficiais

A Probabilidade é um tópico de grande importância em carreiras profissionais de todas as áreas. O ensino de Probabilidade, aliado ao de Estatística, permite a discussão, dentro do ramo da Matemática, de aspectos do dia a dia e de algumas situações pertinentes do mundo em que vivemos. Ao apresentar a Matemática aos alunos, usualmente fica-se estabelecida uma disciplina rígida, rigorosa e exata. Mas a vida cotidiana não é assim: nada é certo, nem sempre o que se espera é o que acontece. Quanto a isso, tem-se que

> A visão do mundo estocástico permite adotar um ponto de vista no qual a aleatoriedade é percebida como um aspecto fundamental, objetivo e real, sendo a utilização dos métodos da teoria da probabilidade necessária para reduzir o caos de um único e imprevisível evento a um padrão mais previsível. (HURTADO; COSTA, 1999 apud NOGUEIRA; BRISOLA, 2013, p. 1930)

A Probabilidade, portanto, mostra-se importante para aproximar a ideia da Matemática como ciência da vida, pois questionamentos como risco e incerteza são inerentes ao ensino-aprendizagem e ao desenvolvimento desse ramo da Matemática, desde seu início. Morgado *et al.* (1991, p. 6) contam que tradicionalmente relaciona-se o início da Teoria das Probabilidades a Blaise Pascal e Pierre de Fermat, no século XVII, envolvendo um jogo de cartas. Apesar disso, há várias outras referências de estudos envolvendo probabilidade muito antes desses dois matemáticos (MORGADO *et al.*, 1991, p. 6), sendo elas principalmente ligadas à análise de situações envolvendo jogos. Esses mesmos autores também afirmam que o desenvolvimento da *Análise Combinatória*, outra importante área da Matemática, se deu, em grande parte, devido à necessidade de resolver problemas de contagem originados em situações da Teoria das Probabilidades.

Os Parâmetros Curriculares Nacionais de Matemática (PCNMat), levavam em conta a importância da aleatoriedade para o ensino de Probabilidade, como apontam Hurtado e Costa (1999 apud NOGUEIRA; BRISOLA, 2013, p. 1930), ao propor novas abordagens e metodologias como suporte à organização e ao planejamento dos docentes brasileiros. O ensino da Probabilidade e da Estatística, nos PCN, aparece inserido no bloco de conteúdos denominado *"Tratamento da Informação"* e é justificado pela demanda social, por sua constante utilização na sociedade atual, pela necessidade de o indivíduo compreender as informações veiculadas, tomar decisões e fazer previsões que influenciam sua vida pessoal e em comunidade (BRASIL, 1997). Nesse bloco, além das noções de Estatística e Probabilidade, destacam-se também as noções de Análise Combinatória (BRASIL, 1997). Ao aproximar os conteúdos de Probabilidade e Combinatória, a proposta dos PCN segue a tendência da década de 1990 para o ensino de probabilidade. Gonçalves (2004, p. 122), através da análise de orientações institucionais publicadas a partir da década de 1970, conclui que o Ensino de Probabilidade no Brasil ocorreu, nas décadas de 70, 80 e 90, por meio das abordagens Clássica e Axiomática, havendo variações apenas nos tipos de tarefas e técnicas apresentadas como exercícios ou exemplos: na década de 70, utilizaram-se as técnicas baseadas na Teoria de Conjuntos para a resolução de problemas enquanto na década de 90, quando foram elaborados os PCN, as técnicas foram ligadas à Análise Combinatória, sendo os anos 80 um período de transição, quando utilizou-se tanto da Teoria de Conjuntos quanto da Análise Combinatória para justificar as técnicas de resolução de problemas envolvendo o cálculo de probabilidades.

Os PCN estabelecem que a principal finalidade para o estudo de Probabilidade é permitir que

> (...) o aluno compreenda que grande parte dos acontecimentos do cotidiano são de natureza aleatória e é possível identificar prováveis resultados desses acontecimentos. As noções de acaso e incerteza, que se manifestam intuitivamente, podem ser exploradas na escola, em situações nas quais o aluno realiza experimentos e observa eventos (BRASIL, 1997, p. 96)

Nos Parâmetros Curriculares do Ensino Médio (PCNEM), aponta-se que a Probabilidade devem ser explorada junto à Estatística, como um conjunto de ideias e procedimentos que permitem aplicar a Matemática em questões do mundo real, mais especialmente aquelas provenientes de outras áreas (BRASIL, 2000). Por entender que a Probabilidade e a Estatística lidam com dados e informações em conjuntos finitos e que utilizam procedimentos que permitem controlar com certa segurança a incerteza e a mobilidade desses dados, os PCNEM (BRASIL, 2000) reafirmam, como os PCNMat (BRASIL, 1997), que o estudo da Análise Combinatória deve ser apenas parte instrumental desse tema.

Dantas (2013, p. 13 a 15) explica que baseado nessas orientações, foi criado em 1997 o ENEM - Exame Nacional do Ensino Médio, inicialmente concebido para avaliar os alunos que findavam seus estudos nessa etapa da Educação Básica. Em 2009, o ENEM passou por uma reformulação e, desde então, passa a ser um importante meio de ingresso no Ensino Superior, uma vez que é utilizado por inúmeras Universidades e Instituições de Ensino Superior como um dos instrumentos de seleção de alunos, muitas vezes em substituição aos vestibulares tradicionais (DANTAS, 2013). O ENEM caracteriza-se por apresentar uma forma inovadora de avaliação, que leva em consideração o desenvolvimento de habilidades e competências pelo aluno. Desde sua implementação, o tema Probabilidade é recorrente em suas provas. Depois de sua reestruturação, o ENEM passou a adotar uma Matriz de Referência (BRASIL, 2009) para a área do conhecimento identificada como *"Matemática e suas Tecnologias"* na qual a Probabilidade, aliada à Estatística, recebeu tratamento importante, passando a fazer parte de uma das sete competências da área analisadas no exame, a saber

> Competência de área 7 – Compreender o caráter aleatório e não determinístico dos fenômenos naturais e sociais e utilizar instrumentos adequados para medidas, determinação de amostras e cálculos de probabilidade para interpretar informações de variáveis apresentadas em uma distribuição estatística.
>
> H27 – Calcular medidas de tendência central ou de dispersão de um conjunto de dados expressos em uma tabela de frequências de dados agrupados (não em classes) ou em gráficos.
>
> H28 – Resolver situação-problema que envolva conhecimentos de Estatística e Probabilidade.
>
> H29 – Utilizar conhecimentos de Estatística e Probabilidade como recurso para a construção de argumentação.
>
> H30 – Avaliar propostas de intervenção na realidade utilizando conhecimentos de Estatística e Probabilidade.(BRASIL, 2009, p. 7)

Ainda segundo Brasil (2009, p. 1), um aluno de ensino médio deve possuir a competência de enfrentar situações-problema, ou seja: selecionar, organizar, relacionar e interpretar dados e informações representadas de diferentes formas para, então, tomar decisões. Visando que tal competência seja desenvolvida, o Currículo Oficial de Matemática do estado de São Paulo (SÃO PAULO, 2011) sugerem que Análise Combinatória e Probabilidade sejam abordadas na segunda série do Ensino Médio e a Estatística, na terceira série. É importante salientar que, embora tais conteúdos não apareçam explicitamente na distribuição de outras disciplinas, seu estudo pode ser feito de maneira interdisciplinar em praticamente qualquer assunto abordado.

Quanto ao tratamento de noções probabilísticas e estatísticas, os PCN (BRASIL, 1998) indicam que a coleta, a organização e descrição de dados são procedimentos utilizados com muita frequência na resolução de problemas e estimulam as crianças a fazer perguntas, estabelecer relações, construir justificativas e desenvolver o espírito de investigação. É sugerido que se desenvolvam atividades relacionadas a assuntos de interesse dos alunos, com proposta de observação de acontecimentos, que promovam situações para se fazerem previsões a fim de que algumas noções de probabilidade sejam desenvolvidas. Também é sugerido que se desenvolva o raciocínio estatístico e probabilístico através da exploração de situações de aprendizagem que levem o aluno a coletar, organizar e analisar informações, formular e fazer inferências convincentes, tendo por base a análise de dados organizados em representações matemáticas diversas. A Teoria da Probabilidade e a Estatística não envolvem apenas números e não se restringem ao ensino de Matemática, devido o fato de englobar diversas áreas do conhecimento, como saúde, política, economia, entre outras. É necessário levar em consideração as rápidas mudanças que vêm ocorrendo no mundo de hoje, e que para isso é fundamental o conhecimento da Probabilidade e da Estatística, para podermos agilizar a tomada de decisões e fazermos previsões. Precisamos saber interpretar resultados criticamente, organizar e representar dados, assim como obter conclusões. Dessa forma, o ensino de Probabilidade e Estatística auxilia na formação do cidadão, ajudando-o a ser mais crítico, a saber formar suas opiniões diante de determinadas situações.

Apesar das orientações oficiais, percebe-se, segundo Carvalho e Oliveira (1999), que os conceitos probabilísticos frequentemente não são estudados nos Ensinos Fundamental e Médio. Segundo os mesmos autores, quando estes conceitos são considerados, sua abordagem é reduzida a uma resolução mecânica de exercícios padrões: na maioria das vezes é suficiente aplicar uma fórmula. Lopes (2003) destaca que o estudo da Teoria da Probabilidade não deve ser baseado somente na definição matemática, já que a maior dificuldade dos alunos não está na definição dos conceitos, mas sim, no modo como o conceito é interpretado e aplicado apropriadamente e em situações específicas. É necessário, portanto, que os alunos desenvolvam um pensamento probabilístico e para que isso aconteça, segundo Gaffuri (2012), é importante conhecer suas particularidades além das concepções de azar e aleatoriedade. Quando isso não acontece, como aponta Silva (2012), os alunos podem apresentar problemas relacionados ao ensino-aprendizagem do conceito de probabilidade no Ensino Médio, a saber:

> (...) a. A ausência de abordagem no processo de ensino-aprendizagem de noções que compõem o campo conceitual probabilístico:
> I. Experimentos determinísticos
> II. Características de um experimento aleatório
> III. Noção de Acaso
> IV. Espaços amostrais não equiprováveis
> b. A abordagem exclusiva da visão laplaciana (clássica) de probabilidades, sem qualquer referência, portanto, à visão freqüentista de probabilidades. Tal aspecto acaba por proporcionar aos alunos apenas uma das visões, uma das faces da teoria probabilística.
> c. A abordagem de noções probabilísticas (Evento, Espaço Amostral, Definição de Probabilidade) utilizando-se a terna "definição-exemplo-exercício", em contraposição à proposta na qual, partindo-se de uma atividade ou situação-problema, atinge- se na seqüência a formalização do conceito.
> d. A abordagem de noções probabilísticas utilizando-se apenas a definição seguida de exemplos, sem qualquer atividade "complementar"(exercícios, testes, etc) com o intuito de retomar e aprofundar as noções estudadas:
> I. Tipos de experimentos
> II. Experimentos aleatórios
> III. Tipos de eventos
> IV. Noções históricas da Teoria das Probabilidades. (SILVA, 2012, p. 18 e 19)

Ao apresentar aos alunos a Probabilidade aplicada a situações presentes no cotidiano deles, eles poderão entender que este tópico da Matemática, além de interdisciplinar, é intuitivo ao ser humano. Os alunos devem ser motivados antes de serem apresentados a novas termologias e propriedades da Teoria da Probabilidade. Dessa forma, aquilo que seria apenas um cálculo teórico passa a tomar um verdadeiro significado na vida do aluno e, consequentemente, em sua aprendizagem. Entender como surgiu o estudo das probabilidades e como ele se desenvolveu ao longo dos anos, percebendo que houve um processo histórico que se iniciou com o cálculo voltado para previsão de vitórias de jogos de azar, contribui para que os alunos tenham uma visão da probabilidade como área viva e aplicada da Matemática, possibilitando que os mesmos enxerguem a probabilidade nos mais diversos ramos dela e de outras ciências como, por exemplo, Economia, Política, Medicina e Biologia.

CAPÍTULO

2

# ASPECTOS TEÓRICOS

Neste capítulo, apresentaremos alguns tópicos relativos à Teoria de Probabilidade. O objetivo desse capítulo é, além de fornecer as bases teóricas para podermos compreender os problemas do Capítulo 3, expor conceitos de Probabilidade a nível superior em uma linguagem acessível a professores de Matemática do Ensino Médio. Com isso, nos propomos a contribuir com a formação continuada desses professores e oferecer instrumentos conceituais para a elaboração de suas aulas.

## 2.1 Conceitos Iniciais

Nos dias atuais, há uma grande quantidade de jogos à nossa disposição, como os inúmeros jogos de loteria (*MegaSena, Lotofácil, Quina, Lotomania e outros).* Ao escolher um desses jogos para apostar, é razoável que queiramos analisar a chance de se ganhar o prêmio. Ao efetuar uma jogada, sabemos de antemão quais os números que podem ser sorteados, bem como quantos serão selecionados. Infelizmente, não sabemos quais os números que efetivamente serão sorteados.

Em ocasiões como essas e muitas outras que vivenciamos no dia a dia, nos deparamos com situações ou acontecimentos que, ao serem repetidos nas mesmas condições, produzem resultados diferentes que não podem ser determinados previamente, mas podem ser previstos. Esses experimentos são chamadas de **experimentos aleatórios**, pois sabemos os resultados possíveis mas não podemos ter certeza deles.

**Exemplo 2.1.1** (Lançamento de uma moeda)**.** Um dos experimentos aleatórios mais simples é aquele que se obtém ao lançar uma moeda. Nesse caso, não se pode dizer com certeza qual face da moeda terminará virada para cima, mas sabe-se que há apenas dois resultados possíveis: cara $(C)$ ou coroa $(K)$.

**Exemplo 2.1.2** (Lançamento duplo de uma moeda)**.** Eventualmente, pode-se lançar uma ou mais moedas e anotar o resultado. Ao lançar duas vezes uma moeda, por exemplo, pode-se observar que há quatro pares de resultados possíveis: $CK$, $CC$, $KK$, $KC$. Esses pares de resultados referem-se às faces obtidas nos dois lançamentos da moeda, sendo a primeira e a segunda letras de cada par correspondentes às faces obtidas no primeiro e segundo lançamentos, respectivamente, supondo que a ordem dos resultados é importante.

**Exemplo 2.1.3** (Lançamento de um dado cúbico comum)**.** Um outro experimento aleatório tão corriqueiro quanto o lançamento de uma moeda é o lançamento de um dado cúbico comum, composto de faces numeradas de 1 a 6. Nesse experimento, analisa-se o número obtido na face superior do dado, sendo 6 resultados possíveis: 1, 2, 3, 4, 5 e 6.

**Exemplo 2.1.4** (Escolher um número natural)**.** Se fosse possível criar um mecanismo capaz de escolher aleatoriamente um número natural, então esse mecanismo seria um experimento aleatório com infinitos resultados possíveis.

**Exemplo 2.1.5** (Tipo sanguíneo de um bebê cujos pais possuem tipos sanguíneos AB e O)**.** Um experimento aleatório que pode-se observar na Biologia é aquele onde trata-se da análise do tipo sanguíneo de um bebê cujos pais tem tipos sanguíneos AB e O. Fatores genéticos contribuem para que o tipo sanguíneo do bebê seja determinado ao acaso. Apesar disso, a Genética nos garante que há apenas dois tipos sanguíneos possíveis para o bebê: $A$ ou $B$.

Para calcular a chance de qualquer um dos resultados de um experimento aleatório acontecer, é preciso conhecer todos os resultados possíveis (ou pelo menos ter uma ideia de quais são todos eles).

**Definição 2.1** (Espaço amostral)**.** O conjunto de todos os resultados possíveis de um experimento aleatório é chamado de *espaço amostral*. Cada resultado do experimento em questão é, portanto, um elemento de seu espaço amostral. Usa-se a letra grega maiúscula ômega ($\Omega$) para representar o espaço amostral de um experimento aleatório.

**Exemplo 2.1.6.** O espaço amostral $\Omega$ do experimento 2.1.1 é formado pelos resultados $C$ e $K$.

Antes de prosseguir, convém citar alguns tópicos relativos ao estudo dos conjuntos.

## 2.2 Espaço Amostral e Conjuntos

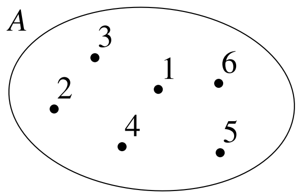
O propósito desse tópico é introduzir as notações e os conceitos de conjuntos associados ao conceito de probabilidade, pois eles estarão intimamente ligados ao longo deste trabalho.

As ideias de conjunto e elemento são tidas como conceitos primitivos da matemática, pois são compreendidas sem a necessidade de definições formais. Costumeiramente, diz-se que um **conjunto** representa uma coleção, reunião ou agrupamento de entes ou objetos, que são chamados de **elementos** do conjunto. Na notação usual, utilizam-se letras maiúsculas do alfabeto moderno para representar um conjunto e letras minúsculas do mesmo alfabeto para representar os seus elementos. Para denotar que um elemento $a$ está em um conjunto $A$, escrevemos $a \in A$ e, no caso contrário, escrevemos $a \notin A$. Ao escrever $a \in A$, também podemos entender que o elemento $a$ possui a propriedade que caracteriza os elementos de $A$, quando seus elementos possuem dada característica em comum. Há três maneiras usuais de se representar conjuntos:

- Listando seus elementos entre chaves:

$$A = \{1,2,3,4,5,6\}$$

- Através de um diagrama de Venn:




- Usando uma característica de seus elementos, quando tal característica é possível de ser estabelecida:

$$A = \{x \in \mathbb{N} \mid x \text{ é um número desenhado em uma face de um dado cúbico comum}\}$$

No exemplo acima, usamos a letra $x$ para designar um elemento do conjunto $A$. No caso do espaço amostral $\Omega$, cada um de seus elementos é chamado *ponto amostral* e será representado pela legra grega minúscula ômega ($\omega$).

De volta aos exemplos anteriores, temos, então, que:

- No exemplo 2.1.1, temos $\Omega_{2.1.1} = \{C,K\}$.
- No exemplo 2.1.2, temos $\Omega_{2.1.2} = \{CK,CC,KK,KC\}$.
- No exemplo 2.1.3, temos $\Omega_{2.1.3} = \{1,2,3,4,5,6\}$.
- No exemplo 2.1.4, temos $\Omega_{2.1.4} = \{1,2,3,4,5,...\}$.
- No exemplo 2.1.5, temos $\Omega_{2.1.5} = \{A,B\}$.

Para representar os elementos de um espaço amostral provenientes de alguns experimentos específicos, como, por exemplo, o lançamento de dois dados, é conveniente utilizar os conceitos de par ordenado e produto cartesiano. **Par ordenado** é uma dupla de elementos $a$ e $b$ na qual designamos $a$ como primeiro elemento e $b$ como segundo elemento, indicando-o por $(a,b)$. Dessa maneira, os pares $(a,b)$ e $(b,a)$ são considerados distintos (excetuando-se o caso em que $a = b$) e os pares $(a,b)$ e $(c,d)$ são iguais se, e somente se, $a = c$ e $b = d$. O conceito análogo ao de par ordenado usando três elementos é chamado de *terno ordenado*: $(a,b,c)$. Já ao utilizar $n$ elementos, compõe-se o que chamamos de *n-upla ordenada*: $(a,b,c,d,\ldots,n)$. O **produto cartesiano** de dois conjuntos $A$ e $B$ é o conjunto de todos os pares ordenados cujo primeiro elemento pertence a $A$ e o segundo elemento pertence a $B$, indicando da seguinte maneira:

$$A \times B = \{(a,b) \mid a \in A \text{ e } b \in B\}.$$

**Exemplo 2.2.1.** O produto cartesiano dos conjuntos $A = \{a,b,c\}$ e $B = \{c,d,e\}$ é o conjunto de pares ordenados

$$A \times B = \{(a,c),(a,d),(a,e),(b,c),(b,d),(b,e),(c,c),(c,d),(c,e)\}.$$

Note que, segundo a definição, tem-se $(d,a) \notin A \times B$.

**Exemplo 2.2.2.** Para o exemplo 2.1.2, podemos reescrever $\Omega_{2.1.2}$ como $\Omega_{2.1.2} = \{(C,K),(C,C),(K,K),(K,C)\}$. Dessa forma, o primeira e o segundo elemento de cada par ordenado correspondem, respectivamente, às faces obtidas no primeiro e segundo lançamentos. Veja que cada par ordenado é um ponto amostral de $\Omega_{2.1.2}$.

**Exemplo 2.2.3.** Considere o experimento "lançar duas vezes o mesmo dado". Usando o $\Omega_{2.1.3}$ para enumerar os elementos do espaço amostral associado, que chamaremos de $\Omega_{2.2.3}$, teremos:

$$\begin{aligned}
\Omega_{2.2.3} &= \Omega_{2.1.3} \times \Omega_{2.1.3} = \{1,2,3,4,5,6\} \times \{1,2,3,4,5,6\} = \\
&= \{(1,1),(1,2),(1,3),(1,4),(1,5),(1,6),(2,1),(2,2),(2,3),(2,4),(2,5),(2,6), \\
&\quad (3,1),(3,2),(3,3),(3,4),(3,5),(3,6),(4,1),(4,2),(4,3),(4,4),(4,5),(4,6), \\
&\quad\quad (5,1),(5,2),(5,3),(5,4),(5,5),(5,6),(6,1),(6,2),(6,3),(6,4),(6,5),(6,6)\}.
\end{aligned}$$

Ao tomar $(3,5) \in \Omega_{2.2.3}$ como exemplo, observa-se que o número 3 refere-se ao número obtido no primeiro lançamento do dado e que o número 5 refere-se ao número obtido no segundo lançamento do dado.

## 2.3 Eventos e Subconjuntos

Chamamos **subconjunto** de um conjunto $A$ um conjunto $A'$ para o qual todos os elementos que o formam também são elementos de $A$. Representamos esse fato usando a notação $A' \subset A$. Por exemplo, se $A = \{a,b,c,d,e\}$ e $A' = \{a,c,e\}$, temos $A' \subset A$, pois $\forall x \in A' \Rightarrow x \in A$. É importante observar que o conjunto vazio, que é aquele que não possui elementos e é denotado por $\emptyset$, é subconjunto de qualquer conjunto.

**Exemplo 2.3.1.** Tomemos o conjunto $V = \{a,e,i,o,u\}$ das vogais do nosso alfabeto atual. Temos, então, que

- $\{a\} \subset V$, pois $a \in V$.

- $\{a,o\} \subset V$, pois $a \in V$ e $o \in V$.
- $V = \{a,e,i,o,u\} \subset V$, pois $a \in V$, $e \in V$, $i \in V$, $o \in V$ e $u \in V$.
- $\emptyset \subset V$.

Chamamos de **eventos** os subconjuntos de um espaço amostral $\Omega$. Qualquer conjunto de elementos do espaço amostral, portanto, é um evento desse espaço amostral.

**Exemplo 2.3.2.** De volta ao exemplo 2.2.3, cujo espaço amostral já foi listado, podemos considerar o conjunto $A = \{(1,4),(2,3),(3,2),(4,1)\}$. Observe que a soma dos números em cada par ordenado de $A$ é igual a 5 e esses são os únicos pares ordenados de $\Omega_{2.2.3}$ com essa propriedade. Assim, como cada elemento de $A$ também é um elemento de $\Omega_{2.2.3}$, isto é, $A \subset \Omega_{2.2.3}$, podemos definir $A$ como um evento de $\Omega_{2.2.3}$, que relaciona os pares ordenados de resultados responsáveis por "*obter a soma* 5". Logo, o evento $A$: "*obter dois números cuja a soma é* 5", ou simplesmente "obter a soma 5", é um evento de $\Omega_{2.2.3}$.

**Exemplo 2.3.3.** Consideremos o experimento aleatório "lançar uma moeda e, em seguida, lançar um dado". Nesse caso, o espaço amostral será dado por:

$$\begin{aligned}\Omega_{2.3.3} &= \{(C,1),(C,2),(C,3),(C,4),(C,5),(C,6),\\ &\quad\ (K,1),(K,2),(K,3),(K,4),(K,5),(K,6)\}\end{aligned}$$

Considerando o evento $A =$ "*obter* cara *e* um número par", temos

$$A = \{(C,2),(C,4),(C,6)\} \subset \Omega_{2.3.3}.$$

Dizemos que um evento $A$ implica o evento $B$ se para todo $\omega \in A$ tivermos $\omega \in B$, ou seja $A \subset B$. Nesse caso, entenderemos que a ocorrência de $A$ garante inevitavelmente a ocorrência de $B$. Além disso, dois eventos $A$ e $B$ serão ditos **iguais** se, e somente se, $A \subset B$ e $B \subset A$, isto é, todo elemento de $A$ também é um elemento de $B$ e vice-versa.

Dizemos que um evento $A$ ocorre quando algum resultado do experimento aleatório pertencer a $A$. Da mesma maneira, dizemos que um evento $A$ não ocorre quando certo resultado do experimento aleatório não pertencer a ele. No caso do exemplo 2.3.2, se no primeiro lançamento do dado for obtido o número 2 e no segundo lançamento for obtido o número 3, então diremos que o evento $A$ em questão ocorreu, pois $(2,3) \in A$. Entretanto, se os resultados obtidos nos dois lançamentos formarem o par $(1,6)$, diremos que $A$ não ocorreu, pois $(1,6) \notin A$.

Chamaremos de **evento elementar** aquele que contém um único elemento (conjunto unitário). No exemplo 2.1.2, o evento $A =$ "obter *cara* nos dois lançamentos" é um evento elementar, pois $A = \{(C,C)\} \subset \Omega_{2.1.2}$.

## 2.4 Eventos e Operações entre Conjuntos

Nesta seção, trataremos das operações entre eventos de um espaço amostral, que são as mesmas entre conjuntos. Para isso, considere dois eventos $A$ e $B$ de um espaço amostral $\Omega$. As operações entre eventos serão denotadas por:

1. **Intersecção**: a intersecção dos eventos $A$ e $B$, denotada $A \cap B$, é o evento no qual os eventos $A$ *e* $B$ ocorrem *simultaneamente*:

$$A \cap B = \{\omega \in \Omega \mid \omega \in A \ \ \mathbf{e} \ \omega \in B\}.$$

2. **União**: a união dos eventos $A$ e $B$, denotada $A \cup B$, é o evento no qual ocorre *pelo menos* um dos eventos $A$ ou $B$, isto é, quando *só* $A$ ou *só* $B$ ou *ambos* $A$ *e* $B$ ocorrem:

$$A \cup B = \{\omega \in \Omega \mid \omega \in A \ \ \mathbf{ou} \ \omega \in B \ \mathbf{ou} \ \omega \in (A \cap B)\}.$$

3. **Complementar**: o evento complementar de $A$, denotado por $A^c$, é o evento no qual $A$ não ocorre:

$$A^c = \{\omega \in \Omega \mid \omega \notin A\}.$$

4. **Diferença**: a diferença entre os eventos $A$ e $B$, denotada nessa ordem por $A - B$, é o evento no qual o evento $A$ ocorre mas o evento $B$ não ocorre, isto é, os elementos da diferença $A - B$ pertencem ao evento $A$ mas não pertencem ao evento $B$:

$$A - B = \{\omega \in \Omega \mid \omega \in A \ \mathbf{e} \ \omega \notin B\} = A \cap B^c.$$

   Note que, em geral, $(A - B) \neq (B - A)$.

5. **Eventos mutualmente exclusivos**: dois eventos $A$ e $B$ são chamados de mutualmente exclusivos (ou **disjuntos**) se ambos não ocorrem simultaneamente, ou seja, o evento $A \cap B = \emptyset$.

**Exemplo 2.4.1.** Considere uma urna que contém bolas idênticas numeradas de 1 a 10. Uma bola é sorteada da urna e seu número é anotado. Tomemos os eventos $A$: "*o número da bola retirada é par*", $B$: "*o número da bola retirada é ímpar*" e $C$: "*o número da bola retirada é um divisor de 12*". Dessa forma, temos:

- $\Omega = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$;
- $A = \{2, 4, 6, 8, 10\}$;
- $B = \{1, 3, 5, 7, 9\}$;
- $C = \{1, 2, 3, 4, 6\}$.

Temos, por exemplo, que:

- $A \cap C = \{2,4,6\}$: números das bolas da urna que são pares **e** divisores de 12.
- $A \cup C = \{1,2,3,4,6,8,10\}$: números das bolas da urna que são pares **ou** divisores de 12.
- $A^c = \{1,3,5,7,9\} = B$: números das bolas da urna que não são pares.
- $C - A = \{1,3\} = C \cap A^c$: números que são divisores de 12 mas não são pares.
- $A \cap B = \emptyset$, pois não há números que sejam simultaneamente pares e ímpares. Por isso, $A$ e $B$ são eventos ditos mutualmente exclusivos.

## *Notações importantes*

Seja $A_1, A_2, \ldots, A_n$ uma sequência **finita** de eventos.

- $\bigcap_{i=1}^{n} A_i$ representa a intersecção desses eventos e corresponde ao evento em que todos os eventos $A_i$ ocorrem simultaneamente:

$$\bigcap_{i=1}^{n} A_i = A_1 \cap A_2 \cap \cdots \cap A_n.$$

- $\bigcup_{i=1}^{n} A_i$ representa a união desses eventos e corresponde ao evento em que pelo menos um dos eventos $A_i$ ocorre:

$$\bigcup_{i=1}^{n} A_i = A_1 \cup A_2 \cup \cdots \cup A_n.$$

Seja $A_1, A_2, \ldots$ uma sequência **infinita** de eventos.

- $\bigcap_{i=1}^{\infty} A_i$ representa a intersecção desses eventos e corresponde ao evento em que todos os eventos $A_i$ ocorrem simultaneamente:

$$\bigcap_{i=1}^{\infty} A_i = A_1 \cap A_2 \cap \ldots$$

- $\bigcup_{i=1}^{\infty} A_i$ representa a união desses eventos e corresponde ao evento em que pelo menos um dos eventos $A_i$ ocorre:

$$\bigcup_{i=1}^{\infty} A_i = A_1 \cup A_2 \cup \ldots$$

### Propriedades importantes

Considere $A$, $B$ e $C$ eventos de um espaço amostral $\Omega$. Pode-se provar que:

1. $A \subset A$.
2. $A \subset B$ e $B \subset C \Rightarrow A \subset C$.
3. $\emptyset \subset A$.
4. $\Omega^c = \emptyset$.
5. $(\Omega \cap A) = A$.
6. $(\Omega \cup A) = \Omega$.
7. $(A \cup A^c) = \Omega$.
8. $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$.
9. $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$.
10. Leis de De Morgan.
    a) $(A \cup B)^c = A^c \cap B^c$.
    b) $(A \cap B)^c = A^c \cup B^c$.
    As leis de De Morgan também se aplicam a uma sequência finita de eventos $A_1, A_2, \ldots, A_n$ e, mais geralmente, a uma sequência infinita enumerável de eventos $A_1, A_2, \ldots$ de $\Omega$. Dessa maneira, temos
    c) $\left(\bigcup_{i=1}^{\infty} A_i\right)^c = \bigcap_{i=1}^{\infty} A_i^c$.
    d) $\left(\bigcap_{i=1}^{\infty} A_i\right)^c = \bigcup_{i=1}^{\infty} A_i^c$.

## 2.5 Definindo Probabilidade

A probabilidade de um evento acontecer tem, hoje, uma definição precisa: a chamada definição *axiomática*. Apesar disso, há outras definições que foram usadas em casos particulares ao longo da história e que ainda hoje são utilizadas em alguns contextos. Essas outras definições encontram-se, atualmente, englobadas e justificadas pela definição axiomática da probabilidade. Esse é o caso da definição *clássica*, da qual trataremos a seguir.

É a definição clássica de Probabilidade a mais comumente apresentada nos livros didáticos de matemática do Ensino Básico. Basta uma rápida análise em dois ou três livros desses para se observar que a probabilidade é apresentada de forma extremamente simplista, imprecisa, geralmente através de uma fórmula e explorando poucas propriedades.

### 2.5.1 Definição Clássica de Probabilidade

Talvez a pergunta mais simples que possamos fazer ao realizarmos um experimento aleatório seja "*qual a probabilidade de certo evento acontecer?*". Ao lançar uma moeda honesta, por exemplo, se fizermos a pergunta "*qual a chance da face que ficar virada para cima ser uma* cara*?*", muitas pessoas podem responder intuitivamente: "50%". Esse tipo de resposta está associado à ideia primitiva de eventos igualmente prováveis, pois se a moeda for honesta, espera-se que a chance de se obter *cara* seja a mesma de se obter *coroa*. Dessa forma, quer-se obter apenas um resultado dentre dois possíveis, podemos pensar que a probabilidade em questão seja "1 em 2". Também podemos pensar que tomamos todos os resultados que nos interessam para que o evento em questão ocorra $(1)$, e dividimos pelo número total de resultados possíveis do experimento $(2)$.

Há duas hipóteses que estamos usando intuitivamente para formar esse raciocínio:

1. o espaço amostral $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$ é finito, portanto enumerável.
2. a probabilidade de cada evento acontecer é a mesma, ou seja, a probabilidade de cada um dos eventos elementares $\{\omega_i\}$, $i = 1, 2, \ldots, n$, é dada por $\frac{1}{n}$.

Ao analisar o lançamento de um dado, um evento que podemos analisar é "obter um número múltiplo de três". Nesse caso, podemos calcular a probabilidade desse evento levando em conta que há dois casos favoráveis dos seis resultados possíveis, nos levando a concluir que a probabilidade em questão seja de $\frac{2}{6} = \frac{1}{3}$. Veja que este evento equivale ao evento "obter os números 3 ou 6", que corresponde à união dos eventos elementares "obter o número 3" e "obter o número 6", que são disjuntos. Assim, ao pensar da maneira que pensamos, estamos intuitivamente supondo que a probabilidade da união de dois eventos disjuntos é igual à soma de suas probabilidades individuais. Dessa forma, podemos levantar uma terceira hipótese que também estamos utilizando intuitivamente:

3. a probabilidade da união de dois eventos disjuntos equivale à soma das probabilidades individuais, ou seja, $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$.

Essas hipóteses foram utilizadas por matemáticos a partir do século XVII para modelar situações geradas em jogos de azar, como o lançamento de moedas e dados. Esses matemáticos definiram a probabilidade de um evento ocorrer como o número de resultados favoráveis a esse evento dividido pelo número total de resultados possíveis do experimento aleatório em questão, isto é, o número de elementos de $\Omega$. Essa ideia, chamada de definição **clássica** da probabilidade, pode ser mais formalmente escrita da seguinte maneira:

**Definição 2.2** (Definição clássica de probabilidade)**.** Consideremos um espaço amostral $\Omega$ finito com $q$ elementos. Seja $A$ um evento de $\Omega$ composto de $p$ elementos. A probabilidade de $A$, que denotaremos $\mathbb{P}(A)$, é definida por

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

onde $p = n(A)$ e $q = n(\Omega)$ representam os números de elementos de $A$ e de $\Omega$, respectivamente.

**Exemplo 2.5.1.** De volta ao exemplo 2.4.1, no qual escolhemos uma entre 10 bolas numeradas de uma urna, onde tínhamos $\Omega = \{1,2,3,4,5,6,7,8,9,10\}$. Seja $C$ o evento "obter um divisor de 12", tal que $C = \{1,2,3,4,6\}$, como já havíamos definido. Supondo que a chance de cada bola ser selecionada na urna seja a mesma, queremos calcular a probabilidade do evento $C$ ocorrer. Como $n(C) = 5$ e $n(\Omega) = 10$, segue da definição clássica que a probabilidade de $C$ ocorrer é

$$\mathbb{P}(C) = \frac{n(C)}{n(\Omega)} = \frac{5}{10} = \frac{1}{2} = 0,5.$$

Supondo que $\mathbb{P}(\{1\}) = \mathbb{P}(\{2\}) = \cdots = \mathbb{P}(\{10\}) = \frac{1}{10}$, ou seja, que os eventos elementares são equiprováveis, então $\mathbb{P}(C)$ também pode ser calculada da seguinte maneira:

$$\begin{aligned}
\mathbb{P}(C) = \mathbb{P}(\{1,2,3,4,6\}) &= \mathbb{P}(\{1\} \cup \{2\} \cup \{3\} \cup \{4\} \cup \{6\}) \\
&= \mathbb{P}(\{1\}) + \mathbb{P}(\{2\}) + \mathbb{P}(\{3\}) + \mathbb{P}(\{4\}) + \mathbb{P}(\{6\}) \\
&= \frac{1}{10} + \frac{1}{10} + \frac{1}{10} + \frac{1}{10} + \frac{1}{10} \\
\mathbb{P}(C) &= \frac{5}{10} = \frac{1}{2} = 0,5.
\end{aligned}$$

A partir da definição 2.2 e das hipóteses mencionadas anteriormente, podemos estabelecer as seguintes propriedades da definição clássica de probabilidade:

1. $\mathbb{P}(A) \geqslant 0,\ \forall A \subset \Omega$;
2. $\mathbb{P}(\Omega) = \dfrac{n(\Omega)}{n(\Omega)} = 1$; e
3. Se $A$ e $B$ são eventos mutuamente exclusivos, então $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$.

**Exemplo 2.5.2.** Tomemos $A = \{\omega_1, \omega_2\}$ e $B = \{\omega_3, \omega_5, \omega_7, \omega_9\}$ eventos em um espaço amostral $\Omega = \{\omega_1, \omega_2, \ldots, \omega_{12}\}$. Segundo a definição clássica de probabilidade, como $n(A) = 2$, $n(B) = 4$ e $n(\Omega) = 12$, temos $\mathbb{P}(A) = \frac{2}{12}$ e $\mathbb{P}(B) = \frac{4}{12}$. Além disso, sabemos que $A \cup B = \{\omega_1, \omega_2\, \omega_3, \omega_5, \omega_7, \omega_9\}$, ou seja, $n(A \cup B) = 6$. Logo, aplicando a definição 2.2, temos

$$\mathbb{P}(A \cup B) = \frac{n(A \cup B)}{n(\Omega)} = \frac{6}{12} = \frac{(2+4)}{12} = \frac{2}{12} + \frac{4}{12} = \mathbb{P}(A) + \mathbb{P}(B).$$

Nesta definição de probabilidade, para calcular a probabilidade de um evento qualquer, é preciso contar o número de eventos elementares do evento em questão, bem como os eventos elementares do espaço amostral. Para tornar mais simples esse processo, pode-se utilizar métodos de contagem conhecidos e princípios de Análise Combinatória, dentre os quais as árvores de possibilidades e tabelas para organizar os dados envolvidos.

A definição clássica de probabilidade é a definição mais encontrada e utilizada pelos livros didáticos de Matemática do Ensino Básico. Talvez por isso também seja a mais utilizada pelos professores do Ensino Básico em sala de aula. Uma explicação para isso é que os casos estudados nessa etapa da escolaridade apresentam, sempre, um espaço amostral $\Omega$ finito.

Entretanto, nem todos os problemas de probabilidade podem ser resolvidos usando a definição clássica de probabilidade. Há dois casos tradicionais em que essa definição não é suficiente (basta negar uma das duas hipóteses "intuitivas" iniciais):

- **$\Omega$ não possui número finito de elementos**.

  **Exemplo 2.5.3.** Sortear uma bola de uma urna com duas bolas (uma preta $P$ e uma branca $B$) repetidamente, com reposição da bola retirada, até se obter uma bola preta.

  $$\Omega = \{P, BP, BBP, BBBP, \ldots, BBBBBBBBBBP, \ldots, BBB \cdots BP, \ldots\}$$

  Nesse caso, o espaço amostral é um conjunto com número infinito enumerável de elementos. Por isso, nessa situação, a definição clássica não é aplicável.

- **$\Omega$ é finito, mas os eventos elementares não são equiprováveis**.

  **Exemplo 2.5.4.** Considere o experimento em que jogamos um dado *viciado*, isto é, um dado onde alguns números tem mais chance de serem obtidos do que outros. Neste caso, vamos considerar que os números 2 e 4, por exemplo, tenham mais chance de serem obtidos do que os demais:

  $$\mathbb{P}(\{2\}) = \frac{1}{5}, \quad \mathbb{P}(\{4\}) = \frac{1}{3}, \quad \mathbb{P}(\{1\}) = \mathbb{P}(\{3\}) = \mathbb{P}(\{5\}) = \mathbb{P}(\{6\}) = \frac{7}{60}.$$

  Nesse caso, a definição clássica não serve para calcular a probabilidade do evento $D$: "obter um número par". De acordo com a definição clássica, havendo três números pares dentre os seis números possíveis, a probabilidade do evento citado seria 3 em 6, ou seja, $\mathbb{P}(D) = \frac{3}{6} = 50\%$. Entretanto, usando a terceira propriedade da definição clássica de probabilidade, definida anteriormente, poderíamos calcular a mesma probabilidade como $\mathbb{P}(D) = \mathbb{P}(\{2\} \cup \{4\} \cup \{6\}) = \mathbb{P}(\{2\}) + \mathbb{P}(\{4\}) + \mathbb{P}(\{6\}) = \frac{1}{5} + \frac{1}{3} + \frac{7}{60} = \frac{39}{60} \neq \frac{3}{6}$, gerando, assim, uma contradição.

Como essa definição parte do princípio de que os eventos elementares são equiprováveis, pode-se perceber que a definição clássica faz uso de si mesma para se definir, num ciclo vicioso. Do ponto de vista formal da matemática, portanto, a definição clássica não seria uma definição propriamente dita. Usaremos essa palavra apenas por questão de organização dessa dissertação.

### 2.5.2 Definição Axiomática de Probabilidade

Para poder aprofundar o conceito de probabilidade e trabalhar resultados mais abrangentes, faz-se necessária a definição axiomática da probabilidade que, de certa forma, justifica a definição clássica (2.2). Para podermos definí-la, é preciso, antes, falarmos de um outro conceito.

**Definição 2.3** ($\sigma$-álgebra)**.** Uma coleção não vazia de subconjuntos de um espaço amostral $\Omega$ que é fechada sob um número infinito enumerável de operações da teoria dos conjuntos é chamada $\sigma$-álgebra de subconjuntos de $\Omega$. De outra forma, podemos dizer que uma coleção $\mathscr{A}$ de subconjuntos de $\Omega$ é uma $\sigma$-álgebra de conjuntos de $\Omega$ se as seguintes propriedades forem satisfeitas:

(i) $\Omega \in \mathscr{A}$;

(ii) Se $A \in \mathscr{A}$, então $A^c \in \mathscr{A}$;

(iii) Se $\{A_n\}_{n \geqslant 1} \in \mathscr{A}$, $n \in \mathbb{N}$, então $\bigcup\limits_{n=1}^{\infty} A_n \in \mathscr{A}$ e $\bigcap\limits_{n=1}^{\infty} A_n \in \mathscr{A}$.

**Exemplo 2.5.5** (A $\sigma$-álgebra mais simples que contém um subconjunto)**.** Em muitas situações, é interessante construir uma $\sigma$-álgebra que tenha entre seus elementos um subconjunto particular $A$ de $\Omega$.

Tomemos $\mathscr{F} = \{\emptyset, A, A^c, \Omega\}$. $\mathscr{F}$ é defininda como a menor $\sigma$-álgebra que contém o conjunto $A$. Qualquer outra $\sigma$-álgebra terá os elementos de $\mathscr{F}$ e, eventualmente, mais alguns (sendo, portanto, maior que $\mathscr{F}$).

**Exemplo 2.5.6** (Como verificar se uma coleção de subconjuntos é uma $\sigma$-álgebra)**.** Tomemos $\Omega = \{a,b,c,d\}$ e as seguintes coleções de subconjuntos:

$$\mathscr{D} = \{\emptyset, \Omega, \{a,b,c\}, \{d\}\} \text{ e } \mathscr{E} = \{\emptyset, \Omega, \{a\}, \{b,c\}, \{a,d\}, \{b,c,d\}\}.$$

Como reconhecer se $\mathscr{D}$ e $\mathscr{E}$ são $\sigma$-álgebras de $\Omega$? Basta verificar os itens (i), (ii) e (iii) da definição 2.3. Como $\Omega$ é finito, basta verificar que qualquer união finita pertence à coleção de subconjuntos para que o item (iii) da Definição 2.3 seja satisfeito.

- Para $\mathscr{D}$, temos:

  (i) $\Omega \in \mathscr{D}$.

  (ii) $\forall A \in \mathscr{D}$, temos $A^c \in \mathscr{D}$, pois

  $$\Omega^c = \emptyset \in \mathscr{D},\ \ \emptyset^c = \Omega \in \mathscr{D},\ \ \{a,b,c\}^c = \{d\} \in \mathscr{D},\ \ e\ \ \{d\}^c = \{a,b,c\} \in \mathscr{D}.$$

  (iii) $\{a,b,c\} \cup \{d\} = \{a,b,c,d\} = \Omega \in \mathscr{D}$ (os demais casos são triviais).

  Como os três itens acima são satisfeitos, podemos concluir que $\mathscr{D}$ é uma $\sigma$-álgebra de $\Omega$.

- Para $\mathscr{E}$, temos:

  (i) $\Omega \in \mathscr{E}$.

  (ii) $\forall A \in \mathscr{E}$, temos $A^c \in \mathscr{E}$, pois

  $$\Omega^c = \emptyset \in \mathscr{E},\ \ \emptyset^c = \Omega \in \mathscr{E},\ \ \{a\}^c = \{b,c,d\} \in \mathscr{E},\ \ \{b,c\}^c = \{a,d\} \in \mathscr{E}$$

  $$\{a,d\}^c = \{b,c\} \in \mathscr{E},\ \ e\ \ \{b,c,d\}^c = \{a\} \in \mathscr{E}.$$

  (iii) Nem toda união de eventos de $\mathscr{E}$ pertence a $\mathscr{E}$, como $\{a\} \cup \{b,c\} = \{a,b,c\} \notin \mathscr{E}$, por exemplo.

  Dessa forma, como nem todos os três itens são satisfeitos, podemos concluir que $\mathscr{E}$ **não** é uma $\sigma$-álgebra de $\Omega$.

**Exemplo 2.5.7.** O conjunto das partes $\mathscr{P}$ de qualquer conjunto finito sempre é uma $\sigma$-álgebra, pois contém todos os subconjuntos desse conjunto. Em particular, para um experimento aleatório com $\Omega$ finito, temos que $\mathscr{P}(\Omega)$ é uma $\sigma$-álgebra de $\Omega$. De fato, veja que

(i) $\Omega \in \mathscr{P}(\Omega)$.

(ii) $\forall A \in \mathscr{P}(\Omega)$, temos $A^c \in \mathscr{P}(\Omega)$, pois $A^c \subset \Omega$.

(iii) A união de quaisquer eventos de $\Omega$ está em $\mathscr{P}(\Omega)$.

Como os três itens acima são satisfeitos, podemos concluir que $\mathscr{P}(\Omega)$ **sempre é** uma $\sigma$-álgebra de $\Omega$.

Para os exemplos que serão abordados neste material, nos quais $\Omega$ é finito, a $\sigma$-Álgebra envolvida é o conjunto das partes $\mathscr{P}(\Omega)$ de todos os eventos (subconjuntos) possíveis de $\Omega$.

Feitas as definições acima, podemos finalmente apresentar a definição a seguir.

**Definição 2.4** (Definição axiomática de probabilidade)**.** A Probabilidade é uma função, denotada por $\mathbb{P}$, definida em uma $\sigma$-álgebra de $\Omega$ assumindo valores no intervalo real $[0,1]$ e satisfazendo os seguintes axiomas:

$(A_1)$ $\mathbb{P}(A) \geqslant 0$, $\forall A \subset \Omega$;

$(A_2)$ $\mathbb{P}(\Omega) = 1$; e

$(A_3)$ Para uma sequência infinita $\{A_n\}_{n \geqslant 1}$ de eventos mutuamente exclusivos (disjuntos) de $\Omega$, temos que

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

Algumas propriedades que seguem dessa definição:

1. $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$.
2. $\mathbb{P}(\emptyset) = 0$.
3. Se $A \subset B$, então $\mathbb{P}(A) \leqslant \mathbb{P}(B)$.
4. $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$, para quaisquer eventos $A$ e $B$ em $\Omega$.

**Exemplo 2.5.8.** Consideremos novamente o exemplo 2.4.1, relativo ao sorteio de uma bola numerada de uma urna, onde $\Omega = \{1,2,3,4,5,6,7,8,9,10\}$. Vamos calcular a probabilidade do evento $D$: "o número da bola sorteada é múltiplo de 3", ou seja, o evento $D = \{3,6,9\}$. Pela definição clássica de probabilidade, temos, então,

$$\mathbb{P}(D) = \frac{n(D)}{n(\Omega)} = \frac{3}{10} = 0,3.$$

Essa afirmação está correta, pois ao utilizar a definição clássica de probabilidades, estamos supondo que todas as bolas tem a mesma chance de serem sorteadas, isto é, que os eventos elementares são equiprováveis. Pela definição axiomática, no caso de eventos elementares equiprováveis, temos que

$$\mathbb{P}(D) = \mathbb{P}(\{3,6,9\}) = \mathbb{P}(\{3\} \cup \{6\} \cup \{9\}).$$

Mas como os eventos elementares $\{3\}$, $\{6\}$ e $\{9\}$ são mutuamente exclusivos (não ocorrem simultaneamente), pelo Axioma $A_3$, temos que

$$\mathbb{P}(D) = \mathbb{P}(\{3\} \cup \{6\} \cup \{9\}) = \mathbb{P}(\{3\}) + \mathbb{P}(\{6\}) + \mathbb{P}(\{9\}) = \frac{1}{10} + \frac{1}{10} + \frac{1}{10} = \frac{3}{10} = 0,3.$$

Assim, pela definição axiomática, concluímos que a probabilidade de escolher uma bola cujo número seja múltiplo de 3 é de $0,3$.

## 2.6 Espaço de Probabilidade

A partir das definições anteriores, podemos falar no que segue.

**Definição 2.5** (Espaço de Probabilidade)**.** Um espaço de probabilidade é representado pela tripla $(\Omega, \mathscr{A}, \mathbb{P})$ no qual $\Omega$ representa o espaço amostral, $\mathscr{A}$ representa uma $\sigma$-álgebra de $\Omega$ e $\mathbb{P}(.)$ é a função de probabilidade, cujo domínio é $\mathscr{A}$ e o contradomínio é o intervalo $[0,1]$.

**Exemplo 2.6.1.** Tomemos $\Omega = \{a,b\}$, $\mathscr{A} = \{\emptyset, \{a\}, \{b\}, \{a,b\}\}$ (conjunto das partes de $\Omega$) e $p$ tal que $0 \leqslant p \leqslant 1$. Definindo $\mathbb{P} : \mathscr{A} \to [0,1]$ tal que $\mathbb{P}(\{a\}) = p$ e $\mathbb{P}(\{b\}) = 1 - p$, temos que $(\Omega, \mathscr{A}, \mathbb{P})$ é um espaço de probabilidade.

## 2.7 Probabilidade condicional

É do conhecimento popular a composição das 52 cartas de um baralho: 4 "naipes" diferentes (ouros, copas, espadas e paus, representados pelos símbolos $\diamondsuit, \heartsuit, \spadesuit$ e $\clubsuit$, respectivamente). Cada naipe é composto formado por 13 cartas enumeradas: $A, 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K$.

**Exemplo 2.7.1.** Suponha que nós queremos calcular a probabilidade do evento $D$: "sortear ao acaso a carta $Q\heartsuit$ (dama de copas)" de um baralho como este. Pela definição clássica, como $D = \{Q\heartsuit\}$, $n(D) = 1$ e $n(\Omega) = 52$, temos que $\mathbb{P}(D) = \frac{1}{52}$. Suponha que, por um descuido nosso, não tenhamos ficados atentos ao "sorteio" da carta e, por isso, o experimento tenha sido realizado sem que saibamos seu resultado. Ao pensar novamente na probabilidade do evento $D$, entretanto, alguém no recinto diz: "a carta sorteada é de copas". Ao sabermos disso, nossa percepção sobre os resultados mudam, pois sabemos que a carta sorteada não pode ser dos outros naipes. Sendo assim, das 52 opções que tínhamos para analisar no $\Omega$ inicial, agora restam-nos apenas as 13 cartas do naipe que sabemos que foi escolhido, dentre as quais apenas uma corresponde ao evento $D$. Dessa forma, nosso $\Omega$ inicial é reduzido ao $\Omega_\heartsuit = \{A\heartsuit, 2\heartsuit, 3\heartsuit, 4\heartsuit, 5\heartsuit, 6\heartsuit, 7\heartsuit, 8\heartsuit, 9\heartsuit, 10\heartsuit, J\heartsuit, Q\heartsuit, K\heartsuit\}$, onde pode-se listar um novo evento $D' = \{Q\heartsuit\} \subset \Omega_\heartsuit$, equivalente ao evento $D$ original. Usando, nesse ponto, a definição clássica de probabilidade, temos então que $\mathbb{P}(D') = \frac{n(D')}{n(\Omega_\heartsuit)} = \frac{1}{13} = 4 \cdot \mathbb{P}(D)$.

Veja que, na situação citada acima, o cálculo da probabilidade do evento em questão *depende* das informações que se tem sobre a realização do experimento. Ao adicionar uma nova informação à situação, a probabilidade de ocorrer o evento pode ser alterada, pois fica *condicionada* à nova dinâmica do experimento (altera-se o $\Omega$). Esse caso nos conduz à definição de Probabilidade Condicional.

**Definição 2.6** (Probabilidade condicional)**.** Seja $(\Omega, \mathscr{A}, \mathbb{P})$ um espaço de probabilidade, com $A, B \in \mathscr{A}$, com $\mathbb{P}(B) > 0$. A probabilidade condicional de qualquer evento $A$ dada a ocorrência de um evento $B$ é definida por

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}. \tag{2.1}$$

Neste caso, escreveremos $\mathbb{P}(A|B)$ para nos referir à "probabilidade da ocorrência do evento $A$ dada a ocorrência do evento $B$" ou, simplesmente, "a probabilidade de $A$ dado $B$".

É possível provar que a probabilidade condicional satisfaz os três axiomas da definição 2.4, a saber

$(A_1)$ $\mathbb{P}(A|B) \geqslant 0$;

$(A_2)$ $\mathbb{P}(\Omega|B) = 1$;

$(A_3)$ Para uma sequência infinita $\{A_n\}_{n\geqslant 1}$ de eventos mutuamente exclusivos (disjuntos) de $\Omega$, temos que

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

**Exemplo 2.7.2.** Considere uma loja de brinquedos onde há bolas e cubos, dentre eles azuis e vermelhos, distribuídos da seguinte maneira: há 5 cubos azuis, 4 cubos vermelhos, 7 bolas vermelhas e 6 bolas azuis, como representado na Tabela 1.

Tabela 1 – Distribuição dos brinquedos da loja, referente ao exemplo 2.7.2

| | azuis | vermelhos |
|---|---|---|
| cubos | 5 | 4 |
| bolas | 6 | 7 |

Ao escolher um desses brinquedos aleatoriamente, queremos saber qual a probabilidade do brinquedo escolhido ser azul. Após efetuada a escolha, entretanto, sabe-se que um cubo foi selecionado. Com essa informação em mãos, vamos calcular qual a chance do brinquedo escolhido ser azul, supondo que cada brinquedo tem a mesma chance de ser selecionado.

Sejam $A$ o evento "escolher um brinquedo azul" e $B$ o evento "escolher um cubo", de forma que $A\cap B$ corresponde ao evento "escolher um cubo azul". Sabe-se que $n(A) = 11$, $n(B) = 9$, $n(A\cap B) = 5$ e $n(\Omega) = 5+4+6+7 = 22$. A probabilidade que queremos calcular é $\mathbb{P}(A|B)$. Para usar a fórmula 2.1 da definição 2.6, partindo da definição clássica (2.2), precisamos calcular as seguintes probabilidades:

- $\mathbb{P}(A\cap B) = \dfrac{n(A\cap B)}{n(\Omega)} = \dfrac{5}{22}$.
- $\mathbb{P}(B) = \dfrac{n(B)}{n(\Omega)} = \dfrac{9}{22}$.

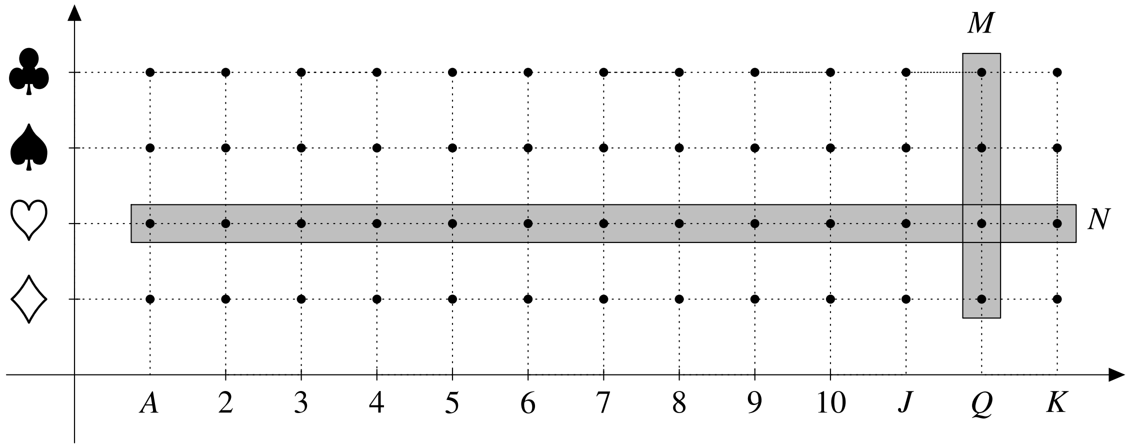
Dessa forma, tem-se

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A\cap B)}{\mathbb{P}(B)} = \frac{\frac{5}{22}}{\frac{9}{22}} = \frac{5}{22}\cdot\frac{22}{9} = \frac{5}{9}.$$

Veja que, ao afirmar que um cubo fora escolhido, não é mais necessário analisar os 22 brinquedos que poderiam ter sido escolhidos. Agora, a análise pode se restringir aos 9 cubos, dos quais 5 são azuis. Supondo que cada cubo, então, tenha a mesma chance $\frac{1}{9}$ de ser escolhido, então de fato a probabilidade de ser escolhido um cubo azul dentre os cubos disponíveis na loja é de $\frac{5}{9}$.

**Exemplo 2.7.3.** De volta ao baralho de 52 cartas (exemplo 2.7.1), considere os eventos $M$: "sortear uma dama ($Q$)" e $N$: "sortear uma carta de copas". Veja que $M$ e $N$ não são mutuamente exclusivos, pois $M\cap N \neq \emptyset$, como pode-se ver na figura 1. Para calcular $\mathbb{P}(M|N)$, afirma-se que o evento $N$ ocorreu. Fazer uma afirmação como essa equivale a dizer que não é necessário

Figura 1 – Distribuição dos eventos elementares no espaço amostral do exemplo 2.7.3.




levar em conta qualquer outro ponto do espaço amostral $\Omega$ que não pertença a $N$, ou seja, podemos considerar o evento $N$ como o novo espaço amostral para o experimento. Dessa maneira, $\mathbb{P}(M|N) = \frac{1}{13}$, pois apenas um dos treze eventos elementares de $N$ está em $M$ (usando a definição clássica 2.2 e supondo que os eventos elementares são equiprováveis). Usando a fórmula 2.1 da definição 2.6, temos, portanto:

$$\mathbb{P}(M|N) = \frac{\mathbb{P}(M \cap N)}{\mathbb{P}(N)} = \frac{\frac{1}{52}}{\frac{13}{52}} = \frac{1}{52} \cdot \frac{52}{13} = \frac{1}{13},$$

que corrobora o resultado anterior.

Observe que a fórmula 2.1 pode ser reescrita das seguintes maneiras, supondo $\mathbb{P}(A) > 0$ e $\mathbb{P}(B) > 0$:

$$\mathbb{P}(A \cap B) = \mathbb{P}(A|B) \cdot \mathbb{P}(B) = \mathbb{P}(B|A) \cdot \mathbb{P}(A) = \mathbb{P}(B \cap A) \tag{2.2}$$

**Exemplo 2.7.4.** De fato, nas mesmas condições do exemplo 2.7.3, temos que

$$\mathbb{P}(N|M) = \frac{\mathbb{P}(N \cap M)}{\mathbb{P}(M)} = \frac{\frac{1}{52}}{\frac{4}{52}} = \frac{1}{52} \cdot \frac{52}{4} = \frac{1}{4}.$$

Daí, note que

$$\mathbb{P}(M \cap N) = \frac{1}{52} = \frac{13}{52} \cdot \frac{1}{13} = \mathbb{P}(N) \cdot \mathbb{P}(M|N)$$

e

$$\mathbb{P}(M \cap N) = \frac{1}{52} = \frac{4}{52} \cdot \frac{1}{4} = \mathbb{P}(M) \cdot \mathbb{P}(N|M),$$

logo, como espera-se,

$$\mathbb{P}(M \cap N) = \mathbb{P}(N \cap M) = \mathbb{P}(N) \cdot \mathbb{P}(M|N) = \mathbb{P}(M) \cdot \mathbb{P}(N|M).$$

### 2.7.1 Teorema do Produto

A expressão 2.2 e sua generalização para uma intersecção de $n$ eventos permitem construir probabilidades em espaços amostrais que representam experimentos realizados em sequência, em que a ocorrência de um evento na $k$-ésima etapa depende das ocorrências nas $k-1$ etapas anteriores. Esse resultado será apresentado a seguir.

**Teorema 2.1** (Teorema do produto)**.** Seja $(\Omega, \mathscr{A}, \mathbb{P})$ um espaço de probabilidade e $A_1, A_2, \ldots, A_n \in \mathscr{A}$ tais que $\mathbb{P}(A_i) > 0$ para todo $i$. Temos que

$$\mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_n) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2|A_1) \cdot \mathbb{P}(A_3|A_1 \cap A_2) \cdot \ldots \cdot \mathbb{P}(A_n|A_1 \cap \cdots \cap A_{n-1}) \tag{2.3}$$

A fórmula 2.3, à primeira vista, pode parecer estranha, mas, como já foi apontado anteriormente, a referida fórmula é a generalização da expressão 2.2. Sua aplicação é simples: para calcular a probabilidade da intersecção de $n$ eventos, basta efetuar o produto das probabilidades condicionais sucessivas.

**Exemplo 2.7.5.** Voltemos ao exemplo 2.4.1, da urna com 10 bolas numeradas de 1 a 10. Suponha que retiremos sucessivamente 4 bolas da urna, sem repor as bolas que forem retiradas, e que os eventos elementares sejam equiprováveis. Qual a probabilidade do número da primeira bola ser par, o da segunda ser ímpar, o da terceira ser par e o número da última, ímpar?
Vamos nomear os eventos da seguinte maneira:

- $A$: "o número da primeira bola é par";
- $B$: "o número da segunda bola é ímpar";
- $C$: "o número da terceira bola é par";
- $D$: "o número da quarta bola é ímpar".

A probabilidade pedida corresponde a $\mathbb{P}(A \cap B \cap C \cap D)$. Pela fórmula 2.3, temos

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

Usando a definição 2.2, temos que $\mathbb{P}(A) = \frac{5}{10}$. Para calcular a probabilidade do número da segunda bola retirada ser ímpar, entretanto, temos que considerar a ocorrência do evento $A$, afinal não estamos efetuando a reposição das bolas retiradas. Ou seja, precisamos calcular $\mathbb{P}(B|A) = \frac{5}{9}$. Da mesma forma, calculamos $\mathbb{P}(C|A \cap B) = \frac{4}{8}$ e $\mathbb{P}(D|A \cap B \cap C) = \frac{4}{7}$. Assim, temos que

$$\mathbb{P}(A \cap B \cap C \cap D) = \frac{5}{10} \cdot \frac{5}{9} \cdot \frac{4}{8} \cdot \frac{4}{7} = \frac{400}{5040} = \frac{5}{63}.$$

### 2.7.2 Teorema da Probabilidade Total e a Fórmula de Bayes

Para apresentar o próximo teorema, é preciso definir o conceito de partição do espaço amostral.

**Definição 2.7** (Partição de $\Omega$)**.** Dizemos que a coleção de eventos $\{A_1, A_2, \ldots, A_n\} \in \mathscr{A}$ forma uma **partição de** $\Omega$ se for composta de eventos mutuamente exclusivos (disjuntos), isto é, $A_i \cap A_j = \emptyset$ para todos $i$ e $j$ distintos, e sua união, denotada por $\bigcup\limits_{i=1}^{n} A_i$, for igual a $\Omega$.

Considerando essa definição, para qualquer evento $B \subset \Omega, B \in \mathscr{A}$, ao considerarmos uma partição $\{A_1, A_2, \ldots, A_n\}$ desse $\Omega$, temos que $B = \bigcup\limits_{i-1}^{n} (B \cap A_i)$. Mas disso conclui-se que

$$\mathbb{P}(B) = \mathbb{P}\left(\bigcup_{i-1}^{n} (B \cap A_i)\right) = \sum_{i-1}^{n} \mathbb{P}(B \cap A_i) = \sum_{i-1}^{n} \mathbb{P}(B|A_i) \cdot \mathbb{P}(A_i),$$

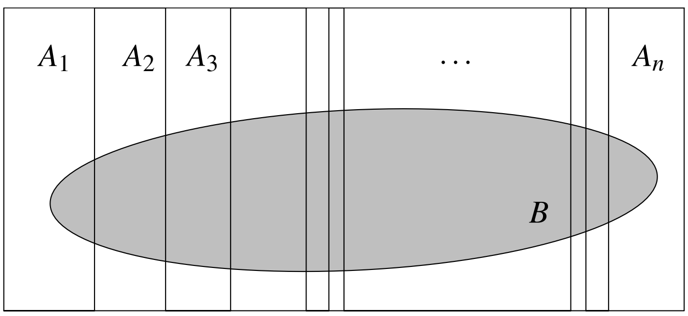
que é o resultado apresentado a seguir.

**Teorema 2.2** (Teorema da Probabilidade Total)**.** Seja $\{A_1, A_2, \ldots, A_n\}$ uma partição de $\Omega$ e $B$ um evento de $\Omega$. Sendo $\mathbb{P}(A_i) > 0$ para todo $i$, temos que

$$\begin{aligned}
\mathbb{P}(B) &= \mathbb{P}\left(\bigcup_{i=1}^{n} (B \cap A_i)\right) \\
&= \mathbb{P}(A_1 \cap B) + \cdots + \mathbb{P}(A_n \cap B) \\
&= \mathbb{P}(A_1) \cdot \mathbb{P}(B|A_1) + \cdots + \mathbb{P}(A_n) \cdot \mathbb{P}(B|A_n) \\
\mathbb{P}(B) &= \sum_{i=1}^{n} \mathbb{P}(B|A_i) \cdot \mathbb{P}(A_i). \qquad (2.4)
\end{aligned}$$

A Figura 2 ilustra a situação do Teorema 2.2.

Figura 2 – Ilustração do Teorema 2.2.




A Fórmula 2.4, obtida do Teorema 2.2, é conhecida como a fórmula das probabilidades totais. Ela permite que calculemos a probabilidade de um evento $B$ quando se conhecem as probabilidades de um conjunto de eventos disjuntos cuja união é o espaço amostral (uma partição de $\Omega$) e as probabilidades condicionais de $B$ dado cada um deles.

**Exemplo 2.7.6.** Em uma apresentação, um mágico precisa escolher uma bola de uma urna, sendo as bolas idênticas em relação à forma e ao peso. A primeira urna possui 5 bolas azuis e 5 bolas vermelhas, a segunda urna possui 3 bolas azuis e 7 bolas vermelhas e a terceira urna, por sua vez, possui 7 bolas azuis e 3 bolas vermelhas. A primeira urna tem $0,25$ de chance de ser escolhida pelo mágico, a segunda urna tem $0,35$ de chance e a terceira urna, $0,4$. Qual a chance do mágico escolher uma bola, nessas condições, e a bola sorteada ser da cor vermelha? Seja $B$ o evento "selecionar uma bola vermelha da urna escolhida" e $A_i$, para $1 \leqslant i \leqslant 3$, o evento "a $i$-ésima urna é escolhida". Temos, então, que

$$B = \bigcup_{i=1}^{3} (B \cap A_i) = (B \cap A_1) \cup (B \cap A_2) \cup (B \cap A_3).$$

Pelos dados do problema, temos que

$$\mathbb{P}(A_1) = \tfrac{25}{100}, \mathbb{P}(A_2) = \tfrac{35}{100}, \mathbb{P}(A_3) = \tfrac{40}{100}, \mathbb{P}(B|A_1) = \tfrac{5}{10}, \mathbb{P}(B|A_2) = \tfrac{7}{10}, \mathbb{P}(B|A_3) = \tfrac{3}{10}$$

Usando a fórmula 2.4 do Teorema 2.2, temos:

$$\begin{aligned}
\mathbb{P}(B) &= \sum_{i=1}^{n} (\mathbb{P}(B|A_i) \cdot \mathbb{P}(A_i)) \\
&= \mathbb{P}(A_1) \cdot \mathbb{P}(B|A_1) + \mathbb{P}(A_2) \cdot \mathbb{P}(B|A_2) + \mathbb{P}(A_3) \cdot \mathbb{P}(B|A_3) \\
&= \frac{25}{100} \cdot \frac{5}{10} + \frac{35}{100} \cdot \frac{7}{10} + \frac{40}{100} \cdot \frac{3}{10} \\
&= \frac{125}{1000} + \frac{245}{1000} + \frac{120}{1000} \\
\mathbb{P}(B) &= \frac{490}{1000} = \frac{49}{100}.
\end{aligned}$$

A fórmula a seguir é chamada de Fórmula de Bayes e permite uma interpretação vasta e profunda, responsável pelo desenvolvimento de uma abordagem da Estatística que é conhecida como *Estatística Bayesiana*.

**Teorema 2.3** (Fórmula de Bayes)**.** Seja $B$ um evento e $\{A_1, A_2, A_3\}$ uma partição do espaço amostral $\Omega$. Assumindo que $\mathbb{P}(A_i) > 0$ para $i \in \{1, 2, 3\}$, temos

$$\mathbb{P}(A_i|B) = \frac{\mathbb{P}(A_i) \cdot \mathbb{P}(B|A_i)}{\mathbb{P}(A_1) \cdot \mathbb{P}(B|A_1) + \mathbb{P}(A_2) \cdot \mathbb{P}(B|A_2) + \mathbb{P}(A_3) \cdot \mathbb{P}(B|A_3)}. \tag{2.5}$$

**Exemplo 2.7.7.** Uma loja de camisetas possui três fornecedores, $A_1$, $A_2$ e $A_3$. A participação dos fornecedores nos produtos da loja é, respectivamente, $0,15$, $0,35$ e $0,50$ e as probabilidades dos fornecedores produzirem camisetas defeituosas são $0,01$, $0,05$ e $0,02$, respectivamente. Uma camiseta é escolhida, ao acaso, no estoque onde ficam todas as camisetas da loja. Infelizmente, o lojista nota que tal camiseta está com um defeito de fabricação. Qual é a probabilidade de esta camiseta tenha sido produzida pela fábrica $A_1$? Seja $B$ o evento "a camiseta não tem defeito" e $B^c$ o evento "a camiseta tem defeito". Sabemos que $\mathbb{P}(A_1) = 0,15$, $\mathbb{P}(A_2) = 0,35$ e

$\mathbb{P}(A_3) = 0,50$, pois a camiseta é escolhida ao acaso do conjunto total de camisetas, observando-se a participação de cada fornecedor no total. Além disso, também sabemos que $\mathbb{P}(B^c|A_1) = 0,01$, $\mathbb{P}(B^c|A_2) = 0,05$ e $\mathbb{P}(B^c|A_3) = 0,02$. Queremos calcular $\mathbb{P}(A_1|B^c)$, isto é, a probabilidade da camiseta escolhida ter sido fornecida pela fábrica $A_1$, sabendo que a peça escolhida é defeituosa. Usando a Fórmula de Bayes (2.5), para $i = 1$, temos:

$$\begin{aligned}
\mathbb{P}(A_1|B^c) &= \frac{\mathbb{P}(A_1)\cdot\mathbb{P}(B^c|A_1)}{\mathbb{P}(A_1)\cdot\mathbb{P}(B^c|A_1)+\mathbb{P}(A_2)\cdot\mathbb{P}(B^c|A_2)+\mathbb{P}(A_3)\cdot\mathbb{P}(B^c|A_3)} \\
&= \frac{\frac{15}{100}\cdot\frac{1}{100}}{\frac{15}{100}\cdot\frac{1}{100}+\frac{35}{100}\cdot\frac{5}{100}+\frac{50}{100}\cdot\frac{2}{100}} = \frac{\frac{15}{10000}}{\frac{15}{10000}+\frac{175}{10000}+\frac{100}{10000}} = \\
\mathbb{P}(A_1|B^c) &= \frac{15}{290} = \frac{3}{58} \approx 0,0517.
\end{aligned}$$

A Fórmula de Bayes nos permite uma diferente interpretação do problema anterior: *como os fornecedores $A_1$, $A_2$ e $A_3$ são responsáveis, respectivamente, por $0,15$, $0,35$ e $0,50$ do estoque de camisetas da loja, se escolhermos uma camisa do estoque ao acaso, as probabilidades de que essa camisa tenha vindo dos fornecedores $A_1$, $A_2$ e $A_3$ são, respectivamente, iguais a $0,15$, $0,35$ e $0,50$. Por outro lado, se escolhemos a camisa ao acaso, desconhecemos seu fornecedor e verificamos que ela é defeituosa, então, levando em conta essa informação proveniente do experimento, a probabilidade de que a peça tenha vindo da fábrica $A_1$ passa a valer cerca de* $0,0517$.

## 2.8 Independência de eventos

A definição de independência de eventos exprime a ideia intuitiva da não-influência de um evento $A$ sobre a ocorrência (ou não) de um outro evento $B$, ou seja, a ocorrência de $A$ não melhora nossa posição para "predizer" a ocorrência de $B$. Em muitas ocasiões, cabe ao observador do experimento aceitar a hipótese de que dois eventos sejam independentes. Intuitivamente, dados dois eventos $A$ e $B$, com $\mathbb{P}(B) > 0$, podemos dizer que o evento $A$ é *independente* do evento $B$ se

$$\mathbb{P}(A|B) = \mathbb{P}(A) \tag{2.6}$$

Formalmente, diz-se que

**Definição 2.8** (Eventos independentes)**.** Dados dois eventos $A$ e $B$ em um $\Omega$, com $\mathbb{P}(B) > 0$, podemos dizer que o evento $A$ é *independente* do evento $B$ se

$$\mathbb{P}(A \cap B) = \mathbb{P}(B)\cdot\mathbb{P}(A) \tag{2.7}$$

A definição formal corrobora a ideia intuitiva, pois usando a definição 2.1 na definição 2.8, temos que

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

Se $A \cap B = \emptyset$, temos que $\mathbb{P}(A \cap B) = \mathbb{P}(\emptyset) = 0 = \mathbb{P}(A) \cdot \mathbb{P}(B)$, o que nos leva a ver que $A$ e $B$ não são independentes a menos que um deles tenha probabilidade zero.

Se um evento $A$ é independente de um evento $B$, considerando $\mathbb{P}(B), \mathbb{P}(A) > 0$, é intuitivo pensar que $B$ também é independente de $A$. De fato, nas condições acima, é fácil ver que

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

Uma outra consequência imediata dessa definição é que o conjunto vazio $\emptyset$ e o espaço amostral $\Omega$ são independentes de qualquer outro evento, pois se $A$ é um evento de $\Omega$, então

$$\mathbb{P}(A \cap \emptyset) = \mathbb{P}(\emptyset) = 0 = \mathbb{P}(\emptyset) \cdot \mathbb{P}(A)$$

e

$$\mathbb{P}(A \cap \Omega) = \mathbb{P}(A) = \mathbb{P}(A) \cdot 1 = \mathbb{P}(A) \cdot \mathbb{P}(\Omega).$$

Veremos, mais adiante, que a independência de eventos depende de como a probabilidade foi definida. Entretanto, como $\mathbb{P}(\Omega) = 1$ e $\mathbb{P}(\emptyset) = 0$, o resultado exposto acima sempre será válido.

**Exemplo 2.8.1.** Em uma urna há duas bolas: uma verde e uma amarela. Retiramos uma bola dessa urna e anotamos sua cor. Antes de efetuar a segunda retirada, devolvemos a bola que retiramos à urna, de forma que essa bola tem novamente, a probabilidade de ser retirada no próximo movimento. Dessa forma, considerando duas retiradas com reposição, adotando $G$ para a bola verde e $Y$ para a bola amarela, nosso espaço amostral é o conjunto de pares ordenados $\Omega = \{(G,Y),(G,G),(Y,Y),(Y,G)\}$, onde o primeiro elemento de cada par ordenado $(i, j)$ representa a primeira bola retirada e o segundo elemento, a segunda bola retirada após a reposição. Vamos supor que cada par de resultados possíveis seja equiprovável, ou seja,

$$\mathbb{P}(\{(G,Y)\}) = \mathbb{P}(\{(G,G)\}) = \mathbb{P}(\{(Y,G)\}) = \mathbb{P}(\{(Y,Y)\}) = 0,25.$$

Consideremos os seguintes eventos:

- $G_1$: "obter bola verde no primeiro lançamento", ou seja, $G_1 = \{(G,Y),(G,G)\}$ e $\mathbb{P}(G_1) = 0,50$;
- $Y_2$: "obter bola amarela no segundo lançamento", ou seja, $Y_2 = \{(G,Y),(Y,Y)\}$ e $\mathbb{P}(Y_2) = 0,50$;
- $G_1 \cap Y_2$: "obter bola verde no primeiro lançamento e bola amarela no segundo lançamento", de forma que $G_1 \cap Y_2 = \{(G,Y)\}$, isto é, $\mathbb{P}(G_1 \cap Y_2) = 0,25$.

Calculando a probabilidade do evento $Y_2$, supondo a ocorrência do evento $G_1$, temos

$$\mathbb{P}(Y_2|G_1) = \frac{\mathbb{P}(Y_2 \cap G_1)}{\mathbb{P}(Y_2)} = \frac{\mathbb{P}(G_1 \cap Y_2)}{\mathbb{P}(Y_2)} = \frac{\frac{25}{100}}{\frac{50}{100}} = \frac{25}{50} = \frac{1}{2} = 0,50 = \mathbb{P}(Y_2).$$

Logo, pode-se concluir que $Y_2$ é independente de $G_1$.

Nosso objetivo com o próximo exemplo é exibir a mesma situação em condições ligeiramente distintas para chegarmos a uma conclusão importante.

**Exemplo 2.8.2.** Vamos considerar o mesmo experimento do exemplo anterior, com os mesmos eventos, mas vamos atribuir probabilidades diferentes para cada evento elementar:

$$\mathbb{P}(\{(G,Y)\}) = 0,10; \mathbb{P}(\{(G,G)\}) = 0,20; \mathbb{P}(\{(Y,G)\}) = 0,30 \text{ e } \mathbb{P}(\{(Y,Y)\}) = 0,40.$$

Assim, temos:

$$\begin{aligned} \mathbb{P}(G_1) &= \mathbb{P}(\{(G,Y)\}) + \mathbb{P}(\{(G,G)\}) = 0,10 + 0,20 = 0,30 \\ \mathbb{P}(Y_2) &= \mathbb{P}(\{(G,Y)\}) + \mathbb{P}(\{(Y,Y)\}) = 0,10 + 0,40 = 0,50 \\ \mathbb{P}(G_1 \cap Y_2) &= \mathbb{P}(\{(G,Y)\}) = 0,10 \end{aligned}$$

Veja que, neste caso, temos

$$\mathbb{P}(Y_2|G_1) = \frac{\mathbb{P}(Y_2 \cap G_1)}{\mathbb{P}(Y_2)} = \frac{\mathbb{P}(G_1 \cap Y_2)}{\mathbb{P}(Y_2)} = \frac{\frac{10}{100}}{\frac{50}{100}} = \frac{10}{50} = \frac{1}{5} = 0,20 \neq \mathbb{P}(Y_2).$$

Assim, como $\mathbb{P}(Y_2|G_1) \neq \mathbb{P}(Y_2)$, temos que, neste caso, $Y_2$ não é independente de $G_1$.

Comparando os exemplos 2.8.1 e 2.8.2 acima, é possível chegar à seguinte conclusão: *a independência de eventos* depende *da distribuição da probabilidade em* $\mathscr{A}$ (no caso, $\mathscr{P}(\Omega)$).

**Exemplo 2.8.3.** Considere o lançamento duplo de uma moeda (2.1.2). Vamos definir os três eventos a seguir:

- $A$: "o primeiro lançamento resulta *cara*";
- $B$: "o segundo lançamento resulta *coroa*";
- $C$: "ambos os lançamentos resultam o mesmo tipo de face".

Supondo que as moedas são balanceadas (os eventos elementares tem a mesma probabilidade de ocorrer), temos que

$$\mathbb{P}(A) = \mathbb{P}(B) = \mathbb{P}(C) = 0,50$$
$$\mathbb{P}(A \cap B) = \mathbb{P}(A \cap C) = \mathbb{P}(B \cap C) = 0,25$$

Observe que, a partir daí, temos

$$\begin{aligned}
\mathbb{P}(A\cap B) &= \mathbb{P}(A)\cdot\mathbb{P}(B)\\
\mathbb{P}(A\cap C) &= \mathbb{P}(A)\cdot\mathbb{P}(C)\\
\mathbb{P}(B\cap C) &= \mathbb{P}(B)\cdot\mathbb{P}(C)
\end{aligned}$$

Assim, podemos afimar que os três eventos são independentes dois a dois, mas não três a três, afinal

$$\mathbb{P}(A\cap B\cap C)=0\neq\mathbb{P}(A)\cdot\mathbb{P}(B)\cdot\mathbb{P}(C).$$

Através dos exemplos anteriores, pode-se perceber que para provar que 2 eventos são independentes, basta provar uma igualdade (a saber, $\mathbb{P}(A\cap B)=\mathbb{P}(A)\cdot\mathbb{P}(B)$). Entretanto, como sugere o último exemplo apresentado, para provar a independência entre 3 eventos, precisamos verificar 4 igualdades, como aponta a definição a seguir.

**Definição 2.9.** Diremos que três eventos $A$, $B$ e $C$ são mutualmente independentes se, e somente se, todas as quatro condições a seguir forem verificadas:

$$\begin{aligned}
\mathbb{P}(A\cap B) &= \mathbb{P}(A)\cdot\mathbb{P}(B)\\
\mathbb{P}(A\cap C) &= \mathbb{P}(A)\cdot\mathbb{P}(C)\\
\mathbb{P}(B\cap C) &= \mathbb{P}(B)\cdot\mathbb{P}(C)\\
\mathbb{P}(A\cap B\cap C) &= \mathbb{P}(A)\cdot\mathbb{P}(B)\cdot\mathbb{P}(C)
\end{aligned}$$

A definição a seguir generaliza a definição anterior, para $n$ eventos.

**Definição 2.10.** Os eventos $A_1,A_2,\ldots,A_n$, tais que $n\geqslant 2$, são coletivamente independentes se

$$\mathbb{P}(A_{i_1}\cap A_{i_1}\cap\cdots A_{i_m})=\mathbb{P}(A_{i_1})\cdot\mathbb{P}(A_{i_2})\cdot\cdots\cdot\mathbb{P}(A_{i_m}),$$

para todo $1\leqslant i_1\leqslant i_2\leqslant\cdots\leqslant i_m\leqslant n$, $m\in\{2,3,4,\ldots\}$.

## 2.8.1 Lema de Borel-Cantelli

Fecharemos este capítulo com o enunciado do Lema de Borel-Cantelli para, com ele, podermos compreender um dos exemplos contra-intuitivos que exploraremos no próximo capítulo. Antes disso, precisamos da definição a seguir.

**Definição 2.11** (Limite superior e inferior)**.** Em termos de ocorrência de eventos em sequência de eventos sucessivos $\{A_n\}$,$n\in\mathbb{N}$, define-se o limite superior ($\limsup$) e o limite inferior ($\liminf$) da seguinte maneira:

- o *limite superior* é o conjunto formado por todos os elementos que pertencem a uma infinidade de conjuntos $A_n$: $\limsup\limits_{n\to\infty} A_n=\bigcap\limits_{n=1}^{\infty}\bigcup\limits_{k=n}^{\infty}A_k$.

- o *limite inferior* é o conjunto formado por todos os elementos que pertencem a cada um dos $A_n$, exceto por um número finito deles: $\liminf_{n\to\infty} A_n = \bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty} A_k$.

Podemos, então, enunciar o seguinte lema:

**Lema 1** (Lema de Borel-Cantelli)**.** Seja $\{A_n\}$, $n \in \mathbb{N}$ uma sucessão de eventos e defina $A = \limsup_{n\to\infty} A_n$.

1. Se $\sum_{n=1}^{\infty} \mathbb{P}(A_n) \leq \infty$, então $\mathbb{P}(A) = 0$.
2. Se $A_1, A_2, \ldots$ forem independentes e $\sum_{n=1}^{\infty} \mathbb{P}(A_n) = \infty$, então $\mathbb{P}(A) = 1$.

**Exemplo 2.8.4.** Seja $A$ o evento "*ganhar na loteria*" e $\mathbb{P}(A) > 0$. Suponha que o experimento $A$ repete-se várias vezes. Podemos definir o seguinte evento:

$$A_k = \text{“ocorre } A \text{ na tentativa k''}, \text{ k = 1, 2, 3, } \ldots$$

Os eventos $A_k$ tem as seguintes propriedades:

- $\mathbb{P}(A_k) = \mathbb{P}(A) > 0$.
- $\sum_{k=1}^{\infty} \mathbb{P}(A_k) = \infty$
- $A_1, A_2, \ldots$ são independentes entre si.

Pelo Lema de Borel-Cantelli, teremos

$$\mathbb{P}\left(\limsup_{k\to\infty} A_k\right) = 1$$

Podemos interpretar esse resultado da seguinte maneira: o evento $\limsup$ corresponde à ocorrência de uma infinidade de eventos $A_1, A_2, \ldots$. Isso significa que se nos mantivermos jogando na loteria de maneira permanente, mais cedo ou mais tarde certamente ganharemos o prêmio com probabilidade 1, sendo que poderemos ganhá-lo até mais de uma vez: basta continuar jogando.

O leitor que desejar aprofundar mais os conceitos ligados à Probabilidade pode consultar Dantas (2004).

CAPÍTULO

3

# ALGUNS PROBLEMAS CONTRA-INTUITIVOS

Apresentamos aqui alguns problemas de probabilidade que envolvem soluções contra-intuitivas. Esses e outros problemas podem ser encontrados em Isaac (1995), Bernstein (1996) e Saldanha (1998).

## 3.1 A irmã do rei

O problema a seguir pode ser utilizado para testar nosso conhecimento de probabilidade condicional. Apesar de curto, esse problema também nos mostrará que devemos ser cuidadosos ao interpretar as informações oferecidas pelo enunciado de um problema. Leia-o a seguir.

> *Um rei é proveniente de uma família de duas crianças. Qual a probabilidade do rei ter uma irmã?*

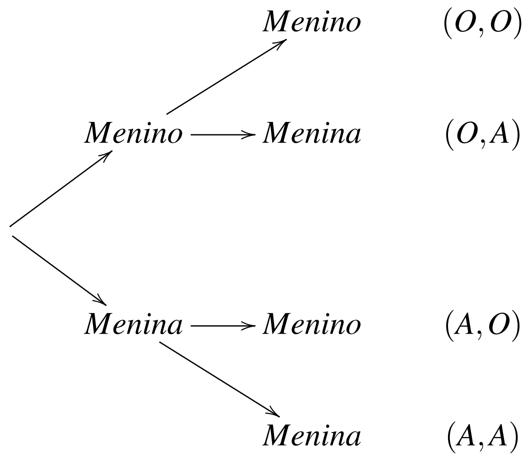
Sabemos que o primeiro passo para resolver um problema envolvendo probabilidade é descrever um espaço amostral adequado. Na situação em questão, podemos considerar como espaço amostral $\Omega$ o conjunto de pares ordenados

$$\Omega = \{(O,O),(O,A),(A,O),(A,A)\},$$

onde o primeiro elemento de cada par ordenado indica a criança mais velha, o segundo elemento de cada par ordenado indica a criança mais jovem e as letras $A$ e $O$ indicam o gênero da criança: $A$ para *menina* e $O$ para *menino*. Organizamos na figura 3 a seguir os resultados possíveis para este experimento.

Vamos supor que cada evento elementar do espaço amostral tenha a mesma probabilidade de ocorrer, isto é, que $\mathbb{P}(\{(O,O)\}) = \mathbb{P}(\{(O,A)\}) = \mathbb{P}(\{(A,O)\}) = \mathbb{P}(\{(A,A)\}) = \frac{1}{4}$.

Figura 3 – Figura ilustrativa do espaço amostral relativo ao problema da Irmã do Rei




Considerando os eventos $R$ : "*uma das crianças é o rei*" e $I$ : "*uma das crianças é uma menina*", pode-se concluir que responder à pergunta do problema, "*qual a probabilidade do rei ter uma irmã?*", é equivalente a calcular a probabilidade do evento $I|R$: "*uma das crianças é uma menina **dado que** uma das crianças é o rei*". Observe que o evento "*uma das crianças é o rei*" é equivalente ao evento "uma das crianças é um menino". Além disso, $I = \{(O,A),(A,O),(A,A)\}$, $R = \{(O,A),(A,O),(O,O)\}$ e $I \cap R = \{(O,A),(A,O)\}$, o que nos permite concluir que $\mathbb{P}(I) = \mathbb{P}(R) = \frac{3}{4}$ e que $\mathbb{P}(I \cap R) = \frac{2}{4}$. Usando a fórmula da probabilidade condicional, temos:

$$\mathbb{P}(I|R) = \frac{\mathbb{P}(I \cap R)}{\mathbb{P}(R)} = \frac{2/4}{3/4} = \frac{2}{3}$$

Esse resultado pode parecer um tanto quanto contraintuitivo. Nossa intuição pode nos levar a acreditar que a probabilidade pedida é $\frac{1}{2}$, pois pode-se argumentar que, como a família possui duas crianças, sendo uma delas o próprio rei, então restam apenas duas possibilidades para o gênero da segunda criança (menino ou menina), das quais apenas uma é favorável ao que se pede no enunciado $\left(\text{o que conduz à incorreta probabilidade de } \frac{1}{2}\right)$.

Se a pergunta do problema fosse "*qual a probabilidade de uma das crianças ser uma menina?*", então a resposta seria $\frac{1}{2}$. Mas está implícita no enunciado a informação de que uma das crianças, o rei, é um menino, fazendo com que possamos eliminar o par $(A,A)$ de $\Omega$. Os outros três resultados de $\Omega$ formam o espaço amostral "atualizado" $\Omega_1 = \{(O,A),(A,O),(O,O)\}$, do qual apenas dois resultados possuem um menino e uma menina.

Para alguns, pode parecer que esse resultado é fruto de um "jogo de números", mas é possível chegar a esse mesmo resultado usando o seguinte raciocínio:

*Como queremos descobrir a chance do **rei** ter uma irmã, fica implícita a informação de que um dos filhos da família é um menino, logo podemos desconsiderar o par $(A,A)$ do espaço amostral, ficando este restrito ao conjunto $\Omega_1 = \{(O,O),(O,A),(A,O)\}$. Supondo que os eventos elementares desse espaço amos-*

*tral são igualmente prováveis, a chance de um dos pares corresponder ao rei e sua irmã (ou à irmã do rei e ele, nessa ordem) é $\frac{2}{3}$.*

Esse problema ilustra como temos que ser cuidadosos no momento de interpretar as informações que um problema nos oferece. É necessário, portanto, levar em conta as ambiguidades do enunciado, pois diferentes interpretações podem nos guiar a espaços amostrais totalmente diferentes e, em consequência, a resultados errados.

## 3.2 Analisando a "Porta dos Desesperados"

Nos anos 80 e 90, vários programas brasileiros de televisão recorreram a um quadro envolvendo prêmios e portas. Um dos mais famosos desses quadros foi a "Porta dos Desesperados", exibido nos programas televisivos do apresentador Sérgio Mallandro. Esse e outros quadros são adaptações de um quadro original do programa "*Let's Make a Deal*", exibido na televisão norte-americana nos anos 70. O quadro do programa original americano deu origem ao "*Problema de Monty Hall*", nome dado em razão do apresentador daquele programa de televisão (SALOMÃO, 2014, p. 36). Esse problema foi publicado em jornais da época e até hoje encontra discussões nas redes sociais, chamando a atenção não só de matemáticos, mas do público em geral. A situação pode ser resumida nas seguintes linhas:

*Um apresentador seleciona um participante da plateia e o coloca em frente a três portas fechadas, uma das quais esconde um prêmio (um carro, por exemplo). Atrás das outras duas portas, entretanto, podem estar escondidos cabras, bodes ou monstros, por exemplo. O participante é, então, convidado a escolher uma das três portas. Se ele escolher a porta que esconde o prêmio, ele o leva para casa. Antes de abrir a porta escolhida pelo participante, o apresentador abre uma das outras duas portas, optando por uma que não esconde o prêmio. Dada essa informação, o apresentador oferece ao participante a opção de trocar a porta que escolhera inicialmente. É mais vantajoso manter a porta escolhida ou trocá-la?*

A resposta mais comum dada a essa pergunta é que trocar a porta é irrelevante, pois cada uma das duas portas fechadas teria a mesma chance de esconder o prêmio. Dessa maneira, pode-se pensar que a chance do participante ganhar o prêmio era de $\frac{1}{3}$ e que, após o apresentador abrir a porta, a chance de ganhar o prêmio aumente para $\frac{1}{2}$. Respostas como essas foram frequentemente publicadas em jornais, inclusive por autoria de matemáticos de renome, como informa Isaac (1995, p. 1). Apesar disso, a resposta correta do problema é contra-intuitiva pois, como veremos, é mais vantajoso trocar de porta. Mais do que isso: veremos que é duas vezes mais provável que o participante ganhe o prêmio se optar por trocar de porta ao invés de não fazê-lo.

Para chegar a esse resultado, é preciso construir um espaço amostral adequado ao experimento. Para isso, vamos melhorar seu enunciado, descrevendo-o matematicamente com mais precisão e exatidão para evitar mal entendidos e ambiguidades. Suponhamos, então, que o participante do jogo sempre decida trocar de porta e analisar as implicações que essa decisão acarreta. Mais à frente, analisaremos o que acontece se o participante sempre optar por manter sua escolha inicial. Dessa forma, podemos dizer que o jogo é formado por três ações:

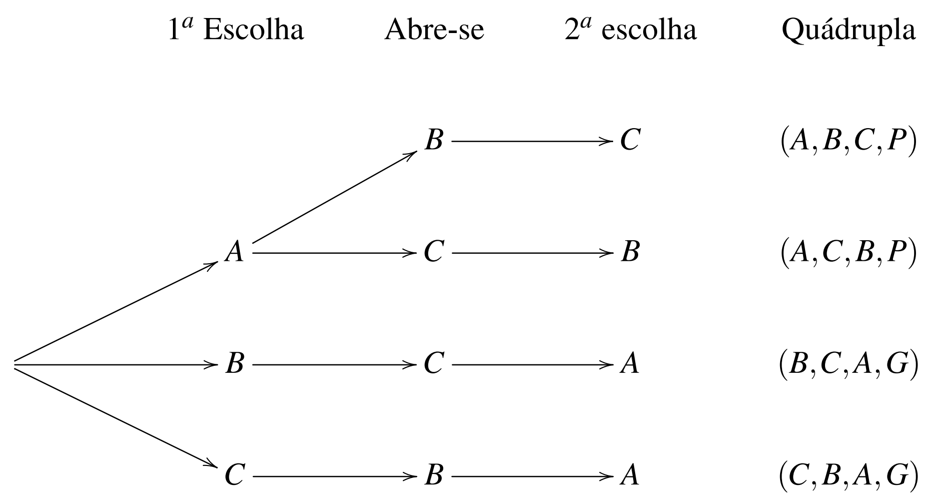
1. o participante faz a primeira escolha de uma das três possíveis portas: $A$, $B$ ou $C$;
2. o apresentador abre uma das duas portas não escolhidas pelo participante, revelando um bode; e
3. o participante troca a porta que escolhera inicialmente pela outra porta disponível.

Sem perder a generalidade do problema, suporemos que o prêmio desejado pelo participante trata-se um carro que está escondido atrás da porta $A$ enquanto cada uma das outras duas portas, $B$ e $C$, esconde um bode. Para facilitar a notação, podemos descrever um resultado desse experimento através de uma quádrupla ordenada $(p,q,r,s)$ onde $p$ indica a porta escolhida inicialmente pelo participante, $q$ indica a porta que o apresentador abriu, $r$ indica a nova porta escolhida pelo participante e $s$ indica se o participante ganhou o carro $(G)$ ou não ganhou o carro $(P)$. Assim, por exemplo, a quádrupla $(A,B,C,P)$ indica que o participante escolheu inicialmente a porta $A$, que esconde o carro, de acordo com nossa suposição inicial; o apresentador abriu a porta $B$ em seguida, revelando um bode; o participante, após isso, trocou sua escolha inicial para a porta $C$, perdendo $(P)$ o carro. O espaço amostral, portanto, pode ser descrito como o conjunto

$$\Omega = \{(A,B,C,P),(A,C,B,P),(B,C,A,G),(C,B,A,G)\}.$$

Observe a organização dos elementos de $\Omega$ na Figura 4.

Figura 4 – Figura ilustrativa do espaço amostral relativo ao problema de Monty Hall




Veja que a última letra da quádrupla é apenas uma convenção que utilizaremos para identificar facilmente o resultado da sequência de movimentos utilizada: é possível descrever todo o jogo apenas com ternos ordenados usando as primeiras três posições de cada uma das quádruplas apresentadas acima. Observe, também que se o participante escolher as portas $B$ e $C$ então, pelas regras fixadas, ele obrigatoriamente trocará de porta e, portanto, ganhará o prêmio, como pode ser visto nos dois últimos resultados apresentados em $\Omega$ acima. Se o participante escolher inicialmente a porta $A$, dependendo da sua escolha de troca, há duas chances diferentes do participante perder, como exibido nos dois primeiros resultados apresentados em $\Omega$.

Para completar a resolução do problema, é preciso definir a probabilidade dos eventos de $\Omega$. Para isso, vamos analisar a situação desde o começo: estando o participante em frente às três portas antes de fazer sua escolha inicial, em quê ele pode se basear para tomar a decisão de qual porta escolher? Assumindo que ele não tem motivos que favorecem a escolha de uma ou outra porta, provavelmente ele escolheria a porta “*ao acaso*”. Dessa maneira, podemos pensar que cada uma das três portas é igualmente provável de ser escolhida inicialmente pelo participante. Diremos, então, que cada porta tem probabilidade $\frac{1}{3}$ de ser escolhida.

Retomemos, agora, a análise de $\Omega$. Suponha que o participante escolha inicialmente a porta $B$, com uma probabilidade igual a $\frac{1}{3}$. O único resultado possível em $\Omega$ com a porta $B$ como escolha inicial é $(B,C,A,G)$. Por isso,

$$\mathbb{P}(\{(B,C,A,G)\}) = \frac{1}{3}.$$

Analogamente, como a porta $C$ também tem a probabilidade $\frac{1}{3}$ de ser escolhida pelo participante na primeira parte do experimento, temos

$$\mathbb{P}(\{(C,B,A,G)\}) = \frac{1}{3}.$$

Por fim, a porta $A$ também pode ser escolhida inicialmente pelo participante com probabilidade $\frac{1}{3}$. Entretanto, ao escolher inicialmente a porta $A$, o participante perderá, podendo fazê-lo de duas maneiras distintas. Assim, podemos afirmar que o evento $A_1$ : “*escolhe-se inicialmente a porta A*” dado por $A_1 = \{(A,B,C,P),(A,C,B,P)\}$, tem probabilidade de $\frac{1}{3}$, mesmo que as probabilidades individuais dos eventos elementares $\{(A,B,C,P)\}$ e $\{(A,C,B,P)\}$ não tenham sido definidas. Veremos que não precisamos definir essas probabilidades individuais para resolver nosso problema. Sejam, então, $\mathbb{P}(\{(A,B,C,P)\}) = u$ e $\mathbb{P}(\{(A,C,B,P)\}) = v$ de forma que $u+v = \frac{1}{3}$. Teremos

$$\mathbb{P}(\{(A,B,C,P),(A,C,B,P)\}) = \mathbb{P}(A_1) = \frac{1}{3}.$$

Sendo $V$ o evento em que o participante *vence*, ou seja, ganha o carro, pode-se observar que $V = \{(B,C,A,G),(C,B,A,G)\}$. Pelas regras fixadas, observando que os eventos elementares são disjuntos, temos que

$$\mathbb{P}(V) = \mathbb{P}(\{(B,C,A,G)\} \cup \{(C,B,A,G)\}) = \mathbb{P}(\{(B,C,A,G)\}) + \mathbb{P}(\{(C,B,A,G)\}) = \frac{1}{3} + \frac{1}{3} = \frac{2}{3}.$$

Da mesma forma, a chance do evento $V^c$ (o participante não ganha o carro) é dada por

$$\mathbb{P}(V^c) = \mathbb{P}(A_1) = \frac{1}{3}.$$

Isso, portanto, responde a questão inicial: de acordo com nossas suposições sobre o jogo, trocar a porta escolhida inicialmente dá ao participante a probabilidade $\frac{2}{3}$ de ganhar e $\frac{1}{3}$ de perder.

Como dito anteriormente, também podemos inverter a suposição inicial e nos perguntarmos: o que acontece quando não se troca a porta escolhida inicialmente? Assim como fizemos, vamos identificar o espaço amostral $\Omega'$ considerando a "não troca" da porta, usando a mesma notação. O espaço amostral $\Omega'$ pode ser escrito como

$$\Omega' = \{(A,B,A,G),(A,C,A,G),(B,C,B,P),(C,B,C,P)\}$$

Veja que o terceiro elemento é igual ao primeiro elemento em cada uma das quádruplas, indicando que não houve troca da porta escolhida inicialmente. Através de argumentos análogos aos que usamos no caso da troca e assumindo que cada porta tem a mesma probabilidade de ser escolhida inicialmente, podemos calcular que

$$\mathbb{P}(V^c) = \mathbb{P}(\{(B,C,B,P)\}) + \mathbb{P}(\{(C,B,C,P)\}) = \frac{1}{3} + \frac{1}{3} = \frac{2}{3}.$$

Logo,

$$\mathbb{P}(V) = \mathbb{P}(\{(A,B,A,G)\}) + \mathbb{P}(\{(A,C,A,G)\}) = 1 - \mathbb{P}(V^c) = 1 - \frac{2}{3} = \frac{1}{3}.$$

Vê-se, então, que trocar a porta escolhida inicialmente dará ao participante a probabilidade $\frac{2}{3}$ de ganhar o carro, enquanto manter a escolha inicial lhe dará apenas $\frac{1}{3}$ de chance de ganhá-lo. Dessa forma, podemos concluir que a chance do participante ganhar o prêmio é maior quando ele opta por trocar de porta (mais precisamente, a chance de ganhar o prêmio dobra ao tomar tal decisão).

O pensamento comum de que o apresentador, após revelar uma porta que esconde um bode, daria ao participante um novo problema, com apenas duas portas e um carro, nos leva ao raciocínio incorreto de que a chance do carro estar em qualquer uma das duas portas seria $0,50$ e que não faria diferença trocar de porta. Mas o que não se leva em questão quando se pensa dessa maneira é que, na verdade, a porta que o apresentador abre depende da porta que foi escolhida inicialmente pelo participante. Como o apresentador sabe desde o início onde está o prêmio, ele nunca abrirá a porta que esconde o carro. Ao abrir uma das portas, portanto, ele não está criando um novo jogo, mas dando uma informação nova sobre o jogo original. Assim, como o participante fez ao escolher inicialmente alguma porta, podemos acreditar que o apresentador abriu uma porta ao acaso. Mas, como dito antes, se o participante tiver escolhido inicialmente uma porta não-premiada, o apresentador não tem nenhuma liberdade de escolha e só pode abrir uma porta: a outra que não está premiada. Se o apresentador abrir uma porta ao acaso, correndo o risco de revelar o carro e assim acabar com o jogo, então só nesse caso não fará diferença trocar de porta escolhida inicialmente. Nesse caso, portanto, a nova informação altera a chance de vitória do participante. Veremos em outro problema, mais adiante, que uma nova informação pode não alterar a probabilidade de algo acontecer.

É possível perceber, portanto, a importância de se converter a descrição do problema original para uma linguagem matemática precisa e assertiva fazendo algumas suposições iniciais condizentes com a situação-problema. O conhecimento da disposição inicial do carro e dos bodes pelo apresentador afeta imediatamente a nossa percepção da situação, daí a importância, por exemplo, de enumerar corretamente o espaço amostral nesse problema e atribuir a cada um de seus elementos a probabilidade correspondente.

## 3.3 O dilema do prisioneiro

O problema a seguir também é interessante para se pensar sobre probabilidade condicional. Há várias maneiras de se enunciar esse problema, sendo uma delas a seguinte:

> *Numa cela da prisão, há três prisioneiros: A, B e C. Dois desses prisioneiros serão soltos e os prisioneiros sabem disso, mas não sabem quais deles receberão esse benefício. O prisioneiro A, então, pede ao guarda que lhe diga a identidade de um dos prisioneiros que será solto: B ou C. O guarda se nega a responder e se explica, dizendo ao prisioneiro A: "*sua probabilidade de ser solto é $\frac{2}{3}$ mas se eu lhe disser, por exemplo, que o prisioneiro $B$ será solto, então só haveriam mais dois prisioneiros para serem escolhidos e sua chance cairia para $\frac{1}{2}$, por isso não irei te dizer qual dos outros receberá a liberdade, senão você ficará preocupado com isso*". O raciocínio do guarda está correto?*

A resposta a esse problema não é tão simples nem intuitiva. Para encontrá-la, é preciso analisar melhor a situação e entender do ponto de vista matemático a afirmação do guarda. Ao fazê-la, o guarda está pensando em todas as possibilidades de dois prisioneiros serem soltos. Essas possibilidades podem ser representadas pelos resultados $AB$, $AC$ e $BC$, onde cada dupla corresponde ao par de prisioneiros que será liberto da prisão. Dessa maneira, o guarda pensou no espaço amostral $\Omega_G = \{AB, AC, BC\}$, que é condizente com a situação do problema. Como há 2 dos 3 resultados de $\Omega_G$ que favorecem a libertação do prisioneiro $A$, podemos perceber que, de acordo com sua afirmação inicial, o guarda pressupôs que o par de prisioneiros a receber a liberdade será selecionado ao acaso. Sendo assim, cada par tem a mesma probabilidade $\frac{1}{3}$ de ser escolhido e, por isso, a primeira afirmação sobre a probabilidade inicial do prisioneiro $A$ receber a liberdade ser $\frac{2}{3}$ é correta. O problema aparece quando o guarda diz que a probabilidade de $A$ ser solto, dado que $B$ será solto, é $\frac{1}{2}$, ou seja:

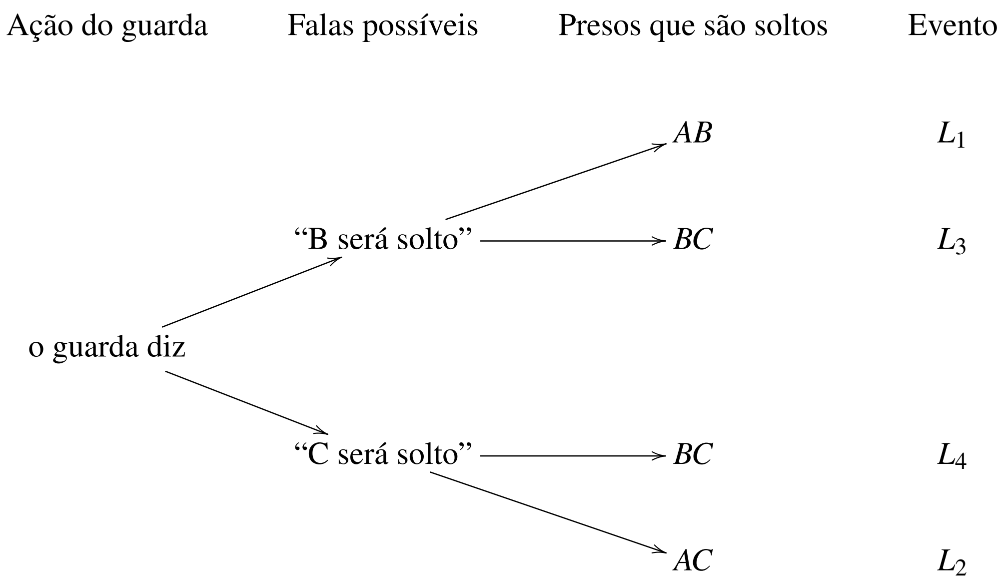
$$\mathbb{P}(\text{“}A \text{ ser solto”} | \text{“o guarda diz que } B \text{ será solto”}) = \frac{1}{2}.$$

Para analisar a veracidade dessa afirmação, a primeira coisa que precisamos levar em conta é que a probabilidade condicional dada acima não pode ser computada em termos do espaço amostral dado anteriormente $(\Omega_G)$, pois nós não tínhamos a informação "*o guarda diz que B será solto*" como condição ao descrevê-lo. Daí a necessidade de descrever um espaço amostral mais complexo $\Omega_L$, incorporando a segunda afirmação do guarda. Considere, então, o espaço amostral dado pelos seguintes resultados:

$$\begin{aligned} L_1 &= \{AB, \text{o guarda diz que } B \text{ será solto}\}; \\ L_2 &= \{AC, \text{o guarda diz que } C \text{ será solto}\}; \\ L_3 &= \{BC, \text{o guarda diz que } B \text{ será solto}\}; \\ L_4 &= \{BC, \text{o guarda diz que } C \text{ será solto}\}. \end{aligned}$$

Observe a organização dos elementos de $\Omega_L$ na Figura 5.

Figura 5 – Figura ilustrativa do espaço amostral relativo ao problema do Dilema do Prisioneiro




Esses eventos dão todas as possibilidades de soltura do par de presos levando em conta a segunda afirmação do guarda. O evento $L_1$ é equivalente ao evento no qual $A$ e $B$ são soltos, por isso tem probabilidade $\mathbb{P}(L_1) = \frac{1}{3}$. Analogamente, $L_2$ corresponde ao evento em que $A$ e $C$ são libertados e, por isso, $\mathbb{P}(L_2) = \frac{1}{3}$. Mas veja só: o evento em que $B$ e $C$ são soltos corresponde ao evento $L_3 \cup L_4$, logo $\mathbb{P}(L_3 \cup L_4) = \frac{1}{3}$. Entretanto, sem mais informações, não podemos determinar as probabilidades individuais de $L_3$ e $L_4$. Supondo que ambos sejam equiprováveis, teríamos $\mathbb{P}(L_3) = \mathbb{P}(L_4) = \frac{1}{6}$, pois $\frac{1}{6} + \frac{1}{6} = \frac{1}{3}$. Dessa forma, sendo $L_1 \cap (L_1 \cup L_3) = L_1$ temos que

$$\begin{aligned} \mathbb{P}(\text{"}A \text{ ser solto"}|\text{"o guarda diz que } B \text{ será solto"}) &= \mathbb{P}(L_1 \,|\, (L_1 \cup L_3)) = \\ &= \frac{\mathbb{P}(L_1)}{\mathbb{P}(L_1 \cup L_3)} = \frac{\frac{1}{3}}{\frac{1}{3} + \frac{1}{6}} = \frac{\frac{1}{3}}{\frac{1}{2}} = \frac{2}{3}, \end{aligned}$$

mostrando que a probabilidade condicional da libertação do prisioneiro $A$ é a mesma que a original, sem considerar a fala do guarda. Observe que o termo $\frac{1}{3}+\frac{1}{6}=\frac{1}{2}$ aparece porque o evento "*o guarda diz que B será solto*" é a união dos eventos $L_1$ e $L_3$, que são disjuntos. O mesmo argumento é válido ao calcular $\mathbb{P}($"$A$ ser solto"$|$"o guarda diz que $C$ será solto"$)=\frac{2}{3}$. Assim, o problema está resolvido: o guarda não muda a probabilidade do prisioneiro $A$ ser solto ao fazer sua afirmação, isto é, o evento "*A ser solto*" é independente tanto do evento "*o guarda diz que B será solto*" quanto do evento "*o guarda diz que B será solto*".

Observe que, diferentemente do exemplo anterior, a nova informação dada pelo guarda não altera a chance do prisioneiro A ser libertado. No problema das portas, a informação dada pelo apresentador alterava a chance do participante ganhar o prêmio. Esses dois problemas são exemplos importantes que pode-se usar para compreender que a interpretação de um problema de probabilidade é parte fundamental para sua compreensão e seu entendimento.

Para completar a análise, vamos considerar o caso em que o guarda sempre identifica $B$ em sua afirmação, mesmo quando ambos $B$ e $C$ são soltos.O guarda, agindo dessa forma, *poderia* mudar o valor da probabilidade do prisioneiro $A$ ser libertado apenas alterando a maneira como ele escolhe outro prisioneiro para fazer sua afirmação. Esse ponto nos leva a nos questionar: "o guarda realmente tem controle do destino do prisioneiro $A$, como ele imagina, simplesmente pela maneira como ele determina sua afirmação?" Se isso fosse verdade, então bastaria o guarda dizer sua afirmação para si mesmo ou para outra pessoa ao invés de dizê-lo ao prisioneiro $A$ para alterar a probabilidade dele ser solto. Entretanto, sabemos que a decisão de quais prisioneiros libertar não cabe ao guarda, mas a outra pessoa, independente da afirmação que o guarda escolher fazer. Esse fato nos sugere que devemos começar nossa análise presumindo a independência do evento "$A$ ser solto" e a afirmação do guarda. Ao fazer tal suposição, a probabilidade condicional na fórmula acima será $\frac{2}{3}$ e a única maneira de isso acontecer é se $L_3$ e $L_4$ tiverem, cada um, a probabilidade de $\frac{1}{6}$, como fizemos em nossa primeira solução. Dessa maneira, a primeira solução é, de fato, aquela que representa o que acontece no "mundo real". Outras soluções, como a que acabamos de apresentar, mesmo que matematicamente corretas, não correspondem à situação real envolvendo esse caso. Veja também que se estivermos interessados na probabilidade *não-condicional* $\mathbb{P}($"$A$ ser solto"$)$ ao invés da probabilidade condicional, então a resposta será, simplesmente,

$$\mathbb{P}(L_1)+\mathbb{P}(L_2)=\frac{2}{3},$$

atestando que esse evento independe da maneira que o guarda escolhe para fazer sua afirmação.

Veja que, nesse caso, a nova informação do guarda não altera a probabilidade analisada, diferentemente do problema explorado anteriormente, no qual a informação dada pelo apresentador altera a chance do participante ganhar o prêmio.

## 3.4 O macaco e a máquina digitadora

### *O problema do macaco*

Uma das histórias mais conhecidas sobre probabilidade no meio acadêmico é a que trata do “macaco na máquina digitadora”.

Imagine que um macaco é colocado em frente a uma máquina digitadora e passa a apertar suas teclas de maneira aleatória, sem parar. Esse macaco é um animal com características únicas: ele não se cansa, não dorme e tem “memória fraca”. Devido a isso, ele nunca para de teclar, não se lembra de nenhuma tecla que já teclou nem conhece a ordem ou posição das teclas da máquina. Dessa maneira, vamos supor que cada tecla que ele aperta não depende de nenhuma tecla que ele já tenha apertado nem influencia ou condiciona a tecla que será apertada a seguir. Assim, diremos que *cada teclar do macaco na máquina é independente dos demais*. Como cada tecla produz um caractere, seja este uma letra, um número, um símbolo, ou mesmo um espaço, o teclar do macaco na máquina produz uma sequência infinita de caracteres aleatórios onde cada caractere é independente dos demais. Uma vez que caracterizamos nosso macaco, podemos pensar além e afirmar que “*nessas condições, em algum momento, o macaco certamente terá digitado a obra completa de* Shakespeare*, isto é, o terá feito com probabilidade de* 100%”. Para um leigo, pode parecer estranho e impossível que essa afirmação seja válida. Ainda assim, veremos que essa afirmação é verdadeira do ponto de vista da Teoria da Probabilidade.

Poderíamos aportuguesar essa história mudando o objetivo final da digitação do macaco: ele poderia escrever o texto completo de “Os Lusíadas”, de Luís de Camões, ou a obra completa de Machado de Assis. Se quisermos, poderemos simplesmente afirmar que, em algum momento, o macaco terá escrito, em ordem de nascimento, o nome completo de todos os membros de cinco gerações da família de alguém. Essas alterações não fariam o problema perder sua generalidade. É provável que alguém logo concorde com a conclusão da história e a complete com o seguinte comentário: “Isso pode acontecer em algum momento, mas vai demorar bastante...”. Esta pessoa ainda poderia afirmar que o tempo para que tal acontecimento ocorresse seria tão grande que, provavelmente, o macaco não sobreviveria ou a máquina quebraria, sem contar a reposição de papeis necessária.

Devido a esses questionamentos, é preciso considerar que o macaco e a máquina digitadora em questão são apenas um modelo para representar um dispositivo que produz uma sequência aleatória de letras e símbolos que poderia ser, também, um aplicativo ou um programa de computador. Se essa história se tratasse de um experimento real, seria muito provável que nossas suposições iniciais, que conduzem à independência da escolha das teclas, não fossem satisfeitas. Afinal, um macaco real é um ser vivo racional que, uma hora ou outra, poderia decidir fazer outra coisa. A metáfora do macaco na máquina digitadora corresponde, portanto, a um modelo teórico-probabilístico, e não a um experimento real com um macaco digitando numa máquina de verdade.

Antes de analisar a digitação da obra completa de Shakespeare, vamos calcular a probabilidade do macaco digitar apenas uma palavra: *GALHO*, por exemplo, a partir das cinco primeiras teclas que ele escolher. Para isso, vamos assumir que $\Omega$ é composto dos eventos *"teclar a tecla* i*"*, sendo cada um destes um evento elementar independente dos demais. Também admitiremos que cada evento elementar é equiprovável, ou seja, cada tecla tem a mesma chance de ser escolhida, como já mencionamos anteriormente. Vamos, então, analisar a ocorrência sucessiva dos eventos elementares *"digitar a letra G"*, *"digitar a letra A"*, *"digitar a letra L"*, *"digitar a letra H"* e *"digitar a letra O"* nas cinco primeiras teclas apertadas pelo macaco. A ocorrência dos eventos citados corresponde à ocorrência do evento $G$ : *"digitar a palavra* GALHO*"*. Supondo que há 50 teclas na máquina digitadora, que os eventos elementares tem a mesma probabilidade $\frac{1}{50}$ de ocorrer e que cada um desses eventos elementares são independente dos demais, podemos calcular a probabilidade do evento $G$: *"digitar a palavra* GALHO *em* 5 *movimentos consecutivos"* da seguinte maneira:

$$\mathbb{P}(G) = \frac{1}{50} \cdot \frac{1}{50} \cdot \frac{1}{50} \cdot \frac{1}{50} \cdot \frac{1}{50} = \frac{1}{50^5} = 0,0000000032 = 3,2 \cdot 10^{-9} > 0.$$

Observe que a chance do macaco digitar a palavra *GALHO* nas condições mencionadas, é menos de uma em $3,2$ bilhões. Analogamente, a probabilidade do evento $G$ não ocorrer é dada por

$$\mathbb{P}(G^c) = 1 - \mathbb{P}(G) = 1 - \frac{1}{50^6} \approx 0,9999999968 > 0$$

Veja, então, que a probabilidade do macaco digitar a palavra *GALHO* é muito pequena, mas não é nula (por isso, o evento $G$ não é impossível). Ao contrário, a probabilidade do evento $G$ não acontecer é muito grande. Entretanto, é preciso levar em consideração que a maneira como calculamos a probabilidade de $G$ ocorrer corresponde à ocorrência do evento apenas na primeira sequência de cinco caracteres consecutivos, um pequeno bloco de caracteres. Como supomos inicialmente que o macaco não para de digitar, é preciso considerar que o evento $G$ pode ocorrer num outro bloco de cinco caracteres. Como supomos inicialmente que cada evento elementar é independente em relação aos demais, podemos estender essa independência aos blocos de 5 caracteres que queremos analisar, afinal nenhum desses blocos relaciona-se com os caracteres dos blocos anteriores ou influencia os caracteres dos próximos blocos. Temos, então, um novo $\Omega_5$, no qual cada evento elementar é composto de um bloco de 5 caracteres que foram teclados sucessivamente, de forma que cada bloco é disjunto dos demais (o primeiro bloco é composto do primeiro ao quinto caractere teclado pelo macaco, o segundo bloco é composto do sexto ao décimo caractere teclado e assim sucessivamente). Admitiremos que a probabilidade da palavra *GALHO* ser teclada em qualquer um desses blocos é $\frac{1}{50^5}$, como acabamos de calcular. Vamos chamar de $G_n^c$ o evento "não digitar a palavra *GALHO* em qualquer um dos primeiros $n$ blocos de 5 caracteres". Temos, então, que a probabilidade de $G_n^c$ ocorrer é dada por

$$\mathbb{P}(G_n^c) = \left(1 - \frac{1}{50^5}\right)^n \tag{3.1}$$

Analisando a equação 3.1, vemos que

$$\begin{aligned} n = 100000 &\quad \to \quad \mathbb{P}(G_n^c) \approx 0,99968 \\ n = 10000000 &\quad \to \quad \mathbb{P}(G_n^c) \approx 0,96851 \\ n = 100000000 &\quad \to \quad \mathbb{P}(G_n^c) \approx 0,72645 \\ n = 1000000000 &\quad \to \quad \mathbb{P}(G_n^c) \approx 0,04076 \\ n = 2000000000 &\quad \to \quad \mathbb{P}(G_n^c) \approx 0,00166 \end{aligned}$$

Observe que, à medida que aumentamos o número $n$ de blocos, a probabilidade de que $G_n^c$ ocorra diminui, isto é, a probabilidade da palavra *GALHO* não ser digitada nesses $n$ blocos diminui. É de se pensar, portanto, que se tomarmos $n$ suficientemente grande, a probabilidade de $G_n^c$ ocorrer fica tão pequena quanto desejarmos. De fato,

$$1 - \frac{1}{50^5} = \frac{50^5 - 1}{50^5} < 1.$$

Por isso, quando $n$ cresce, $\left(1 - \frac{1}{50^5}\right)^n$ tende a zero. Daí,

$$\lim_{n\to\infty} \mathbb{P}(G_n^c) = \lim_{n\to\infty} \left(1 - \frac{1}{50^5}\right)^n = 0$$

É possível perceber, através desse raciocínio, que usando um número $n$ suficientemente grande de blocos, a chance da palavra *GALHO* ser digitada é 1. Vamos usar os mesmos argumentos para ampliar esse resultado para toda a obra de Shakespeare.

De modo mais geral, então, vamos supor que há $t$ teclas no teclado da máquina em frente ao macaco. Vamos continuar supondo que os movimentos do macaco não param nem "guardam memória" das teclas ou do teclado, de forma que cada tecla que ele aperta na máquina é uma ação independente das demais e que cada uma dessas $t$ teclas tem probabilidade $\frac{1}{t} = t^{-1}$ de ser apertada pelo macaco, produzindo uma enorme sequência de caracteres aleatórios.

Dessa maneira, a produção da obra de Shakespeare corresponde a uma grande sequência de caracteres que, quando lida, forma todos os livros desse autor, em uma ordem particular (cronológica, por exemplo). Precisamos, então, escolher uma sequência de caracteres consecutivos suficientemente grande produzida pelo macaco na máquina para que nela caibam todos os escritos de Shakespeare. Suponha, então, que essa sequência é composta de um número finito de caracteres $n$, por exemplo. Considere que o macaco digite $n$ caracteres aleatoriamente. Para que esses $n$ caracteres correspondam à obra de Shakespeare na ordem escolhida, o macaco precisa apertar os $n$ caracteres exatamente na sequência esperada. Caso isso ocorra, diremos que o experimento obteve sucesso.

Como cada tecla tem $\frac{1}{t}$ chance de ser escolhida pelo macaco e cada movimento dele é um evento independente dos demais, a chance de que ele digite os $n$ caracteres esperados, em ordem, corresponde a $\left(\frac{1}{t}\right)^n = t^{-n}$.Veja que $t^{-n}$ é uma probabilidade muito pequena, sendo, entretanto, um número positivo.

Como o macaco sempre continua a trabalhar apertando sucessivamente as teclas da máquina, podemos escolher um desses movimentos para iniciar nossa contagem das $n$ primeiras ações. Após a $n$-ésima tecla, poderemos apontar se a obra de Shakespeare foi (ou não) digitada completamente. Consideraremos que a obra completa só será produzida se cada caractere digitado corresponder ao caractere esperado para aquela posição. Se ao menos um caractere for digitado incorretamente, então essa sequência será desconsiderada, mesmo que o macaco já tenha digitado uma das obras corretamente. Diremos, então, que o macaco terá sucesso em sua tarefa após $n$ teclas apertadas se não houver nenhuma falha (caractere errado). A probabilidade disso ocorrer é dada, portanto, por $\left(\frac{1}{t}\right)^n = t^{-n} = p > 0$ e a probabilidade de isso não acontecer, $1 - p$.

O macaco continua a digitar, por isso podemos observar os caracteres digitados da $n+1$ tentativa até a $2n$, sendo esta a segunda sequência de $n$ caracteres digitados que ocorre logo após o primeiro bloco de $n$ tentativas. Esse segundo bloco também consiste de $n$ teclas apertadas, cada uma teclada independentemente das demais, com a mesma probabilidade $\frac{1}{t}$ de ser teclada corretamente em cada tentativa, por isso a chance de produzir a obra completa nesse segundo bloco também é de $\left(\frac{1}{t}\right)^n = t^{-n} = p > 0$. Analogamente, a chance de obter sucesso nos próximos $n$ caracteres digitados (da tecla $2n+1$ até a $3n$ também é $\left(\frac{1}{t}\right)^n = t^{-n} = p > 0$. Dessa maneira, estamos subvidindo a sequência infinita de caracteres em sequências menores disjuntas, compostas de $n$ caracteres, para os quais analisamos o sucesso (ou não) do evento "digitar a obra completa de Shakespeare". Dessa maneira, é fácil ver que a $i$-ésima sequência de caracteres será considerada do caractere $(i-1)n+1$ ao caractere $i-n$. Passamos a considerar um $\Omega_B$ composto de blocos disjuntos de $n$ caracteres teclados pelo macaco sucessivamente, independentes uns dos outros cada um com probabilidade $p = \frac{1}{t^n} > 0$ de apresentar a sequência de caracteres desejada. Dessa forma, podemos definir o evento

$$S_i = \{\text{a obra completa é digitada no } i\text{-ésimo segmento}\}.$$

Também é possível notar que cada evento $S_i$ é independente. De fato, o $i$-ésimo segmento de teclas considera um conjunto de caracteres digitados pelo macaco que não se soprepõe com nenhum $j$-ésimo segmento, $i \neq j$. Como cada movimento do macaco é independente dos demais, é intuitivo deduzir que essas sequências disjuntas de caracteres aleatórios digitados sejam independentes, pois uma das nossas hipóteses é o macaco não reter memória sobre seus movimentos anteriores, de forma que as teclas apertadas em uma sequência de caracteres não influenciam as teclas que serão tecladas em qualquer outra sequência. Uma vez compreendida a independência dos eventos $S_i$, o problema é reconduzido a algo parecido com o que fizemos anteriormente. O sucesso na empreitada de digitar a obra completa de Shakespeare no segmento $i$ de caracteres digitados corresponde ao evento $S_i$, sendo que esse evento tem uma probabilidade positiva $p$. Assim, a probabilidade do *primeiro* sucesso ocorrer na $i-$ésima sequência de caracteres é

$S_i = (1-p)^{i-1} \cdot p$, pois considera-se que as primeiras $i-1$ sequências de caracteres falharam e que cada $S_i$ é independente dos demais.

Seja $S$ o evento no qual o sucesso vai ocorrer em pelo menos um segmento. $S$ é a união disjunta dos eventos $S_i$ definidos até o primeiro sucesso no segmento $i$ e, assim,

$$\mathbb{P}(S) = p + (1-p)p + (1-p)^2 p + \cdots$$

Essa expressão corresponde a uma série geométrica cujo primeiro elemento é $p > 0$ e a razão é $(1-p) < 1$. Essa soma pode ser calculada usando a fórmula da soma dos termos de uma progressão geométrica infinita de razão $(1-p)$ cujo primeiro termo é $p$:

$$\mathbb{P}(S) = \frac{p}{1-(1-p)} = \frac{p}{1-1+p} = \frac{p}{p} = 1.$$

Dessa forma, a probabilidade do macaco acertar a sequência de caracteres que formam a obra completa de Shakespeare em algum $S_i$ é 1, ou seja, é certo que em algum momento ele atingirá nosso objetivo.

Também podemos apresentar a solução de uma maneira diferente. Vamos olhar pelo lado contrário. Se uma falha acontecer em cada um de $m$ segmentos de caracteres, podemos calcular que a probabilidade de $m$ falhas é $(1-p)^m$. À medida que $m$ aumenta, essa probabilidade tende a zero, ou seja, a probabilidade do evento em que há pelo menos um sucesso em um número suficientemente grande de $m$ tentativas aumenta e tende a 1. Isso é equivalente a dizer que $\mathbb{P}(S) = 1$.

## *O problema do macaco e o Lema de Borel Cantelli*

Pode ser difícil acreditar que o macaco irá digitar as obras de Shakespeare com certeza em algum momento. Mesmo assim, é possível ir ainda mais além: pode-se provar um teorema assegurando, com certeza, que o macaco digitará a obra completa de Shakespeare não apenas uma vez, mas *infinitas vezes*. Entretanto, se pensarmos quanto tempo demorará para que toda a obra de Shakespeare seja digitada apenas uma vez, veremos que o macaco provavelmente passará mais tempo digitando do que o Sol sobreviverá como uma estrela (então, teríamos que mover o macaco e sua caixa digitadora para algum outro lugar no universo).

O problema do macado na máquina digitadora exemplifica o Lema 1, de Borel Cantelli. Definindo o evento $S$: *digitar a obra completa de Shakespeare em determinada ordem*, podemos dizer que há outros eventos $S_k$: *ocorre $S$ na sequência $k$ de caracteres*, para $k = 1, 2, 3, \ldots$. Assim, nota-se que

- para cada valor de $k$, a probabilidade $\mathbb{P}(S_k) = p > 0$, logo $\sum_{k=1}^{\infty} \mathbb{P}(S_k) = \infty$; e
- os eventos $S_1, S_2, \ldots$ são independentes entre si, de acordo com nossas suposições iniciais.

Assim, pela segunda parte do Lema 1, de Borel Cantelli, temos que

$$\mathbb{P}\left(\limsup_{k\to\infty} S_k\right) = 1$$

O evento $\limsup_{k\to\infty} S_k$ corresponde à ocorrência de uma infinidade de eventos $S_1$, $S_2$, .... Isso significa que, se o macaco continuar a digitar infinitamente de maneira permanente, em algum momento devemos obter $S$ em algum $S_k$ com probabilidade 1. E mais do que isso: significa, também, que isso ocorrerá tantas vezes quanto quisermos (ou tivermos paciência de esperar).

A história do macaco é uma narrativa simples de compreender e nos ensina uma importante lição: *eventos raros, de fato, ocorrerão*. Um evento raro é aquele que tem uma probabilidade muito pequena de ocorrer. Os mesmos argumentos usados para mostrar a certeza do sucesso obtido pelo macaco ao escrever a obra completa de Shakespeare também servem para mostrar que qualquer evento raro com certeza ocorrerá, eventualmente, se o experimento que o produz for repetido independentemente para sempre. Entretanto, podemos nos perguntar: quanto tempo demorará para que tal evento raro ocorra? É possível mostrar que se há uma probabilidade $a$ de sucesso, então o tempo médio de espera até o primeiro sucesso é cerca de $a^{-1}$ tentativas. Isso nos dá uma maneira de estimar quanto tempo teremos que esperar para que o macaco complete essa tarefa. Para que o macaco digite a palavra "*GALHO*", por exemplo, supondo que seu teclado seja formado por 50 teclas, a probabilidade é de cerca de $a = \frac{1}{50^5}$, o que implica na necessidade de $a^{-1} = 50^5 = 312500000$ tentativas. Considerando que o macaco leve $1s$ para digitar cada tecla, o tempo necessário seria $312500000s = 120$meses 16dias 21h 33min 56s. Não é de se estranhar, portanto, que, para digitar completamente a obra de Shakespeare pela primeira vez, o macaco demore mais tempo do que durará toda duração do Sistema Solar. Apesar disso, existem eventos raros que não são tão extremos como esse exemplo e modelos probabilísticos mais aplicados à vida real.

## 3.5 Algumas variações do problema de Monty Hall

### *1ª variação*

O problema de Monty Hall, que exploramos anteriormente, poderia ser realizado com quatro portas, sendo uma delas premiada. Adaptando o enunciado, poderíamos estabelecer o funcionamento do problema da seguinte maneira:

1. o participante faz a escolha inicial de uma das quatro possíveis portas: $A$, $B$, $C$ ou $D$;

2. o apresentador abre **duas** das três portas não escolhidas pelo participante, revelando duas portas com bodes; e

3. o participante é convidado a trocar a porta que escolhera inicialmente pela outra porta agora disponível.

Com as mesmas suposições que fizemos antes, consideraremos que todas as quatro portas tem a mesma chance de serem escolhidas inicialmente pelo participante. Ao escolher uma dessas portas, o participante tem $\frac{1}{4}$ de probabilidade de acertar a porta que esconde o carro. Após o apresentador revelar duas portas com bodes, o participante enfrenta um novo dilema, pois o prêmio, agora, está atrás da porta que ele escolheu inicialmente ou não. Assim, podemos considerar o evento $I$...: "*o carro está escondido na porta escolhida inicialmente*" e afirmar, com razão, que $\mathbb{P}(I) = \frac{1}{4}$. Seguindo nessa linha de raciocínio, observamos que o evento complementar $I^c$: "*o carro não está escondido na porta escolhida inicialmente*" tem $\mathbb{P}(I^c) = 1 - \mathbb{P}(I) = \frac{3}{4}$ de chance de ocorrer. Mas se $I^c$ ocorre, então, devido ao fato do apresentador ter aberto duas portas, a chance do carro não estar escondido na porta escolhida inicialmente, isto é, estar atrás da outra porta disponível, é $\frac{3}{4}$. Dessa forma, o participante tem três vezes mais chances de ganhar o carro ao trocar de porta do que se ele não o fizer.

Podemos ampliar essa situação ao imaginá-la sendo desenvolvida com 1000 portas. Após o participante escolher uma porta, o apresentador abre outras 998 portas, revelando 998 prêmios ruins. Restam, então, apenas duas portas: a que o participante escolhera inicialmente e uma outra oferecida para troca. No início, com 1000 portas, a chance de ganhar o carro com a escolha inicial era muito pequena, pois a porta com o carro tinha apenas a probabilidade de $\frac{1}{1000}$ de ser escolhida, supondo que cada porta tem a mesma probabilidade de esconder o carro atrás de si. Se permanecer com a porta inicialmente escolhida e ela estar com o prêmio, o participante manterá seus $\frac{1}{1000} = 0,1\%$ de chance de ganhar o prêmio.Mas se ele escolheu a porta que não está com o prêmio, ele terá $\frac{999}{1000} = 99,9\%$ de chance de perder o prêmio se mantiver sua escolha inicial. Ou seja: o participante tem $99,9\%$ de chance de ganhar o prêmio se trocar de porta.

## 2ª variação

A dinâmica do jogo poderia ser alterada fazendo apenas uma pequena mudança em suas regras:

1. o participante faz a escolha inicial de uma das quatro possíveis portas: $A$, $B$, $C$ ou $D$;

2. o apresentador abre **uma** das três portas não escolhidas pelo participante, revelando um bode;

3. o participante é convidado a trocar a porta que escolhera inicialmente por uma das duas outras portas disponíveis;

4. o apresentador abre **uma** das duas portas não escolhidas pelo participante, revelando outro bode;

5. o participante é convidado a trocar a segunda porta que escolhera pela única porta disponível naquele momento;

6. abre-se a porta escolhida pelo participante e, então, revela-se o prêmio escondido nela, sendo este o carro ou não.

Vamos analisar o que acontece quando o participante sempre opta por trocar a porta escolhida quando lhe é dada essa opção. Como fizemos anteriormente, suporemos que o carro encontra-se atrás da porta $A$ enquanto os bodes encontram-se atrás das portas $B$, $C$ e $D$, sem perder a generalidade do problema. Usando a mesma notação anterior, podemos descrever um resultado desse experimento através de uma sêxtupla $(p,q,r,s,t,u)$ onde $p$ indica a porta escolhida inicialmente pelo participante, $q$ indica a primeira porta que o apresentador abriu, $r$ indica a primeira porta para a qual o participante trocou sua escolha inicial, $s$ indica a segunda porta que o apresentador abriu, $t$ indica a escolha final feita pelo participante e $u$ dá o resultado obtido com o experimento: se o participante ganhou o carro $(G)$ ou ganhou um bode $(P)$. Assim, por exemplo, a sêxtupla $(A,B,C,D,A,G)$ indica que o participante escolheu inicialmente a porta $A$, o apresentador abriu a porta $B$ em seguida, revelando um bode; o participante, após isso, trocou sua escolha inicial para a porta $C$, fazendo com que o apresentador abrisse a porta $D$ a seguir e que o participante tenha escolhido a porta $A$ ao fim do experimento, ganhando $(G)$ o carro, de acordo com nossa suposição inicial. Listar todos os elementos do espaço amostral desse experimento não é tão intuitivo, devido ao grande número de situações possíveis. Para auxiliar, listamos as opções possíveis na tabela da Figura 6.

O espaço amostral, portanto, pode ser descrito como o conjunto de sêxtuplas

$$\begin{aligned}\Omega \;=\; & \{(A,B,C,D,A,G),(A,B,D,C,A,G),(A,C,B,D,A,G),(A,C,D,B,A,G),(A,D,B,C,A,G),\\ & (A,D,C,B,A,G),(B,C,A,B,D,P),(B,C,A,D,B,P),(B,C,D,B,A,G),(B,D,A,B,C,P),\\ & (B,D,A,C,B,P),(B,D,C,B,A,G),(C,B,A,C,D,P),(C,B,A,D,C,P),(C,B,D,C,A,G),\\ & (C,D,A,B,C,P),(C,D,A,C,B,P),(C,D,B,C,A,G),(D,B,A,C,D,P),(D,B,A,D,C,P),\\ & (D,B,C,D,A,G),(D,C,A,B,D,P),(D,C,A,D,B,P),(D,C,B,D,A,G)\}.\end{aligned}$$

Como antes, a última letra da sêxtupla é apenas uma convenção que utilizaremos para identificar facilmente o resultado da sequência de movimentos utilizada. Observe, também que se o participante escolher inicialmente a porta $A$, então, pelas regras fixadas, ele obrigatoriamente trocará de porta duas vezes e, dessa forma, ganhará o prêmio, como pode ser visto nos seis primeiros resultados apresentados em $\Omega$. Se o participante escolher inicialmente alguma outra porta, dependendo das suas escolhas de troca, há outras seis chances diferentes de ganhar, como exibido entre os demais resultados apresentados em $\Omega$.

Falta, assim, definir a probabilidade de cada evento elementar de $\Omega$. Como antes, vamos supor que o participante sempre fará suas escolhas de portas "*ao acaso*". Dessa maneira, podemos

Figura 6 – Opções possíveis para a 2ª variação do problema das portas

| Escolhe | Abre | Escolhe | Abre | Escolhe | Sigla |
|---|---|---|---|---|---|
| A | B | C | D | A | ABCDA G |
| | | D | C | A | ABDCA G |
| | C | B | D | A | ACBDA G |
| | | D | B | A | ACDBA G |
| | D | B | C | A | ADBCA G |
| | | C | B | A | ADCBA G |
| B | C | A | B | D | BCABD P |
| | | | D | B | BCADB P |
| | | D | B | A | BCDBA G |
| | D | A | B | C | BDABC P |
| | | | C | B | BDACB P |
| | | C | B | A | BDCBA G |
| C | B | A | C | D | CBACD P |
| | | | D | C | CBADC P |
| | | D | C | A | CBDCA G |
| | D | A | B | C | CDABC P |
| | | | C | B | CDACB P |
| | | B | C | A | CDBCA G |
| D | B | A | C | D | DBACD P |
| | | | D | C | DBADC P |
| | | C | D | A | DBCDA G |
| | C | A | B | D | DCABD P |
| | | | D | B | DCADB P |
| | | B | D | A | DCBDA G |

pensar que cada uma das quatro portas é igualmente provável de ser escolhida inicialmente pelo participante. Considerando isso, portanto, diremos que cada porta tem probabilidade $\frac{1}{4}$ de ser escolhida inicialmente.

Retomemos, agora, a análise das probabilidades individuais em $\Omega$. A porta $A$ pode ser escolhida inicialmente pelo participante com probabilidade de $\frac{1}{4}$. Mas em $\Omega$, ao escolher inicialmente a porta $A$, o participante obrigatoriamente ganhará, podendo fazê-lo de seis maneiras distintas. Assim, podemos afirmar que o evento $A_1$ : "*escolhe-se inicialmente a porta* $A$" dado por

$$\begin{aligned} A_1 &= \{(A,B,C,D,A,G),(A,B,D,C,A,G),(A,C,B,D,A,G), \\ &\qquad (A,C,D,B,A,G),(A,D,B,C,A,G),(A,D,C,B,A,G)\}, \end{aligned}$$

tem probabilidade $\frac{1}{4}$, mesmo que as probabilidades individuais dos seis eventos elementares que o compõem não tenham sido definidas. Veremos que não precisamos definir essas probabilidades individuais para resolver nosso problema. Sejam, então,

$$\mathbb{P}(\{(A,B,C,D,A,G)\}) = u, \quad \mathbb{P}(\{(A,B,D,C,A,G)\}) = v, \quad \mathbb{P}(\{(A,C,B,D,A,G)\}) = w,$$
$$\mathbb{P}(\{(A,C,D,B,A,G)\}) = x, \quad \mathbb{P}(\{(A,D,B,C,A,G)\}) = y, \quad \mathbb{P}(\{(A,D,C,B,A,G)\}) = z,$$

de forma que $u+v+w+x+y+z=\frac{1}{4}$. Teremos

$$\mathbb{P}(A_1)=u+v+w+x+y+z=\frac{1}{4}.$$

Os demais eventos de $\Omega$ onde o participante ganha o carro são

$$(B,C,D,B,A,G),\quad (B,D,C,B,A,G),\quad (C,B,D,C,A,G),$$
$$(C,D,B,C,A,G),\quad (D,B,C,D,A,G),\ \text{ e }\ (D,C,B,D,A,G).$$

Vamos calcular a probabilidade de um deles: $\mathbb{P}(\{(B,C,D,B,A,G)\})$. Para isso, suporemos que as escolhas do participante ou do apresentador são feitas, em cada etapa, ao acaso. Podemos observar que o experimento é composto de etapas que ocorrem sucessivamente, sendo que cada etapa a partir da segunda depende do resultado observado na etapa anterior. Para esse evento, portanto, podemos usar o Teorema do Produto (2.1) da seguinte maneira:

- a probabilidade do participante escolher inicialmente a porta $B$ é de $\frac{1}{4}$;
- após feita a escolha inicial do participante, ao apresentador cabe apenas escolher entre as portas $D$ e $C$ para abrir e revelar o bode e por isso, supondo que sua escolha seja ao acaso, a probabilidade do apresentador abrir a porta $C$ é $\frac{1}{2}$;
- a seguir, o participante pode escolher ao acaso uma de duas portas - $A$ ou $D$, o que leva à probabilidade de $\frac{1}{2}$ de trocar a porta $B$ pela porta $D$;
- resta ao apresentador abrir a porta $B$ com certeza, ou seja, com probabilidade 1; e
- ao participante resta apenas a porta $A$ como opção para trocar, que deve ser feita com probabilidade 1, de acordo com nossas suposições.

Pelo Teorema do Produto (2.1), portanto, a probabilidade do evento $\{(B,C,D,B,A,G)\}$ é, portanto, dada por

$$\mathbb{P}(\{(B,C,D,B,A,G)\})=\frac{1}{4}\cdot\frac{1}{2}\cdot\frac{1}{2}\cdot 1\cdot 1=\frac{1}{16}.$$

Analogamente, pode-se mostrar que

$$\mathbb{P}(\{(B,D,C,B,A,G)\})=\mathbb{P}(\{(C,B,D,C,A,G)=\mathbb{P}(\{(C,D,B,C,A,G)\})=$$
$$=\mathbb{P}(\{(D,B,C,D,A,G)\})=\mathbb{P}(\{(D,C,B,D,A,G)\})=\mathbb{P}(\{(B,C,D,B,A,G)\})=\frac{1}{16}.$$

Usando as mesmas ideias, vamos calcular $\mathbb{P}(\{(B,C,A,B,D,P)\})$:

- a probabilidade do participante escolher inicialmente a porta $B$ é de $\frac{1}{4}$;

- após feita a escolha inicial do participante, ao apresentador cabe apenas escolher uma das portas $D$ e $C$ para abrir e revelar o bode e por isso, supondo que sua escolha seja ao acaso, a probabilidade do apresentador abrir a porta $C$ é $\frac{1}{2}$;

- a seguir, o participante pode escolher ao acaso uma de duas portas - $A$ ou $D$, o que leva à probabilidade de $\frac{1}{2}$ de trocar a porta $B$ pela porta $A$;

- ao apresentador cabe, agora, escolher uma das duas portas disponíveis: $B$ ou $D$ para abrir e, como essa escolha é feita ao acaso, a porta $B$ tem probabilidade $\frac{1}{2}$ de ser aberta pelo apresentador; e

- ao participante resta apenas a porta $D$ como opção para trocar, que deve ser feita com probabilidade 1, de acordo com nossas suposições.

Pelo Teorema do Produto (2.1), portanto, a probabilidade do evento $\{(B,C,A,B,D,P)\}$ é dada por

$$\mathbb{P}(\{(B,C,A,B,D,P)\}) = \frac{1}{4}\cdot\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}\cdot 1 = \frac{1}{32}.$$

Os 11 eventos ainda não citados, nos quais o participante perde o carro, tem todos a mesma probabilidade $\frac{1}{32}$ de ocorrer e sua justificativa é análoga ao caso exibido acima.

Dessa maneira, podemos elencar as seguintes probabilidades:

- Seja $A_B$ o evento em que o participante termina com a porta $A$ escolhida tendo escolhido inicialmente a porta $B$, então $A_B = \{(B,C,D,B,A,G),(B,D,C,B,A,G)\}$ e

$$\begin{aligned}\mathbb{P}(A_B) &= \mathbb{P}(\{(B,C,D,B,A,G)\}\cup\{(B,D,C,B,A,G)\})\\ &= \mathbb{P}\{(B,C,D,B,A,G)\}+\mathbb{P}\{(B,D,C,B,A,G)\}\\ &= \frac{1}{16}+\frac{1}{16}=\frac{1}{8}.\end{aligned}$$

- Analogamente, sendo $A_C$ o evento em que o participante termina com a porta $A$ escolhida tendo escolhido inicialmente a porta $C$, então $A_C = \{(C,B,D,C,A,G),(C,D,B,C,A,G)\}$ e

$$\begin{aligned}\mathbb{P}(A_C) &= \mathbb{P}(\{(C,B,D,C,A,G)\}\cup\{(C,D,B,C,A,G)\})\\ &= \mathbb{P}\{(C,B,D,C,A,G)\}+\mathbb{P}\{(C,D,B,C,A,G)\}\\ &= \frac{1}{16}+\frac{1}{16}=\frac{1}{8}.\end{aligned}$$

- Dessa maneira, sendo $A_D$ o evento em que o participante termina com a porta $A$ escolhida tendo escolhido inicialmente a porta $D$, então $A_D = \{(D,B,C,D,A,G),(D,C,B,D,A,G)\}$

e

$$
\begin{aligned}
\mathbb{P}(A_D) &= \mathbb{P}(\{(D,B,C,D,A,G)\} \cup \{(D,C,B,D,A,G)\}) \\
&= \mathbb{P}\{(D,B,C,D,A,G)\} + \mathbb{P}\{(D,C,B,D,A,G)\} \\
&= \frac{1}{16} + \frac{1}{16} = \frac{1}{8}.
\end{aligned}
$$

Observe que os eventos $A_1$, $A_B$, $A_C$ e $A_D$ são eventos mutuamente exclusivos, pois $A_1 \cap A_B \cap A_C \cap A_D = \emptyset$. Sendo $V$ o evento em que o participante *vence*, ou seja, ganha o carro, pode-se observar que $V = A_1 \cup A_B \cup A_C \cup A_D$. Pelas regras fixadas, como os eventos são disjuntos, temos que

$$
\begin{aligned}
\mathbb{P}(V) &= \mathbb{P}(A_1 \cup A_B \cup A_C \cup A_D) \\
&= \mathbb{P}(A_1) + \mathbb{P}(A_B) + \mathbb{P}(A_C) + \mathbb{P}(A_D) \\
&= \frac{1}{4} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{5}{8}.
\end{aligned}
$$

Considemos que o evento $V^c$: *o participante não ganha o carro* é dado pela união dos 12 eventos elementares onde o participante tem como escolha final uma das portas $B$, $C$ ou $D$. Como todos esses eventos tem probabilidade $\frac{1}{32}$ e são disjuntos, segue que

$$\mathbb{P}(V^c) = 12 \cdot \frac{1}{32} = \frac{3}{8} = 1 - \frac{5}{8} = 1 - \mathbb{P}(V).$$

Isso, portanto, responde a questão inicial: se o participante sempre trocar a porta escolhida com essas regras determinando o jogo, então ele terá mais chance de ganhar o carro do que de ganhar um bode, de acordo com nossas suposições iniciais. De fato, se ele sempre trocar de porta, terá $\frac{5}{8} - \frac{3}{8} = \frac{2}{8} = 25\%$ a mais de chances de ganhar o carro.

Pelo exposto até aqui, pode-se pensar que trocar de porta é vantajoso em qualquer situação, independente da regra do jogo. Veremos a seguir um caso onde essa suposição não é válida.

### 3ª variação

Alteramos mais uma vez a regra do jogo, que funcionará da seguinte maneira:

1. o participante faz a primeira escolha de uma das quatro possíveis portas: $A$, $B$, $C$ ou $D$;

2. o apresentador abre **uma** das três portas não escolhidas pelo participante, revelando um bode;

3. o participante é convidado a trocar a porta que escolhera inicialmente por uma das duas outras portas disponíveis;

4. abre-se a porta escolhida pelo participante e, então, revela-se o prêmio escondido nela, sendo este bom ou ruim.

Vamos, novamente, analisar o que acontece quando o participante sempre opta por trocar a porta escolhida quando lhe é dada essa opção. Mais uma vez, suporemos que o carro encontra-se atrás da porta $A$ enquanto os bodes encontram-se atrás das portas $B$, $C$ e $D$, sem perder a generalidade do problema. Usando as mesmas notações, um resultado desse experimento pode ser descrito através de uma quádrupla $(p,q,r,s)$ onde $p$ indica a porta escolhida inicialmente pelo participante, $q$ indica a porta que o apresentador abriu, $r$ indica a escolha final feita pelo participante e $s$ dá o resultado obtido com o experimento: se o participante ganhou o carro $(G)$ ou ganhou um bode $(P)$. Assim, por exemplo, a quádrupla $(A,B,C,P)$ indica que o participante escolheu inicialmente a porta $A$, o apresentador abriu a porta $B$, revelando um bode; o participante, então, trocou sua escolha inicial pela porta $C$, fazendo com que ele, ao fim do experimento, perdesse $(P)$ o carro, de acordo com nossa suposição inicial. Mais uma vez, para auxiliar a enumeração dos elementos do espaço amostral, lista-se as opções possíveis na tabela da Figura 7.

O espaço amostral, portanto, pode ser descrito como o conjunto de quádruplas

$$\begin{aligned}\Omega = \{ & (A,B,C,P),(A,B,D,P),(A,C,B,P),(A,C,D,P),(A,D,B,P),(A,D,C,P),\\ & (B,C,A,G),(B,C,D,P),(B,D,A,G),(B,D,C,P),\\ & (C,B,A,G),(C,B,D,P),(C,D,A,G),(C,D,B,P),\\ & (D,B,A,G),(D,B,C,P),(D,C,A,G),(D,C,B,P)\}.\end{aligned}$$

Novamente, a última letra da quádrupla é apenas uma convenção que utilizaremos para identificar facilmente o resultado da sequência de movimentos utilizada. Falta, agora, definir uma probabilidade nos eventos de $\Omega$. Vamos supor que o participante continua fazendo a escolha de sua porta "*ao acaso*". Assim, podemos pensar que cada uma das quatro portas é igualmente provável de ser escolhida inicialmente pelo participante. Considerando isso, portanto, diremos que cada porta tem probabilidade $\frac{1}{4}$ de ser escolhida inicialmente.

A porta $A$, portanto, pode ser escolhida inicialmente pelo participante com probabilidade $\frac{1}{4}$. Mas em $\Omega$, ao escolher inicialmente a porta $A$, o participante obrigatoriamente perderá, podendo fazê-lo de seis maneiras distintas. Assim, podemos afirmar que o evento $A_1$ : "*escolhe-se inicialmente a porta A*" dado por

$$A_1 = \{(A,B,C,P),(A,B,D,P),(A,C,B,P),(A,C,D,P),(A,D,B,P),(A,D,C,P)\},$$

tem probabilidade $\mathbb{P}(A_1) = \frac{1}{4}$, mesmo que as probabilidades individuais dos seis eventos elementares que o compõem não tenham sido definidas. Também nessa ocasião, veremos que não precisamos definir essas probabilidades individuais para resolver nosso problema. Sejam, então, $\mathbb{P}(\{(A,B,C,P)\}) = a$, $\mathbb{P}(\{(A,B,D,P)\}) = b$, $\mathbb{P}(\{(A,C,B,P)\}) = c$, $\mathbb{P}(\{(A,C,D,P)\}) = d$,

Figura 7 – Opções possíveis para a 3ª variação do problema das portas

| Escolhe | Abre | Escolhe | Sigla |
|---|---|---|---|
| A | B | C | ABC P |
| | | D | ABD P |
| | C | B | ACB P |
| | | D | ACD P |
| | D | B | ADB P |
| | | C | ADC P |
| B | C | A | BCA G |
| | | D | BCD P |
| | D | A | BDA G |
| | | C | BDC P |
| C | B | A | CBA G |
| | | D | CBD P |
| | D | A | CDA G |
| | | B | CDB P |
| D | B | A | DBA G |
| | | C | DBC P |
| | C | A | DCA G |
| | | B | DCB P |

$\mathbb{P}(\{(A,D,B,P)\}) = e$, $\mathbb{P}(\{(A,D,C,P)\}) = f$ de forma que $a+b+c+d+e = \frac{1}{4}$. Teremos

$$\mathbb{P}(A_1) = a+b+c+d+e = \frac{1}{4}.$$

De volta ao $\Omega$, veja que há outras 6 ocasiões onde o participante não ganha o carro: $(B,C,D,P)$, $(B,D,C,P)$, $(C,B,D,P)$, $(C,D,B,P)$, $(D,B,C,P)$, $(D,C,B,P)$. Vamos calcular a probabilidade de uma delas: $\mathbb{P}(\{(B,C,D,P)\})$. Vamos manter as suposiões anteriores de que as escolhas do participante ou do apresentador são feitas ao acaso. Como o experimento é composto de três etapas que ocorrem sucessivamente, sendo que a segunda e a terceira etapas dependem do resultado observado na etapa anterior, podemos, portanto, usar o Teorema do Produto (2.1) da seguinte maneira:

- a probabilidade do participante escolher inicialmente a porta $B$ é de $\frac{1}{4}$;
- após feita a escolha inicial do participante, ao apresentador cabe apenas escolher uma entre as portas $D$ e $C$ para abrir e revelar o bode e por isso, supondo que sua escolha seja ao acaso, a probabilidade do apresentador abrir a porta $C$ após a primeira escolha do participante é $\frac{1}{2}$;
- o participante, então, escolhe uma das duas portas restantes disponíveis, por isso a probabilidade de escolher a porta $D$ é $\dfrac{1}{2}$.

Pelo Teorema do Produto (2.1), portanto, a probabilidade do evento $\{(B,C,D,P)\}$ é dada por

$$\mathbb{P}(\{(B,C,D,P)\}) = \frac{1}{4}\cdot\frac{1}{2}\cdot\frac{1}{2} = \frac{1}{16}.$$

Analogamente, pode-se mostrar que

$$\begin{aligned}\mathbb{P}(\{(B,D,C,P)\}) = \mathbb{P}(\{(C,B,D,P) = \mathbb{P}(\{(C,D,B,P)\}) = \mathbb{P}(\{(D,B,C,P)\}) = \\ = \mathbb{P}(\{(D,C,B,P)\}) = \mathbb{P}(\{(B,C,D,P)\}) = \frac{1}{16}.\end{aligned}$$

Usando as mesmas ideias, vamos calcular $\mathbb{P}(\{(B,C,A,G)\})$:

- a probabilidade do participante escolher inicialmente a porta $B$ é de $\frac{1}{4}$;
- após feita a escolha inicial do participante, ao apresentador cabe apenas escolher uma entre as portas $D$ e $C$ para abrir e revelar o bode e por isso, supondo que sua escolha seja ao acaso, a probabilidade do apresentador abrir a porta $C$ após a primeira escolha do participante é $\frac{1}{2}$;
- o participante, então, escolhe uma das duas portas restantes disponíveis, por isso a probabilidade de escolher a porta $A$ ao final do experimento é $\frac{1}{2}$.

Pelo Teorema do Produto (2.1), portanto, a probabilidade do evento $\{(B,C,A,G)\}$ é dada por

$$\mathbb{P}(\{(B,C,A,G)\}) = \frac{1}{4}\cdot\frac{1}{2}\cdot\frac{1}{2} = \frac{1}{16}.$$

Os demais eventos onde o participante ganha o carro são tais que

$$\begin{aligned}\mathbb{P}(\{(B,D,A,G)\}) = \mathbb{P}(\{(C,B,A,G)\}) = \mathbb{P}(\{(C,D,A,G)\}) = \mathbb{P}(\{(D,B,A,G)\}) = \\ = \mathbb{P}(\{(D,C,A,G)\}) = \mathbb{P}(\{(B,C,A,G)\}) = \frac{1}{16}.\end{aligned}$$

e suas justificativas são análogas ao caso exibido acima.

Dessa maneira, podemos elencar as seguintes probabilidades:

- Seja $A_B$ o evento em que o participante termina com a porta $A$ escolhida tendo escolhido inicialmente a porta $B$: $A_B = \{(B,C,A,G),(B,D,A,G)\}$ e

$$\begin{aligned}\mathbb{P}(A_B) &= \mathbb{P}(\{(B,C,A,G)\}\cup\{(B,D,A,G)\}) \\ &= \mathbb{P}\{(B,C,A,G)\} + \mathbb{P}\{(B,D,A,G)\} \\ &= \frac{1}{16}+\frac{1}{16} = \frac{1}{8}.\end{aligned}$$

- Analogamente, sendo $A_C$ o evento em que o participante termina com a porta $A$ escolhida tendo escolhido inicialmente a porta $C$, então $A_C = \{(C,B,A,G),(C,D,A,G)\}$ e

$$\begin{aligned}\mathbb{P}(A_C) &= \mathbb{P}(\{(C,B,A,G)\} \cup \{(C,D,A,G)\}) \\ &= \mathbb{P}\{(C,B,A,G)\} + \mathbb{P}\{(C,D,A,G)\} \\ &= \frac{1}{16} + \frac{1}{16} = \frac{1}{8}.\end{aligned}$$

- Por último, sendo $A_D$ o evento em que o participante termina com a porta $A$ escolhida tendo escolhido inicialmente a porta $D$, então $A_D = \{(D,B,A,G),(D,C,A,G)\}$ e

$$\begin{aligned}\mathbb{P}(A_D) &= \mathbb{P}(\{(D,B,A,G)\} \cup \{(D,C,A,G)\}) \\ &= \mathbb{P}\{(D,B,A,G)\} + \mathbb{P}\{(D,C,A,G)\} \\ &= \frac{1}{16} + \frac{1}{16} = \frac{1}{8}.\end{aligned}$$

Observe que os eventos $A_B$, $A_C$ e $A_D$ são eventos mutuamente exclusivos, pois $A_B \cap A_C \cap A_D = \emptyset$. Sendo $V$ o evento em que o participante *vence*, ou seja, ganha o carro, pode-se observar que $V = A_B \cup A_C \cup A_D$. Pelas regras fixadas, como os eventos são disjuntos, temos que

$$\begin{aligned}\mathbb{P}(V) &= \mathbb{P}(A_B \cup A_C \cup A_D) \\ &= \mathbb{P}(A_B) + \mathbb{P}(A_C) + \mathbb{P}(A_D) \\ &= \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{3}{8}.\end{aligned}$$

Vamos considerar o evento $V^c$: *o participante não ganha o carro*, que é dado pela união do evento $A_1$ com os 6 eventos elementares onde o participante tem como escolha final uma das portas $B$, $C$ ou $D$. Como todos esses 6 eventos tem probabilidade $\frac{1}{16}$ e são disjuntos, segue que

$$\mathbb{P}(V^c) = \mathbb{P}(A_1) + 6 \cdot \frac{1}{16} = \frac{1}{4} + \frac{3}{8} = \frac{5}{8} = 1 - \frac{3}{8} = 1 - \mathbb{P}(V)$$

Veja, então, que chegamos à resposta do questionamento inicial. Nessa variação do problema original, se o participante sempre trocar de porta, ele terá menos chances de ganhar o carro do que de ganhar um bode. É possível chegar à conclusão geral de que pode não ser vantajoso ao candidato sempre optar por trocar a porta escolhida: tudo depende da dinâmica e das regras do jogo.

CAPÍTULO

4

# ORIENTAÇÕES DIDÁTICAS

Nesse capítulo, faremos algumas sugestões de abordagens didáticas para a utilização em sala de aula dos exemplos que tratamos neste trabalho. Tais sugestões podem ser detalhadas pelos professores de Matemática que as utilizarem e adaptadas para o cotidiano de acordo com a percepção da realidade de suas salas de aula.

## 4.1 Usando o problema de Monty Hall em sala de aula

### *Objetivos*

- Objetivos Gerais
  - desenvolver conceitos básicos em probabilidade;
  - motivar o estudo de probabilidade.
- Objetivos específicos
  - compreender a necessidade de entender um problema do ponto de vista matemático, a partir da elucidação de suas situações envolvidas e do uso de vocabulário matemático específico;
  - compreender os conceitos de probabilidade, espaço amostral, evento e probabilidade condicional.

### *Público-alvo*

- Alunos do Ensino Médio

## Materiais e Recursos Necessários

- 3 caixas de papelão decoradas outros objetos iguais que permitam esconder algo dentro (copos plásticos, por exemplo);
- 1 objeto que sirva como prêmio bom (um bombom, um carrinho de brinquedo, um troféu ou uma medalha, por exemplo);
- 2 objetos que sirvam como prêmio ruim (um rolo de papel higiênico, um caderno usado, uma camisa furada ou uma caneta que não escreve mais, por exemplo);
- Envelopes coloridos e papéis em branco;
- Material para anotações (caderno, lápis, borracha e caneta, por exemplo);
- Lousa e giz colorido ou quadro branco e canetas para quadro branco;
- Projetor multimídia conectado a um computador equipado com software de edição de planilhas dinâmicas e acesso à Internet, se o professor assim o desejar.

## Duração

- Três aulas.

## Desenvolvimento das aulas

*Aula 1*

Inicialmente, o professor deve preparar as caixas, escondendo os prêmios de forma que saiba onde está o prêmio bom, sem que os alunos percebam. No decorrer dessa atividade, o professor será o *apresentador do jogo*. Sem deixar que a Matemática apareça como coadjuvante, o professor pode aproveitar esse início de atividade para cativar a atenção dos alunos simulando um programa de auditório: usando jargões de apresentadores de TV, pedindo palmas, executando trilhas sonoras de suspense. Veja a Figura 8 para ter ideia de uma aplicação dessa atividade feita pelo autor dessa pesquisa, por exemplo. Acreditamos que o empenho e a dedicação do professor nessa etapa será fundamental para manter os alunos conectados ao desenvolver da atividade.

Para de iniciar o jogo, o professor convida um aluno para fazer o papel do *participante*. O professor, então, anuncia que uma das caixas esconde um prêmio bom enquanto as demais escondem prêmios ruins. Depois disso, o professor solicita que o aluno participante escolha uma das três caixas. Neste momento, o professor pode dialogar com a turma e pedir a opinião dos demais alunos. Após o participante efetuar a escolha da caixa inicial, o docente pergunta à turma qual a chance do participante ter escolhido a caixa que contém o bom prêmio e escreve essa probabilidade na lousa, acompanhado dos dizeres “Probabilidade inicial:”. Espera-se que

Figura 8 – Exemplo de aula explorando o Problema de Monty Hall


Fonte: Acervo pessoal do autor dessa pesquisa.

os alunos deem respostas intuitivas, como *uma chance em três*. Se isso não ocorrer, o professor deve encaminhar um raciocínio para que os alunos tirem essa conclusão. É importante recordar essa ideia a cada momento do jogo, bem como reforçar que a escolha da porta não depende de ninguém senão do aluno participante.

A seguir, o professor abre uma das caixas que não tem o prêmio bom e dá ao participante a chance de trocar a caixa que escolhera pela outra caixa agora disponível, incentivando os demais alunos a darem sugestões. Feita a escolha, pergunta-se à classe de alunos qual a chance da nova caixa escolhida esconder o prêmio bom. O professor deve apenas escutar as respostas apresentadas e não dar a resposta correta. Pode-se problematizar e propor uma rápida discussão entre os alunos sobre esse ponto da atividade. A seguir, revela-se o prêmio do aluno e anota-se o resultado em uma tabela com as colunas: “*Trocou?*” e “*Ganhou?*”, às quais se responde com *sim* ou *não*, como a Tabela 2 a seguir. Se julgar necessário, o professor pode repetir mais uma vez o procedimento do jogo.

Dando sequência à aula, o professor deve questionar os alunos sobre as probabilidades que foram ditas durante a execução do jogo, perguntando a eles se acreditam ser melhor para o jogador trocar de porta ou se isso não faz diferença. O professor pode promover um novo debate entre os alunos, separando-os em grupos (de 4 ou 5 alunos) e dando um tempo determinado (cerca de 5 minutos) para que cada grupo discuta a possibilidade do jogador trocar de porta ou não, pedindo que um representante de cada grupo compartilhe a conclusão do grupo com a turma toda.

Antes de dar a resposta ao problema, o professor deve solicitar que cada grupo repita o procedimento com seus membros, fazendo um rodízio entre apresentador e participante, e

Tabela 2 – Tabela de resultados

| Trocou a porta? | Ganhou o carro? |
|---|---|
| | |
| | |
| | |
| | |
| | |
| | |

anotem os resultados numa tabela. Se o tempo de aula não for suficiente, o professor pode pedir aos alunos que façam ou terminem a simulação em casa, anotando os resultados em uma tabela como a Tabela 2 numa folha à parte. É fundamental que o professor não relacione a execução da atividade com o problema original nesse momento, pois os alunos podem, através de alguma referência, pesquisar na Internet e descobrir a resposta da situação antes da finalização da atividade.

*Aulas 2 e 3*

A segunda aula é iniciada a partir da retomada das experiências da aula anterior. O professor deve recolher as resultados dos experimentos feitos pelos alunos e anotar os dados na lousa ou em uma planilha eletrônica exposta com auxílio do projetor multimídia.

Sempre questionando a sala, o professor deve oferecer condições para que os alunos tirem suas próprias conclusões quanto à porcentagem de sucesso dos alunos que trocaram a porta e à porcentagem de sucesso dos alunos que não o fizeram.

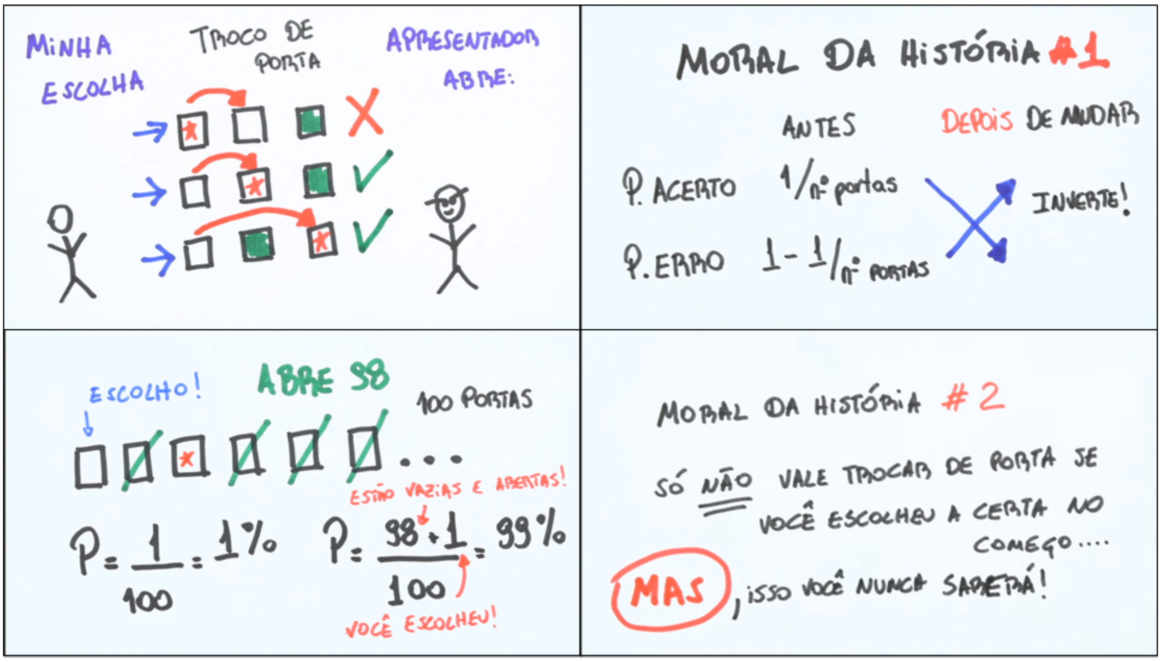
É de se esperar que o resultado se aproxime do resultado real. Caso contrário, o professor pode propor um novo desafio e coletar novos dados.

O professor deve retomar os resultados obtidos na aula anterior, relembrando o que foi feito e solicitando os resultados dos experimentos feitos pelos grupos em casa. A seguir, ele os resume na lousa (ou em uma planilha eletrônica preparada previamente) em uma tabela, o número de pessoas que trocaram de porta e ganharam o prêmio bom, que trocaram de porta e não ganharam prêmio bom, que não trocaram de porta e ganharam o prêmio bom e o número de pessoas que não trocaram de porta mas ganharam o prêmio bom.

Espera-se que os resultados corroborem a informação que conhecemos sobre o problema inicial: que a probabilidade de ganhar o prêmio bom é $\frac{2}{3}$ ao trocar de porta. Caso contrário, o professor deve ter uma lista pronta para apresentar para os alunos e organizar as conclusões.

Se os alunos demonstrarem surpresa com resultados, o professor deve fazer proveito dessa situação para mantê-los motivados e atentos à sua explicação. Caso os alunos não esbocem

Figura 9 – Algumas imagens do vídeo “*Me Salva! O famoso Problema de Monty Hall!*”




Fonte: Montagem feita pelo autor através do vídeo disponível em <https://www.youtube.com/watch?v=Hh7pDPnKK-4>, acessado em 30 dez. 2015.

reação de surpresa (alguém pode ter pesquisado a resposta na internet), peça que algum desses alunos explique para a sala porque a probabilidade que é intuitiva a todos não confere com a probabilidade calculada no problema.

O professor, então, pode explicar a importância de listar os casos possíveis e atribuir-lhes a correta probabilidade para se chegar a resposta da pergunta, definindo o que é espaço amostral e probabilidade condicional a partir das situações do problema. Usando o mesmo raciocínio com o qual explicamos na Seção 3.2 do Capítulo 3, o professor pode utilizar uma árvore de possibilidades para mostrar aos alunos uma forma eficiente de listar os elementos do espaço amostral.

Antes de terminar essa aula, sugerimos que o professor repita todo o procedimento usando, dessa vez, quatro caixas e pedindo que os alunos anotem todos os elementos do espaço amostral e atribuam a probabilidade correta a cada um deles. Após isso, o professor pode exibir o vídeo *Me Salva! O famoso Problema de Monty Hall!*, disponível em <https://www.youtube.com/watch?v=Hh7pDPnKK-4>. Trata-se de um vídeo curto (apenas $2min40seg$) onde se expõe o problema e a explicação, ampliando o resultado para 100 portas. Algumas imagens do vídeo em questão estão apresentadas na Figura 9.

Na terceira aula, o professor pode pedir a participação dos alunos da turma para resolver o problema com quatro caixas, usando ideias de probabilidade condicional para, enfim, propor exercícios para fixação do novo conteúdo. Sugerimos que o professor utilize problemas simples e vá transformando-os em problemas mais elaborados.

As outras versões do problema de Monty Hall que tratamos na Seção 3.5 podem ser desenvolvidas como um projeto ou atividade em sala. O professor também pode utilizar alguns vídeos, que pode encontrar em sites de vídeo na internet para concluir esse tema. Após o cálculo das probabilidades, o professor pode utilizar as tabelas montadas durante a atividade para citar a definição frequentista de probabilidade.

### *Avaliação*

Como avaliação dos conhecimentos desenvolvidos nesta aula, propõe-se que o professor solicite uma pesquisa mais aprofundada sobre o Paradoxo de Monty Hall, envolvendo a história do problema, além de solicitar exercícios e problemas para fixação do conteúdo pelos alunos.

## 4.2 A irmã do rei e a análise de $\Omega$

### *Objetivos*

- Objetivo Geral
  - compreender a importância de identificar corretamente os eventos do espaço amostral para calcular a probabilidade de um evento.
- Objetivos específicos
  - construir uma estratégia de resolução de problemas envolvendo probabilidade a partir da identificação dos pontos do espaço amostral;
  - compreender a importância de se resolver problemas de probabilidade pela definição ao invés de fórmulas pré-definidas;
  - aprimorar a compreensão do conceito de probabilidade condicional.

### *Público-alvo*

- Alunos do Ensino Médio.

### *Materiais e Recursos Necessários*

- Material para anotações (caderno, lápis, borracha e caneta, por exemplo);
- Lousa e giz colorido ou quadro branco e canetas para quadro branco;
- Projetor multimídia conectado a um computador equipado com software de edição de planilhas dinâmicas ou de apresentações multimídia, se o professor assim o desejar.

## Duração

- Uma aula.

## Desenvolvimento da aula

Ao iniciar a aula, o professor deve pedir que os alunos se reúnam em grupos de 4 ou 5 alunos. Em seguida, o professor enuncia o problema da irmã do rei (Seção 3.1). Ao final da exposição do problema, que pode ser oral, em uma apresentação multimídia ou distribuindo o enunciado do problema impresso, o professor deve indicar aos alunos que o resultado pode não ser o que se espera de antemão. Os alunos devem, então, ter um tempo de 7 a 10 minutos para que discutam entre si. Ao final, um representante de cada grupo compartilhará com a sala a conclusão obtida.

O professor deve atuar, durante a execução dessa parte da aula, como um mediador: oferecendo caminhos, sugestões e estratégias para atingir a resposta correta, nunca oferecendo a resposta de prontidão nem conferindo se a resposta está certa ou errada. A conclusão final se dará após o debate da sala (10 min), quando, enfim, o professor encaminhará a resposta correta. O professor deve enfatizar a importância de identificar os elementos do espaço amostral durante sua explicação e pode aproveitar para trabalhar com os alunos a elaboração de árvores de possibilidades para facilitar a listagem dos elementos do espaço amostral.

A seguir, o professor pode solicitar que os alunos resolvam outros problemas nos quais é necessário indicar os elementos do espaço amostral:

- *Qual a chance de obtermos a soma* 8 *a partir do experimento de lançar duas vezes o mesmo dado?*

- *Num escritório, trabalham* 8 *homens e* 9 *mulheres. Sabe-se que há* 5 *pessoas que trabalham nesse escritório que têm mais de* 40 *anos de idade enquanto as demais tem menos do que essa idade. Se há* 7 *mulheres com menos de* 40 *anos dentre os trabalhadores desse escritório, qual a chance de sortear alguém entre eles que seja do sexo masculino e tenha menos de* 40 *anos?*

Usando a mesma dinâmica, o professor pode desenvolver o Dilema do Prisioneiro, exposto nesse trabalho na Seção 3.3. Esse outro problema é interessante para discutir situações que teoricamente podem ocorrer mas não ocorrem na prática. Essa discussão será interessante para introduzir o assunto da Seção 4.3.

## Avaliação

Como avaliação dos conhecimentos desenvolvidos nesta aula, propõe-se que o professor trabalhe uma lista de exercícios e problemas para favorecer o desenvolvimento e fixação dos conteúdos pelos alunos.

## 4.3 Os macacos digitadores e outras ideias sobre probabilidade

### Objetivos

- Objetivos Gerais
  - levar os alunos a perceber que o estudo da probabilidade vai além do que se ensina e se aprende na escola;
  - motivar o estudo de probabilidade através do problema do macaco na máquina digitadora.
- Objetivos específicos
  - desenvolver o pensamento abstrato e crítico e o uso da imaginação para resolver um problema;
  - conhecer o Lema de Borel Cantelli.

### Público-alvo

- Alunos do Ensino Médio.

### Materiais e Recursos Necessários

- Uma máquina de digitar ou um teclado de computador e um macaco de pelúcia para compor o "cenário" da aula;
- Material para anotações (caderno, lápis, borracha e caneta, por exemplo);
- Lousa e giz colorido ou quadro branco e canetas para quadro branco;
- Projetor multimídia conectado a um computador equipado com software de edição de apresentações e conexão à internet, se o professor assim o desejar.

### Duração

- Uma aula.

## Desenvolvimento da aula

Usando o teclado e o macaco de pelúcia, o professor introduz a história do macaco na máquina digitadora, como exposto na Seção 3.4, sem mencionar a obra de Shakespeare. Ao fazer isso, o professor deve usar técnicas que possibilitem o envolvimento dos alunos na contagem da história: use uma linguagem que não seja difícil de ser entendida, use expressões faciais e sonoras para manter a atenção dos alunos.

Organize os alunos em duplas e peça que eles, com o uso de uma calculadora, calculem a chance do macaco digitar, em movimentos consecutivos, o primeiro nome de cada um. Se preferir, use sua criatividade e peça que os alunos calculem a probabilidade de digitar outras palavras, usando o processo que descrevemos para escrever a palavra "GALHO" (p. 67 e p. 67).

Após isso, o professor pode comentar sobre o tempo necessário para que o macaco digite uma palavra, como fizemos na p. 71, e pedir que cada dupla calcule o tempo estimado que o macaco levaria para digitar o nome de um dos integrantes da dupla, convertendo para uma unidade de medida adequada. O professor deve solicitar, então, que os alunos compartilhem entre si os resultados obtidos.

Após isso, o professor pode levantar um debate sobre o que os alunos acharam do tempo e da história, incentivando-os a se expressarem perguntando se acreditam que aquilo é possível ou não e pedindo justificativas. Nesse momento, o professor deve falar sobre a ocorrência de eventos raros e citar o fim do problema do macaco, relativo à digitação da obra completa de Shakespeare.

Ao final, o professor deve apresentar a explicação matemática do problema do macaco, usando o Lema de Borel Cantelli (Lema 1) ou a ideia da soma dos termos de uma progressão geométrica, como exposto em nossa explicação. Para terminar a aula, o professor pode exibir o vídeo *Probabilidade de morrer (Law of large numbers)*, disponível em <https://www.youtube.com/watch?v=p0AjmFAvkZc>. Trata-se de um vídeo de 7*min* onde se trata de eventos raros e sua ocorrência, citando o problema do macaco. Algumas imagens do vídeo em questão estão apresentadas na Figura 10.

Figura 10 – Algumas imagens do vídeo "*Probabilidade de morrer (Law of large numbers)*"


Fonte: Montagem feita pelo autor através do vídeo disponível em <https://www.youtube.com/watch?v=p0AjmFAvkZc>, acessado em 30 dez. 2015.

## Avaliação

Como avaliação dos conhecimentos desenvolvidos nesta aula, propõe-se que o professor solicite uma pesquisa mais aprofundada sobre o Problema do Macaco na Máquina Digitadora. Uma lista de exercícios ou uma simples avaliação somativa também podem ser utilizados para como instrumentos de avaliação.

CAPÍTULO

5

# CONSIDERAÇÕES FINAIS

Neste trabalho, vimos que a Probabilidade pode promover a compreensão de grande parte dos acontecimentos do cotidiano que são de natureza aleatória. O acaso e a incerteza, que se manifestam intuitivamente em nós, muitas vezes são deixados de lado ao se ensinar Probabilidade no Ensino Básico, tornando a abordagem mecânica e repetitiva.

Os problemas analisados aqui tratam exatamente de expor situações onde o acaso gera resultados inesperados, que podem ser explicados através da nossa compreensão da realidade da situação. Em todos os casos citados, está implícita a necessidade de desenvolver a habilidade de enumerar todos os resultados possíveis de um experimento aleatório e entender como eles se relacionam em cada contexto. As situações exploradas favorecem o desenvolvimento das habilidades de leitura, compreensão e abstração, muitas vezes aliadas à intuição, necessárias na resolução problemas. O desenvolvimento de tais habilidades é fundamental para contribuir com o processo de tomada de decisões, o que colabora enormemente para que as competências e habilidades relativas à Probabilidade apontadas em Brasil (2009) sejam desenvolvidas.

Quanto à parte teórica, foram expostos diversos exemplos de situações-problema a fim de abordar experimentos aleatórios e suas características, tipos de eventos, experimentos com espaços amostrais não equiprováveis, a noção de acaso e outros conceitos. Através da leitura atenta do Capítulo 2 e da utilização das sugestões apresentadas no Capítulo 4, que podem ser trabalhadas como atividades complementares, o professor do Ensino Médio pode obter subsídios para desenvolver melhor suas aulas de Probabilidade, evitando assim as dificuldades relacionadas ao ensino-aprendizagem apontadas por Silva (2012), como listamos na p. 28.

Pode-se notar, também, que as sugestões de orientações didáticas que trouxemos no Capítulo 4 são bastante gerais e flexiveis, para que, ao serem implementadas em sala de aula, possam tornar-se um processo desafiador e criativo para alunos e professores, podendo ser melhoradas e consolidadas por eles. Uma perspectiva de atuação em uma pesquisa futura é acompanhar esse processo de melhoramento e consolidação, mensurando-o e avaliando-o. Cabe a este autor começar o processo através de sua própria prática docente e divulgá-lo entre seus colegas e à comunidade de Educação Matemática, seus principais públicos-alvo.

# REFERÊNCIAS

BERNSTEIN, P. L. ***Against the Gods***. New York: John Wiley & Sons, 1996. Disponível em: <https://matrixtrainings.files.wordpress.com/2014/09/against-the-gods-the-remarkable-story-of-risk-1996-peter-l-bernstein.pdf>. Acesso em: 29 dez. 2015. Citado na página 57.

BRASIL. Ministério da Educação e dos Desportos. **Parâmetros Curriculares Nacionais: Matemática**. Brasília, DF, 1997. Disponível em: <http://portal.mec.gov.br/seb/arquivos/pdf/livro03.pdf>. Acesso em: 30 dez. 2015. Citado na página 26.

____. **Parâmetros curriculares nacionais: Ensino Médio. Parte III - Ciências da Natureza, Matemática e suas Tecnologias**. Brasília, 2000. 58 p. Disponível em: <http://portal.mec.gov.br/seb/arquivos/pdf/ciencian.pdf>. Acesso em: 30 dez. 2015. Citado na página 26.

BRASIL. Ministério da Educação e dos Desportos. Secretaria de Educação Fundamental. **Parâmetros curriculares nacionais: terceiro e quarto ciclos do ensino fundamental - introdução aos parâmetros curriculares nacionais**. [S.l.], 1998. 174 p. Disponível em: <http://portal.mec.gov.br/seb/arquivos/pdf/introducao.pdf>. Acesso em: 30 dez. 2015. Citado na página 27.

BRASIL. Ministério da Educação e Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Matriz de Referência para o ENEM 2009**. Brasília, DF, 2009. 26 p. Disponível em: <http://ensinomediodigital.fgv.br/resources/pdf/matriz_novoenem.pdf>. Acesso em: 30 dez. 2015. Citado 2 vezes nas páginas 27 e 93.

CARVALHO, D. L.; OLIVEIRA, P. C. Quatro concepções de probabilidade manifestadas por alunos ingressantes na licenciatura em matemática: clássica, frequentista, subjetiva e formal. In: REUNIÃO ANUAL DA ANPED, 25., 2002, Caxambu. **Anais...** Caxambu: Associação Nacional de Pós-Graduação e Pesquisa em Educação, 1999. Disponível em: <http://25reuniao.anped.org.br/excedentes25/dionelucchesicarvalhot19.rtf>. Acesso em: 30 dez. 2015. Citado na página 28.

DANTAS, C. A. B. **Probabilidade: Um Curso Introdutório**. 2. ed. São Paulo, SP: Edusp, 2004. Citado na página 55.

DANTAS, E. A. **Probabilidade: Uma reflexão teórico-prática no ensino da Matemática**. Dissertação (Mestrado) — Universidade Federal de Campina Grande, Centro de Ciências e Tecnologia, Campina Grande, 2013. Disponível em: <http://www.mat.ufcg.edu.br/PROFmat/TCC/Emanuel.pdf>. Acesso em: 30 dez. 2015. Citado na página 26.

GAFFURI, S. L. **Ensino e Aprendizagem de Probabilidade Através da Metodologia de Resolução de Problemas**. Dissertação (Mestrado) — Centro Universitário Franciscano de Santa Maria, Santa Maria, 2012. Disponível em: <http://tede.unifra.br/tde_busca/arquivo.php?codArquivo=207>. Acesso em: 30 dez. 2015. Citado na página 28.

GONÇALVES, M. C. **Concepção de Professores e o Ensino de Probabilidade na Escola Básica**. Dissertação (Mestrado) — Pontifícia Universidade Católica de São Paulo, Centro de

Ciências e Tecnologia, São Paulo, 2004. Disponível em: <http://www.sapientia.pucsp.br/tde_busca/arquivo.php?codArquivo=4525>. Acesso em: 30 dez. 2015. Citado na página 26.

ISAAC, R. ***The Pleasures of Probability***. New York: Springer-Verlag, 1995. Citado 2 vezes nas páginas 57 e 59.

LOPES, C. A. E. **O conhecimento profissional dos professores e suas relações com estatística e probabilidade na educação infantil.** 281 p. Tese (Doutorado) — Faculdade de Educação, Universidade Estadual de Campinas, Campinas, 2003. Disponível em: <http://www.pucrs.br/famat/viali/tic_literatura/teses/Lopes_CAE.pdf>. Acesso em: 30 dez. 2015. Citado na página 28.

MORGADO, A. C. d. O.; CARVALHO, J. B. P. d.; CARVALHO, P. C. P.; FERNANDEZ, P. **Análise Combinatória e Probabilidade**. 9. ed. Rio de Janeiro: IMPA/VITAE, 1991. Citado na página 25.

NOGUEIRA, A.; BRISOLA, M. A probabilidade no ensino médio: a importância dos jogos como ferramenta didática. In: ACTA LATINOAMERICANA DE MATEMáTICA EDUCATIVA, 26., 2013, Cidade do México. **Anais...** Cidade do México: Comité Latinoamericano de Matemática Educativa A. C., 2013. p. 1929–1937. Disponível em: <http://funes.uniandes.edu.co/4566/1/NogueiraAtividadesALME2013.pdf>. Acesso em: 30 dez. 2015. Citado na página 25.

SALDANHA, N. C. Como perder amigos e enganar pessoas. **Eureka!**, Sociedade Brasileira de Matemática (Olimpíada Brasileira de Matemática), n. 1, p. 41–50, 1998. Citado na página 57.

SALOMÃO, M. S. **Estudo e Generalizações do Paradoxo de Monty Hall na Educação Básica**. Dissertação (Mestrado) — Universidade Federal de Juiz de Fora, Instituto de Ciências Exatas, Juiz de Fora, 2014. Disponível em: <https://sca.profmat-sbm.org.br/tcc_get.php?cpf=03298792780&d=20160103115432&h=2b7088639d29e3e0cab010d0384502d69d970556>. Acesso em: 30 dez. 2015. Citado na página 59.

SÃO PAULO. Secretaria da Educação. **Currículo do Estado de São Paulo: Matemática e suas tecnologias**. São Paulo, 2011. 72 p. Coordenação Maria Inês Fini. Disponível em: <http://www.educacao.sp.gov.br/a2sitebox/arquivos/documentos/238.pdf>. Acesso em: 30 dez. 2015. Citado na página 27.

SILVA, I. d. A. **Probabilidades: a visão laplaciana e a visão frequentista na introdução do conceito**. Dissertação (Mestrado) — Pontifícia Universidade Católica de São Paulo, São Paulo, 2012. Disponível em: <http://www.sapientia.pucsp.br/tde_arquivos/3/TDE-2007-06-29T08:50:26Z-3675/Publico/dissertacao_ismael_araujo_silva.pdf>. Acesso em: 30 dez. 2015. Citado 2 vezes nas páginas 28 e 93.