UNIVERSIDADE FEDERAL FLUMINENSE

INSTITUTO DE MATEMÁTICA E ESTATÍSTICA

MESTRADO PROFISSIONAL EM MATEMÁTICA-PROFMAT

ALGUMAS APLICAÇÕES DA DISTRIBUIÇÃO BINOMIAL

FÁBIO NASCIMENTO DOS SANTOS

Niterói/2015

FÁBIO NASCIMENTO DOS SANTOS

ALGUMAS APLICAÇÕES DA DISTRIBUIÇÃO BINOMIAL

Trabalho de conclusão de curso de Pós-Graduação Profissional em Matemática da Universidade Federal Fluminense, como requisito parcial à obtenção do Grau de Mestre.

Orientadora: Profa. Dra. Cecília de Souza Fernandez

Niterói/2015

Ata dos trabalhos finais da Comissão Examinadora da monografia apresentada por Fábio Nascimento dos Santos.

Aos onze dias do mês de agosto de dois mil e quinze, reuniram-se na sala de seminário do GMA – 5º andar do Instituto de Matemática e Estatística da Universidade Federal Fluminense, os membros da Comissão Examinadora constituída pelos professores Cecília de Souza Fernandez da Universidade Federal Fluminense; Victor Augusto Giraldo da Universidade Federal do Rio de Janeiro e José Rodrigo de Moraes da Universidade Federal Fluminense, sob a presidência do primeiro, para prova pública de defesa da dissertação intitulada **"Algumas aplicações da distribuição binomial",** apresentada pelo mestrando Fábio Nascimento dos Santos. A defesa da dissertação atende às exigências contidas no Regulamento Específico do Curso de Mestrado Profissional em Matemática – PROFMAT - da Universidade Federal Fluminense. A dissertação foi elaborada sob a orientação da professora Cecília de Souza Fernandez. O mestrando Fábio Nascimento dos Santos fez a exposição de seu trabalho durante 50 minutos, iniciando às 09:00h e concluindo às 09:50h. A seguir, respondeu as questões formuladas pelos integrantes da Comissão Examinadora. Terminada a arguição, realizou-se a reunião da Comissão Examinadora, que apresentou parecer no sentido da aprovação do mestrando Fábio Nascimento dos Santos, considerando-se o trabalho apresentado e a forma com que se houve na apresentação na defesa do mesmo. Para constar, foi lavrada a presente ata que vai assinada pela Secretária Administrativa do Mestrado Profissional em Matemática, pelos membros da Banca Examinadora e pelo mestrando.

Niterói, 11 de agosto de 2015.

Cecília de Souza Fernandez

Profª. Cecília de Souza Fernandez

Prof. Victor Augusto Giraldo

José Rodrigo de Moraes

Prof. José Rodrigo de Moraes

Fábio Nascimento dos Santos

Fábio Nascimento dos Santos (mestrando)

Mariana Lattanzi

Mariana Lattanzi (secretária)

AGRADECIMENTOS

Sou realmente grato a todos aqueles que contribuíram para que eu concretizasse mais esta etapa na minha vida acadêmica, na qual fomenta tudo no que venho me tornando ao longo dos anos.

Agradeço à minha família, em especial aos meus pais, cuja a importância é imensurável em todo processo, aos meus amigos Alexander e Luciano por toda a ajuda a mim dada, os quais tiveram nesta empreitada um papel fundamental. Agradeço também, aos professores do Profmat-UFF e em especial à Professora Cecília, a qual foi paciente, compreensiva, muito atenciosa e dedicada à minha causa, acima de tudo obrigado a todos.

Resumo

Neste trabalho, vamos apresentar a noção de probabilidade e algumas de suas propriedades. Vamos também apresentar o conceito de distribuição de probabilidade binomial, que se baseia no Binômio de Newton, assunto tratado no Ensino Médio e , em geral, apresentado nos livros didáticos sem alguma aplicação. Escolhemos aplicar a distribuição binomial em problemas da área de saúde. Acreditamos que estas aplicações podem motivar as aulas sobre probabilidade ou sobre o Binômio de Newton, ilustrando a aplicabilidade da Matemática nas áreas de ciências biológicas e ciências médicas.

Palavras chaves: Probabilidade, Distribuição Binomial e Ensino.

Abstract

In this work, we present the notion of probability and some of its properties. We will also present the concept of the binomial probability distribution, which is based on Newton's Binomial Theorem, subject of the High School and, in general, presented in didactic books without any application. We choose to apply the binomial distribution in problems about healthcare. We believe that these applications can motivate lessons on probability or on the Binomial Newton, showing the applicability of Mathematics in the areas of life sciences and medical sciences.

Key-word: Probability, Binomial Distribution and Teaching.

Sumário

1. INTRODUÇÃO

O ensino de Probabilidade e Estatística no Brasil, que até a década de 90 do século XX estava, em geral, restrito ao Ensino Superior, foi incorporado à Educação Básica com a publicação dos Parâmetros Curriculares Nacionais, em 1997. Isto se deu ao entendimento de que a influência da Probabilidade e Estatística no mundo atual é grande. De fato, muitos meios de comunicação como jornais, revistas, televisão e internet fazem uso de dados analíticos com fins descritivos ou inferências. Gráficos, tabelas e pesquisas de opinião integram e enriquecem os conjuntos de informações desses meios de comunicação. Para que o público compreenda estas informações, o que influencia na tomada de decisões e na previsão de certas situações, é necessário o conhecimento de conceitos e resultados básicos de Probabilidade e Estatística. No Ensino Fundamental, inicia-se com o tratamento descritivo dos dados, através do ensino de gráficos e tabelas. Índices descritivos numéricos de centro, como média, mediana e moda, também são introduzidos no Ensino Fundamental. Já no Ensino Médio, estuda-se noções básicas de probabilidade. Estas noções são, em geral, apresentadas após o estudo de análise combinatória e o estudo do teorema do Binômio de Newton.

Neste trabalho, vamos apresentar a noção de probabilidade e algumas de suas propriedades. Vamos também apresentar o conceito de distribuição de probabilidade binomial, que se baseia no Binômio de Newton, assunto tratado no Ensino Médio e , em geral, apresentado nos livros didáticos sem alguma aplicação. Vamos aplicar a distribuição binomial em problemas da área da saúde. Acreditamos, assim, estarmos oferecendo ao professor da Escola Básica uma bibliografia complementar àquela oferecida pelos livros didáticos, podendo motivar as aulas sobre probabilidade ou sobre o Binômio de Newton, mostrando a aplicabilidade da Matemática nas áreas de ciências biológicas e ciências médicas.

No Capítulo 2, apresentamos a noção de probabilidade e a noção de probabilidade em espaços amostrais finitos equiprováveis. Definimos probabilidade condicional, eventos independentes e apresentamos alguns exemplos.

No Capítulo 3, estudamos variáveis aleatórias e distribuições de probabilidades. Vamos apresentar a definição e dar exemplos de variável aleatória e vamos fazer uma breve explanação do que é uma distribuição de probabilidades.

No Capítulo 4, enunciaremos o Teorema de Binômio de Newton. Uma prova do Binômio de Newton usando o Princípio da Indução Matemática é apresentada. A seguir apresentamos a distribuição binomial.

No Capítulo 5, faremos aplicações da distribuição binomial em problemas das áreas da saúde.

No Capítulo 6, faremos algumas considerações finais, apresentando um pouco sobre o trabalho do matemático americano Steven Smale na área biológica e também sobre o PenCalc, um software desenvolvido na USP utilizado para se determinar riscos de certas doenças genéticas.

2. NOÇÕES SOBRE PROBABILIDADE

A Teoria das Probabilidades é um ramo da Matemática no qual o conceito de probabilidade é a ferramenta principal. A Teoria das Probabilidades estuda modelos probabilísticos. Estes modelos matemáticos permitem fazer predições sobre a frequência com que certos resultados podem ser esperados de ocorrer quando um experimento for repetido um certo número de vezes. Por exemplo, o modelo matemático para se estudar a herança da característica genética cor no processo evolutivo de uma determinada espécie de flor deve ser um modelo probabilístico, pois o modelo a ser usado no estudo deve predizer a cor de maior frequência após um certo número de cruzamentos dentro da mesma população. Um outro exemplo, é na investigação da qualidade da produção de uma determinada peça de carro pela indústria. O modelo matemático que deve ser usado para se predizer a porcentagem de peças defeituosas que são esperadas no processo industrial deve ser um modelo probabilístico.

Na Teoria das Probabilidades, os experimentos que, repetidos sob as mesmas condições, produzem resultados geralmente diferentes são chamados de *experimentos aleatórios*. Como exemplo de um experimento aleatório temos o lançamento de uma moeda para cima, observando a face que irá ficar virada para cima e a face que irá ficar virada para baixo após a queda. A escolha de um aluno dentre 30 alunos de uma classe também é um exemplo de experimento aleatório. Já um experimento é chamado *determinístico* quando, repetido em condições semelhantes, conduz a resultados essencialmente idênticos. Como exemplo de um experimento determinístico temos o lançamento de um dado viciado. Abaixo, temos o exemplo da fabricação de um tal dado.

Dado viciado

Fonte : http://pt.vegasmaster.com/quebrando-as-regras-historias-de-fraudes-em-casinos/

A estratégia mais engenhosa requer equipamentos de ponta.

i. Com uma broca fina, perfura-se o dado. Antes de perfurar, porém, é preciso apoiá-lo em uma prensa para evitar acidentes, caso a broca escorregue.

ii. Para viciar o dado no 6, a face 1 é perfurada. O furo deve ir até 1/4 da altura do cubo para evitar rachaduras. Preenche-se o buraco com algum material mais denso que o dado.

iii. A face perfurada deve ser oposta à do número escolhido para cair mais vezes. O objetivo é fazer com que essa face fique mais pesada e caia mais vezes para baixo. Pinta-se o furo para disfarçar a intervenção.

Infelizmente trapacear com um dado viciado é uma estratégia tão antiga quanto o próprio jogo.

Os fenômenos chamados aleatórios acontecem constantemente em nosso cotidiano. São frequentes perguntas tais como:

Vai chover? Qual será a temperatura máxima amanhã? Qual será o número de acertadores da Mega da Virada? Qual será a população brasileira em 2030?

A seguir, vamos apresentar as noções de espaço amostral e de probabilidade em espaços amostrais. Terminamos observando que, neste texto, vamos trabalhar com espaços amostrais finitos.

2.1 ESPAÇO AMOSTRAL

Consideremos um experimento. O conjunto de todos os possíveis resultados do experimento é chamado *espaço amostral* e é usualmente denotado por Ω. Por exemplo, consideremos o experimento de lançarmos duas moedas. Os possíveis resultados deste experimento são: (Cara, Coroa), (Cara, Cara), (Coroa, Cara), (Coroa, Coroa). Assim,

$$\Omega = \{(\text{Cara, Coroa}), (\text{Cara, Cara}), (\text{Coroa, Cara}), (\text{Coroa, Coroa})\},$$

onde o primeiro elemento do par ordenado indica o resultado do lançamento da primeira moeda e o segundo elemento do par ordenado indica o resultado do lançamento da segunda moeda. Observamos que denotando *Cara* por K e denotando *Coroa* por C, podemos escrever $\Omega = \{(K,C), (K,K), (C,K), (C,C)\}$.

Vejamos um outro exemplo. Consideremos o experimento de retirar duas bolas consecutivas e sem reposição de uma urna que contém 3 bolas brancas, 2 bolas azuis e 4 bolas vermelhas. Temos que o espaço amostral Ω é dado por

$$\Omega=\{(b,b),(b,a),(b,v),(a,b),(a,a),(a,v),(v,b),(v,a),(v,v)\},$$

onde b denota ocorrência de bola branca, a denota ocorrência de bola azul e v denota ocorrência da bola vermelha. Observamos que o primeiro elemento do par ordenado se refere a primeira bola retirada da urna e o segundo elemento do par ordenado se refere a segunda bola retirada.

Os elementos de um espaço amostral são chamados *eventos elementares* ou *eventos simples.* Os subconjuntos de um espaço amostral são chamados *eventos.* Vejamos a seguir alguns exemplos.

Exemplo 1. No lançamento de duas moedas, vimos que $\Omega=\{(K,C), (K,K), (C,K), (C,C)\}$. Os eventos elementares são (C,C), (C,K),(K,K) e (K,C). O subconjunto de Ω dado por $\{(C,C), (C,K)\}$ é um evento. Este evento se traduz como a ocorrência de coroa no lançamento da primeira moeda.

Exemplo 2. No experimento de retirar duas bolas consecutivas e sem reposição de uma urna que contém 3 bolas brancas, 2 bolas azuis e 4 bolas vermelhas, vimos que $\Omega=\{(b,b),$

(b,a), (b,v), (a,b), (a,a), (a,v), (v,b), (v,a), (v,v)}. Os eventos elementares são (b,b), (b,a), (b,v), (a,b), (a,a), (a,v), (v,b), (v,a) e (v,v). Temos que{(b,b), (a,b), (v,b)} é um evento. Este evento se traduz como a ocorrência de bola branca na segunda retirada de duas bolas da urna.

2.2 PROBABILIDADE EM ESPAÇOS AMOSTRAIS FINITOS

Em geral, existem duas abordagens para o estudo da Teoria das Probabilidades. Uma abordagem é não rigorosa e tenta desenvolver no estudante um pensamento intuitivo que o permita "pensar probabilisticamente". A outra abordagem é matematicamente rigorosa, porém usa Teoria da Medida, tópico avançado da Matemática Pura, e, dessa forma, longe do alcance do público alvo do presente trabalho. Assim, neste texto usaremos a abordagem intuitiva para tratar do conceito de probabilidade.

Consideremos um espaço amostral **Ω** formado por k elementos; digamos **Ω**=$\{a_1, a_2, a_3, \ldots, a_k\}$. Vamos associar a cada um dos elementos de **Ω** um número real, $p\{(a_i)\}$, ou simplesmente p_i, chamado de *probabilidade do evento simples* $\{a_i\}$ tal que :

$$i)\ 0 \le p_i \le 1\ para\ todo\ 1 \le i \le k;$$

$$ii)\ \sum_{i=1}^{k} p_i = 1,\ isto\ é,\ p_1 + p_2 + \ldots + p_k = 1.$$

Em seguida, consideremos A um evento qualquer de **Ω**. Definimos *probabilidade do evento A,* e denotamos por P(A), da seguinte forma:

$$iii)\ Se\ A = \phi,\ P(A) = 0;$$

$$iv)\ Se\ A \ne \phi,\ P(A) = \sum_{a_i \in A} p_i.$$

Observemos que se A ≠ Ø, a definição nos dá que a probabilidade de A é a soma das probabilidades dos eventos simples que constituem A. Mas, surge a seguinte pergunta: como obter os números $p_i,\ 1 \le i \le k$? Para respondermos esta pergunta, vamos precisar do conceito de frequência relativa. Seja **Ω** um espaço amostral e $\{a_i\}$ um evento simples. Se, após a realização de um experimento, o resultado é a_i, dizemos que o evento $\{a_i\}$ ocorreu. Agora, suponhamos que realizamos repetidamente o experimento N vezes. Nós podemos contar o número de vezes que o evento $\{a_i\}$ ocorreu nas N vezes que o experimento foi realizado. Este número é chamado frequência relativa de $\{a_i\}$ e é denotada por f_i. Para N pequeno, esse número é variável. Mas, quando N cresce, a frequência relativa tende a se estabilizar. Assim, embora não possamos predizer o resultado de um experimento, nós podemos, para um valor de N suficientemente grande, predizer a frequência relativa do resultado a_i. O número p_i, probabilidade do evento $\{a_i\}$, deve então ser um número próximo da frequência relativa do evento $\{a_i\}$ quando tomamos N grande. Denotando por n_i o número de vezes que ocorre o evento $\{a_i\}$ realizando o experimento N vezes, temos, portanto,

$$f_i = \frac{n_i}{N} \qquad \text{e} \qquad |p_i - f_i| \cong 0, \quad \text{para N grande.}$$

Como exemplo, consideremos o experimento de lançar uma moeda e observamos a face de cima. Em torno de 1750, Georges-Louis Leclerc, conde de Buffon, lançou uma moeda 4048 vezes e observou o resultado "cara" 2048 vezes. Ou seja, ele obteve como frequência relativa do evento {cara} o número $\frac{2048}{4048} \cong 0{,}5059$.

No experimento "lançamento de uma moeda", $\Omega=\{K,C\}$, onde K denota cara e C denota coroa. Se

$$p_1 = p(K) \qquad \text{e} \qquad p_2 = p(C),$$

então, por $i)$ e $ii)$, devemos ter $0 \leq p_1 \leq 1, 0 \leq p_2 \leq 1$ e $p_1 + p_2 = 1.$ De fato, como veremos na próxima seção consideramos

$$p_1 = \frac{1}{2} \qquad \text{e} \qquad p_2 = \frac{1}{2}.$$

Observemos que p_1 é um número próximo da frequência relativa obtida na experiência de Leclerc. Poderíamos atribuir valores quaisquer para p_1 satisfazendo $i)$ e $ii)$. Porém, se p_1

assumir um valor longe da proporção de vezes que um certo resultado ocorre em um número grande de repetições do experimento, então o modelo probabilístico construído não se mostrará útil[1].

2.3 PROBABILIDADE EM ESPAÇOS AMOSTRAIS FINITOS EQUIPROVÁVEIS

Neste trabalho, estamos interessados não apenas em espaços amostrais finitos, mas também em espaços amostrais nos quais, para seus eventos simples, são atribuídos a mesma probabilidade p. Tais espaços amostrais são chamados de *espaços amostrais equiprováveis.*

Seja Ω um espaço amostral equiprovável com k elementos. Digamos $\Omega = \{a_1, a_2,...,a_k\}$. Como definir P(A), onde $A \subset \Omega$? Ora, como Ω é equiprovável, $p(\{a_1\}) = ... = p(\{a_k\})$. Pelo fato de $\sum_{j=1}^{k} p(\{a_j\}) = 1$, segue que $p(\{a_j\}) = \frac{1}{k}$ para todo $1 \leq j \leq k$. Consideremos $A \subset \Omega$; digamos $A = \{a_{j_1}, a_{j_2},...,a_{jr}\}$ com r elementos. De *iv*),

$$P(A) = \sum_{k=1}^{r} p(\{a_{j_k}\}) = \frac{r}{k} = \frac{n(A)}{k},$$

onde n(A) denota o número de elementos do conjunto A. Denotando por n(Ω) o número de elementos de Ω, temos

$$P(A) = \frac{n(A)}{n(\Omega)}, \qquad (1)$$

qualquer que seja $A \subset \Omega$, $A \neq \phi$. Se $A = \phi$, consideramos $P(A) = 0$.

[1] Por i) e ii), vimos que podemos definir a probabilidade de um evento simples de um espaço amostral finito de várias maneiras. Se tomarmos p_i como o número para o qual a frequência relativa do evento $\{a_i\}$ tende a se estabilizar depois de um número grande de vezes em que o experimento é repetido, temos o que chamamos de “abordagem da probabilidade pela frequência relativa” ou “método frequentista”.

Podemos dizer que a proporção $\frac{n(A)}{n(\Omega)}$ é igual aos "casos favoráveis" sobre os "casos possíveis". De fato, a definição de probabilidade como quociente do número de "casos favoráveis" sobre o número de " casos possíveis" foi a primeira definição formal de probabilidade e apareceu pela primeira vez na obra *líber de ludo alial* de Jerônimo Cardano (1501-1576). Esta abordagem é chamada por alguns autores de "abordagem clássica da probabilidade". Observamos se **Ω** for infinito, a definição dada por Cardano não pode ser adotada. Vejamos a seguir alguns exemplos.

Exemplo 3. Uma bola será retirada de uma sacola contendo 5 bolas verdes e 7 bolas amarelas. Qual é a probabilidade desta bola ser verde?
O espaço amostral possui 12 elementos, que é o número total de bolas na sacola. Sendo **Ω** o espaço amostral e A o evento da retirada de uma bola verde temos

$$P(A) = \frac{n(A)}{n(\Omega)} = \frac{5}{12} = 0{,}416.$$

Exemplo 4. Numa urna existem 30 bolas numeradas de 1 a 30. Retirando-se 1 bola ao acaso, qual é a probabilidade de que seu número seja um múltiplo de 4?

O espaço amostral é **Ω** = { 1, 2, 3, ..., 30}. Portanto n(Ω) = 30. O evento A={4, 8, 12, 16, 20, 24, 28} é o evento ocorrência de um múltiplo de 4 entre 1 e 30. Como n(A) = 7, segue que

$$P(A) = \frac{7}{30} = 0{,}233.$$

Exemplo 5. Ache a probabilidade de que, quando um casal tem três filhos, exatamente dois deles sejam meninos. Suponhamos que meninos e meninas sejam igualmente prováveis.

Consideremos A o evento dois meninos em três nascimentos. Temos que

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

onde Ω é o espaço amostral, que consiste em 8 maneiras diferentes pelas quais as três crianças podem ocorrer. Elas são:

1°	2°	3°
menino	Menino	menino
menino	**Menino**	**menina**
menino	**Menina**	**menino**
menina	**Menino**	**menino**
menino	Menina	menina
menina	Menino	menina
menina	Menina	menino
menina	Menina	menina

O que desejamos está representado nas linhas 2, 3 e 4 a tabela acima. Então, como n(A) = 3, segue que P(A) = 3/8 = 0,375.

Observamos que, ao expressarmos o valor de uma probabilidade, devemos dar ou a fração ou o decimal exato ou arredondar o resultado final para três algarismos significativos. A vantagem de se apresentar uma probabilidade em forma decimal é que facilmente a podemos converter em porcentagem: basta multiplicar o resultado por 100.

Como consequência imediata de (1), temos, para um espaço amostral finito e equiprovável Ω, que

$$(a)\ 0 \le P(A) \le 1,\ para\ todo\ A \subset \Omega;$$

$$(b)\ P(\Omega) = 1.$$

Finalizamos esta seção observando que é um erro bastante comum usar (1) para espaços amostrais que não são equiprováveis. Por exemplo, se o espaço amostral $\Omega = \{a_1, a_2\}$, onde a_1 é o evento passar na prova de Álgebra e a_2 é o evento não passar na prova de Álgebra, então não significa que $P(\{a_1\}) = P(\{a_2\}) = \frac{1}{2}$. De fato, não há 50% de probabilidade de passar na prova de Álgebra se o aluno não estudou a matéria. Também, não há 50% de probabilidade de não passar na prova se o aluno estudou e compreendeu toda a matéria dada pelo professor. Assim, não podemos supor que todo espaço amostral finito é um espaço equiprovável.

2.4 CLASSIFICAÇÃO DE EVENTOS

Podemos classificar os eventos por vários tipos. Vejamos alguns deles.

2.4.1 Evento Simples

Um evento A de um espaço amostral **Ω** é dito *evento simples* quando ele é um conjunto unitário. Por exemplo, A={ 5 } é um evento simples do lançamento de um dado cuja face para cima é divisível por 5. Nenhuma das outras possibilidades é divisível por 5.

2.4.2 Evento Certo

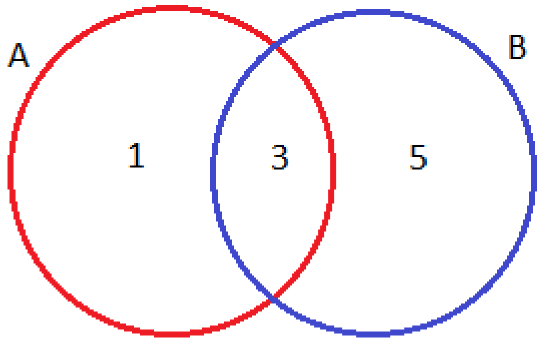
Um evento A de um espaço amostral **Ω** é dito *evento certo* quando A = **Ω**. Por exemplo, consideremos o experimento de lançarmos um dado. O evento A dado pela ocorrência de um número menor ou igual a 6 no lançamento do dado é o conjunto A={1,2,3,4,5,6}.

2.4.3 Evento Impossível

Um evento A de um espaço amostral **Ω** é dito *evento impossível* quando A=Ø. Em outras palavras, A é um evento que nunca ocorre. Por exemplo, no lançamento de dois dados, o evento ocorrência da soma dos números das faces para cima ser 15 é um evento impossível.

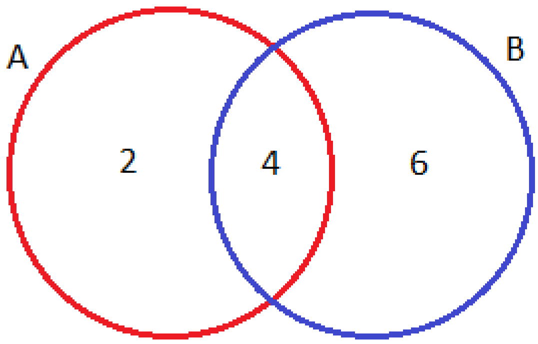
2.4.4 Evento União

Sejam dois eventos de um espaço amostral Ω . O evento $A \cup B$ é chamado *evento união* e é o evento em que A ocorre ou B ocorre. Por exemplo, seja A = **{** 1, 3 **}** o evento de ocorrência da face superior no lançamento de um dado, ímpar e menor ou igual a 3 e B = { 3, 5 **}**, o evento de ocorrência da face superior, ímpar e maior ou igual a 3. Então $A \cup B$ = **{** 1, 3, 5 **}** é o evento de ocorrência da face superior ímpar.



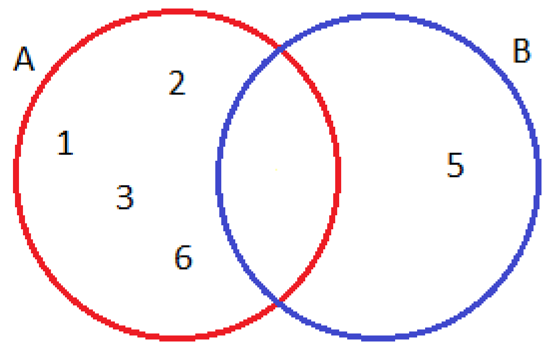
2.4.5 Evento Interseção

Sejam A e B dois eventos de um espaço amostral Ω. O evento $A \cap B$ é chamado *evento interseção* e é o evento em que A ocorre e B ocorre. Por exemplo, seja A = { 2, 4} o evento de ocorrência da face superior no lançamento de um dado, par e menor ou igual a 4 e B = { 4, 6 }, o evento de ocorrência da face superior, par e maior ou igual a 4. Então $A \cap B$ ={ 4 } é o evento de ocorrência da face múltipla de 4.



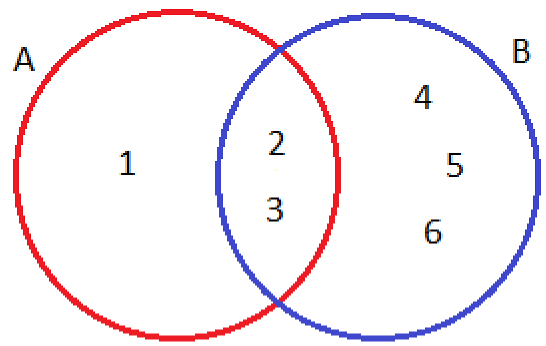
2.4.6 Eventos Mutuamente Exclusivos

Sejam A e B dois eventos de um espaço amostral Ω. A e B são ditos *eventos mutuamente exclusivos* quando $A \cap B = \phi$. Por exemplo, seja A = { 1, 2, 3, 6 } o evento de ocorrência da face superior no lançamento de um dado, um número divisor de 6 e B={5}, o evento de ocorrência da face superior, um número múltiplo de 5. Os eventos A e B são mutuamente exclusivos, pois $A \cap B = \phi$, isto é, os eventos não possuem elementos em comum.



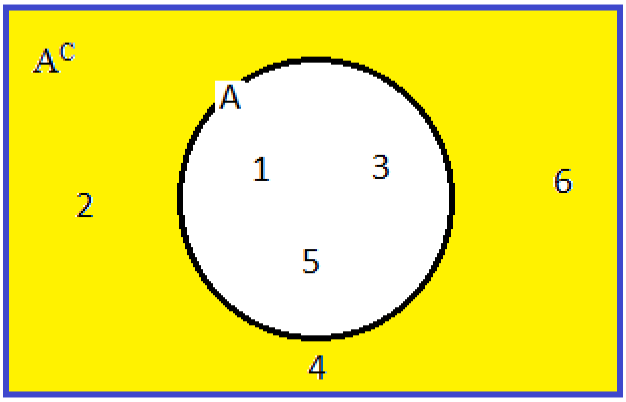
2.4.7 Evento Diferença

Sejam A e B dois eventos de um espaço amostral Ω. O evento A-B é chamado *evento diferença* de A e B e é o evento que contém os elementos de A que não estão em B. Por exemplo, sejam A = {1, 2, 3} e B = {2, 3, 4, 5, 6}, então o conjunto A-B={1}.



2.4.8 Evento Complementar

Seja A um evento de um espaço amostral Ω. O evento A^C é chamado *evento complementar* de A e é o evento B tal que $A \cup B = \Omega$ e $A \cap B = \phi$. Em outras palavras, A^c é o evento que ocorre quando A não ocorre. Por exemplo, seja A = { 1, 3, 5 } o evento de ocorrência de um número ímpar na face superior no lançamento de um dado. O seu evento complementar é A^C = { 2, 4, 6 }, que é o evento de ocorrência de um número par na face superior no lançamento de um dado.



A seguir, vamos apresentar alguns resultados sobre a noção de probabilidade e os tipos de eventos.

Proposição 1. *Sejam A e B dois eventos de um espaço amostral Ω com n elementos. Se A está contido em B, então P(A) ≤ P(B).*

Demonstração. Se A = B, P(A) = P(B) e, portanto, P(A) ≤ P(B). Se A está contido em B e é diferente de B, consideremos $A=\{a_1,a_2,\ldots,a_r\}$ e $B=\{a_1, a_2,\ldots, a_r, a_{r+1},\ldots, a_{r+q}\}$. Então,

$$P(A) = \frac{r}{n} \leq \frac{r+q}{n} = P(B).$$

Portanto, $P(A) \leq P(B)$.

Se A = Ø, temos P(A) = 0. Assim, neste caso também, P(A) ≤ (B). ■

Proposição 2. Se A e B são eventos de um mesmo espaço amostral Ω, então

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

Demonstração. Vamos dividir a demonstração em dois casos:

Caso 1. Suponhamos $A \cap B = \phi$, ou seja, suponhamos que A e B são eventos mutuamente exclusivos.

Se A={a_1, a_2,...,a_p} e B={b_1, b_2,...,b_m}, então $A \bigcup B$ ={ a_1, a_2,...,a_p, b_1, b_2,...,b_m}. Como $A \cap B = \phi$, $A \bigcup B$ tem p + m elementos. Portanto,

$$P(A \cup B) = \frac{n(A \cup B)}{n} = \frac{p+m}{n} = \frac{p}{n} + \frac{m}{n} = \frac{n(A)}{n} + \frac{n(B)}{n} =$$

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

já que $P(A \cap B) = 0.$

Caso 2. Suponhamos agora $A \cap B \neq \phi$.

Temos que

$$P(A \cup B) = P(A - B) + P(A \cap B) + P(B - A), \qquad (2)$$

pelo Caso 1, já que A-B, A∩B e B-A são eventos mutuamente exclusivos. Também, pelo Caso 1,

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

e

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

donde

$$P(A - B) = P(A) - P(A \cap B) \qquad (3)$$

e

$$P(B - A) = P(B) - P(A \cap B). \qquad (4)$$

Substituindo (3) e (4) em (2), obtemos

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

como desejado. ■

A expressão dada na Proposição 2 é usualmente chamada de "regra da adição". Observemos que quando A e B são eventos mutuamente exclusivos, $P(A \cup B) = P(A) + P(B)$.

Proposição 3. *Se A é um evento de um espaço amostral* **Ω**, *então* $P(A^C) = 1 - P(A)$.

Demonstração. Como $A \cap A^C = \varnothing$ e $A \cup A^C = \Omega$, decorre, da Proposição 2, que $P(\Omega) = P(A) + P(A^C)$. Como $P(\Omega) = 1$, $1 = P(A) + P(A^C)$ e, portanto, $P(A^C) = 1 - P(A)$. ■

Vamos a seguir dar alguns exemplos do uso das propriedades acima apresentadas, que podem simplificar o cálculo das probabilidades.

Exemplo 6. Uma pesquisa mostrou que 58% dos brasileiros acreditam que há vida fora da Terra. Qual é a probabilidade de se sortear uma pessoa que não tenha essa crença?

Chamemos A o evento " brasileiro que acredita em Vida fora da Terra". O evento "brasileiro que não tem essa crença" é dado por AC. Como P(AC) = 1 – P(A), segue que $P(A^C) = 1 - P(A) = 1 - 0{,}58 = 0{,}42$.

Exemplo 7. Em um grupo de 500 estudantes, 80 estudam Engenharia, 150 estudam Economia e 20 estudam Engenharia e Economia. Se um aluno é escolhido ao acaso, qual é a probabilidade de que:

a) ele estude Economia e Engenharia?

b) ele estude Economia ou Engenharia?

c) ele não estude Engenharia e nem Economia?

Chamemos A o evento " aluno que estuda Engenharia " e B " aluno que estuda Economia". Temos que:

$A \cap B$ é o evento " aluno que estuda Engenharia e Economia"; $A \cup B$ é o evento " aluno que estuda Engenharia ou Economia" e $A^C \cap B^C = (A \cup B)^C$ [2] é o evento " aluno que não estuda Engenharia e nem Economia".

a) É dado que 20 alunos estudam Engenharia e Economia. Portanto,

$$P(A \cap B) = \frac{20}{500} = 0{,}04.$$

b) Pela Proposição 2,

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

Como $P(A) = \frac{80}{500}$ e $P(B) = \frac{150}{500}$, segue que $P(A \cup B) = \frac{80}{500} + \frac{150}{500} - \frac{20}{500} = \frac{210}{500} = 0{,}42.$

c) Como $A^c \cap B^c = (A \cup B\ )^c$, segue da Proposição 3 que $P((A \cup B\ )^c) =$
$= 1 - P(A \cup B\ ) = 1 - 0{,}42 = 0{,}58.$

Em outras palavras:

4% dos alunos estudam Economia e Engenharia;

42% dos alunos estudam Economia ou Engenharia;

58% dos alunos não estudam Economia e nem Engenharia.

2.5 PROBABILIDADE CONDICIONAL E EVENTOS INDEPENDENTES

Vamos introduzir o conceito de probabilidade condicional através do seguinte exemplo hipotético:

"Um centro de estudos do coração realizou um estudo para avaliar a ocorrência de morte após a realização de uma angioplastia. O estudo envolveu 4.050 pacientes. Dentre os 3.250 homens no estudo, 25 morreram no hospital após o procedimento e dentre as 800 mulheres no estudo, 12 morreram no hospital após o procedimento."

A tabela abaixo descreve os dados acima apresentados:

---

[2] As propriedades $A^C \cap B^C = (A \cup B)^C$ e $A^C \cup B^C = (A \cap B)^C$ são conhecidas como as leis de De Morgan.

Angioplastia e mortalidade após procedimento

Sexo do paciente	Resultado do pós-operatório		
	Morreu	Sobreviveu	Total
Masculino	25	3.225	3.250
Feminino	12	788	800
total	37	4.013	4.050

Analisemos as seguintes perguntas:

a) Qual é a probabilidade de um paciente selecionado aleatoriamente ser homem?

b) Qual é a probabilidade de morte após a angioplastia dado que o paciente é homem?

Chamemos de H o evento de ocorrer um homem e de M o evento de ocorrer morte após a angioplastia. Denotamos o evento de ocorrer morte após a angioplastia dado que o paciente é homem por M/H. Pela definição de probabilidade dada em (1),

$$P(H) = \frac{3.250}{4.050}$$

e

$$P(M/H) = \frac{25}{3.250}.$$

Quando calculamos $P(M/H)$, probabilidade de morte após a angioplastia dado que o paciente é homem, vemos que esta probabilidade é diferente da probabilidade de morte após a angioplastia, que é dada por

$$P(M) = \frac{37}{4.050}.$$

Observamos que ao calcularmos $P(M/H)$, estamos restringindo nosso espaço amostral original, que tem 4.050 elementos. De fato, estamos calculando uma probabilidade chamada probabilidade condicional, ou seja, uma probabilidade baseada em uma condição. Esta condição reduz o espaço amostral original **Ω** para um espaço amostral **Ω**’. No exemplo dado, **Ω**’ tem 3.250 elementos.

Vamos à definição:

Sejam A e B dois eventos de uma espaço amostral Ω. Suponhamos $P(B) > 0$. Definimos a *probabilidade condicional de A dado que B ocorre,* e denotamos por $P(A/B)$, como

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

No cálculo de $P(A/B)$, B se torna o novo espaço amostral Ω'. Notemos que, em nosso exemplo,

$$P(M/H) = \frac{25}{3.250} = \frac{\frac{25}{4.050}}{\frac{3.250}{4.050}} = \frac{P(M \cap H)}{P(H)}.$$

As probabilidades condicionais podem ser particularmente complicadas e, às vezes, exigem muito cuidado. Leia com atenção a seguinte história conhecida:

Os três alunos

Três alunos A, B e C, são candidatos a uma viagem patrocinada pelo Paulo, dono da escola. Paulo decide que apenas um aluno viajará. Ele informa ao diretor da escola sobre sua escolha, mas pede que o nome do aluno escolhido permaneça em segredo por alguns dias.

No dia seguinte, A tenta falar com o diretor para que lhe revele quem foi o contemplado, mas o diretor se recusa. A então pergunta se B ou C não foi escolhido. O diretor pensa por um tempo, então diz a A que B não será escolhido.

Raciocínio do diretor: cada aluno tem 1/3 de chance de ser escolhido. Obviamente, B ou C não deverá ser escolhido; deste modo não dei para A nenhuma informação quanto a se A viajará.

Raciocínio de A: considerando que B não viajará, então A ou C viajará. Então, minhas chances de ser contemplado com a viagem aumentaram para ½ .

Deve ficar claro que o raciocínio do diretor está correto, mas vejamos por quê. Digamos que A, B e C denotam os eventos em que A, B ou C viajará, respectivamente. Sabemos que $P(A) = P(B) = P(C) = \frac{1}{3}$. Digamos que W denota o evento em que o diretor diz que B não viajará. Daí, utilizando probabilidade condicional, A pode atualizar sua probabilidade de viajar para

$$P(A/W) = \frac{P(A \cap W)}{P(W)}.$$

O que está acontecendo pode ser resumido na tabela a seguir:

Aluno escolhido	O diretor diz que é A	
A	B não viaja	cada um com igual prababilidade
A	C não viaja	
B	C não viaja	
C	B não viaja	

Utilizando esta tabela, podemos calcular

$$
\begin{aligned}
&P(W) = P(o\ diretor\ diz\ que\ B\ não\ viajará)\\
&= P(o\ diretor\ diz\ que\ B\ não\ viajará\ e\ que\ A\ viajará)\\
&+ P(o\ diretor\ diz\ que\ B\ não\ viajará\ e\ que\ C\ viajará)\\
&+ P(o\ diretor\ diz\ que\ B\ não\ viajará\ e\ que\ B\ viajará)\\
&= \frac{1}{6} + \frac{1}{3} + 0 = \frac{1}{2}.
\end{aligned}
$$

Portanto, utilizando o raciocínio do diretor, temos

$$P(A/W) = \frac{P(A \cap W)}{P(W)}$$

$$= \frac{P(o\ diretor\ diz\ que\ B\ não\ viajará\ e\ que\ A\ viajará)}{P(o\ diretor\ diz\ que\ B)} = \frac{\frac{1}{6}}{\frac{1}{2}} = \frac{1}{3}.$$

No entanto, A interpreta erroneamente o evento W como sendo igual ao evento B^C e calcula

$$P(A/B^C) = \frac{P(A \cap B^C)}{P(B^C)} = \frac{\frac{1}{3}}{\frac{2}{3}} = \frac{1}{2}.$$

(Adaptado de “Inferência Estatística” de George Casella e Roger L. Berger, tradução da 2° edição norte-americana, págs 20 e 21) .

Da definição de $P(A/B)$, onde A e B são dois eventos de um espaço amostral Ω, temos

$$P(A \cap B) = P(A/B).P(B). \qquad (5)$$

Notemos que podemos escrever também

$$P(A \cap B) = P(B/A).P(A), \qquad (6)$$

uma vez que $P(A \cap B) = P(B \cap A).$ Igualando (5) e (6) , obtemos

$$P(A/B) = P(B/A).\frac{P(A)}{P(B)}. \qquad (7)$$

A expressão em (7) é um caso particular[3] do Teorema de Bayes, que é enunciado a seguir. Este resultado não será provado aqui, mas uma prova do resultado pode ser encontrada em [5].

Teorema de Bayes: *Seja* Ω *um espaço amostral finito e seja* $\{A_1, A_2, ..., A_n\}$ *uma partição de* Ω*. Seja* A *um evento de* Ω*. Então, para todo* $1 \le i \le n$*,*

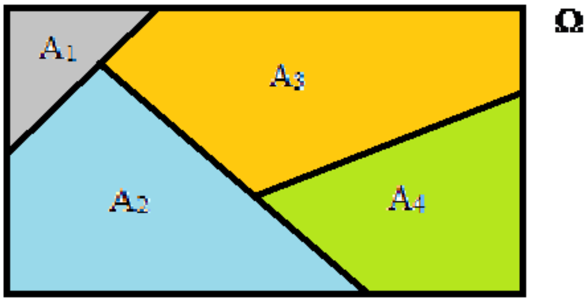
$$P(A_i / A) = \frac{P(A_i).P(A/A_i)}{\sum_{j=1}^{n}[P(A_j).P(A/A_j)]}.$$

Observamos que por uma partição de um conjunto não vazio Ω entendemos um conjunto $\{A_1, A_2, ..., A_n\}$, onde $\phi \ne A_J \subset \Omega$ para todo $1 \le j \le n, A_i \cap A_j = \phi$ para todo $1 \le i, j \le n$ com $i \ne j$ e $\bigcup_{j=1}^{n} A_j = \Omega$. Podemos ilustrar graficamente a situação acima da seguinte forma:

---

[3] A expressão em (7) é um caso particular do Teorema de Bayes.A partição considerada é $\{B, \Omega - B\}$. Daí, temos

$$P(A/B) = \frac{P(A).P(B/A)}{P(B).P(B/B) + P(\Omega - B).P(B/\Omega - B)} = P(B/A).\frac{P(A)}{P(B)}$$, já que $P(B/B) = 1$ e $P(B/\Omega - B) = 0$.



Na figura acima, Ω foi particionado em quatro partes menores, não vazias, duas a duas disjuntas e com a união delas sendo todo o conjunto Ω.

Sejam A e B dois eventos de um espaço amostral Ω. Dizemos que os eventos A e B são *independentes* quando $P(A/B) = P(A)$. Isto significa, por (7), que $P(B/A) = P(B)$. Portanto, A e B são independentes quando sabendo que B tenha ocorrido não afeta a probabilidade de A ocorrer (e vice-versa). Em outras palavras, a ocorrência de A é independente de B ocorrer ou não ocorrer. Dizemos que A e B são *eventos dependentes* quando não são independentes. Vejamos a seguir alguns exemplos.

Exemplo 8. Lançando dois dados honestos simultaneamente, qual é a probabilidade de obtermos um número par no primeiro dado e 5 no segundo dado?

São lançados dois dados. A probabilidade de obtermos um número par no primeiro dado é

$$P(A) = \frac{n(A)}{n(\Omega)} = \frac{3}{6},$$

onde A={2, 4, 6} e $n(\Omega) = 6$.

A probabilidade de obtermos o número 5 no segundo dado é

$$P(B) = \frac{n(B)}{n(\Omega)} = \frac{1}{6},$$

onde B={5} e $n(\Omega) = 6$.

Como o primeiro evento independe do segundo evento, a probabilidade de obtermos, simultaneamente, número par no primeiro e 5 no segundo é

$$P(A\cap B)=P(A).P(B)=\frac{3}{6}.\frac{1}{6}=\frac{3}{36}=\frac{1}{12}.$$

Exemplo 9. Antes de um jogo de futebol entre as equipes A e B, 3 dos 11 jogadores da equipe A e 4 dos 11 jogadores da equipe B ingeriram drogas estimulantes cujo consumo não é permitido pelas regras. O regulamento prevê que 2 jogadores serão sorteados aleatoriamente de cada uma das duas equipes e serão encaminhados ao exame antidoping. Qual é a probabilidade de que nenhum dos jogadores de A e de B selecionados para o antidoping esteja dopado?

A probabilidade de selecionarmos dois jogadores (j_1 e j_2) da equipe A que não estejam dopados é $P(j_1\, j_2)=P(j_1).P(j_2/j_1)=\frac{8}{11}.\frac{7}{10}=0{,}509$, já que temos que selecionar o primeiro jogador que é dado por $\frac{8}{11}$ e o segundo jogador que é dado por $\frac{7}{10}$. Notemos que há uma relação de dependência, onde o resultado $\frac{7}{10}$ depende do primeiro resultado que é $\frac{8}{11}$.

A probabilidade de selecionarmos dois jogadores (j_1 e j_2) da equipe B que não estejam dopados é $P(j_1\, j_2)=P(j_1).P(j_2/j_1)=\frac{7}{11}.\frac{6}{10}=0{,}382$, já que temos que selecionar o primeiro jogador que é dado por $\frac{7}{11}$ e o segundo jogador que é dado por $\frac{6}{10}$. Notemos que há uma relação de dependência, onde o resultado $\frac{6}{10}$ depende do primeiro resultado que é $\frac{7}{11}$.

Como a probabilidade de selecionarmos dois jogadores de A que não estejam dopados é um evento independente do evento de selecionarmos dois jogadores de B que não estejam dopados então, a probabilidade de que nenhum dos jogadores de A e de B selecionados para o antidoping esteja dopado é

$$0{,}509\;.\;0{,}382=0{,}194.$$

Na área da biológica existem vários exemplos de “eventos dependentes” e “eventos independentes”. Por exemplo, “cabelos claros” e “pele clara” são eventos dependentes, porque a probabilidade de uma pessoa ter cabelo claro é maior se a pessoa tem pele clara. Já “cabelos claros” e “estatura baixa” são eventos independentes, porque a probabilidade de uma pessoa ter cabelo claro não aumenta (ou diminui) com a estatura.

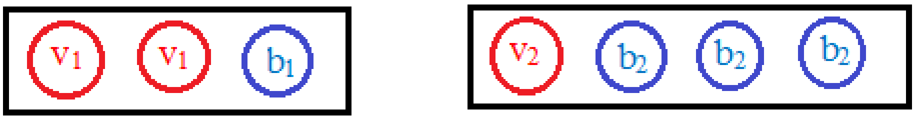
Terminamos esta seção observando que podemos expressar a probabilidade da interseção de dois eventos A e B de um espaço amostral **Ω** por uma expressão. Esta expressão é conhecida como a “regra da multiplicação”. De fato, por (6),

$$P(A \cap B) = P(A).P(B/A). \qquad (8)$$

Se A e B são eventos independentes, podemos escrever (8) da seguinte forma:

$$P(A \cap B) = P(A).P(B). \qquad (9)$$

Exemplo 10. Consideremos duas urnas. Suponhamos que a urna 1 contém duas bolas vermelhas e uma bola branca e a urna 2 contém uma bola vermelha e três brancas. Tira-se uma bola de cada urna. Qual é a probabilidade de que as bolas retiradas sejam de cores diferentes?



urna 1 urna 2

Chamemos v_1, b_1, v_2 e b_2 respectivamente de bola de cor vermelha da urna 1, bola de cor branca da urna 1, bola de cor vermelha da urna 2 e bola de cor branca da urna 2. Se A é o evento “bolas retiradas de cada urna são de cores diferentes”, então A é a união dos eventos B e C, onde B é o evento “ v_1 e b_2 ” e C é o evento “ v_2 e b_1 ”. Como B e C são mutuamente exclusivos,

$$P(A) = P(B) + P(C).$$

Agora, pelo fato de B ser a interseção de dois eventos independentes,

$$P(B) = P(v_1).P(b_2) = \frac{2}{3}.\frac{3}{4} = \frac{1}{2}.$$

Analogamente,

$$P(C) = P(b_1).P(v_2) = \frac{1}{3}.\frac{1}{4} = \frac{1}{12}.$$

Portanto,

$$P(A) = \frac{7}{12}.$$

Exemplo 11. Ache a probabilidade de um casal ter pelo menos uma menina entre 3 crianças. Suponhamos que meninos e meninas sejam igualmente prováveis.

Seja A o evento “ pelo menos uma das 3 crianças é menina”. O complementar de A é “nenhuma menina”. Ou seja, o complementar de A é “todas as três crianças são meninos”. Assim,

$$P(A^C) = P(M_1 \cap M_2 \cap M_3),$$

onde M_i $(1 \le i \le 3)$ denota o evento “a i-ésima criança é menino”. Como $M_i \cap M_j = \phi$ para todo $1 \le i,\ j \le 3$ com $i \ne j$, segue que

$$P(A^C) = P(M_1).P(M_2).P(M_3) = \frac{1}{2}.\frac{1}{2}.\frac{1}{2} = \frac{1}{8}.$$

Portanto,

$$P(A) = 1 - P(A^C) = 1 - \frac{1}{8} = \frac{7}{8}.$$

Observamos que o princípio usado neste exemplo pode ser resumido da seguinte forma: "P(pelo menos um) = 1- P(nenhum)".

3. VARIÁVEIS ALEATÓRIAS E DISTRIBUIÇÃO DE PROBABILIDADES

Em Matemática, por uma variável entendemos um valor numérico qualquer em um dado conjunto universo. Em geral, denotada por x, a variável varia no conjunto universo considerado de modo a verificar ou não uma dada sentença aberta. Por exemplo, a sentença aberta

$$x^2 + 1 = 0$$

com $x \in R$, ou seja, com x variando no conjunto dos números reais, é sempre falsa. Se assumirmos $x \in C$, ou seja, x variando no conjunto dos números complexos, a sentença é verdadeira para $x = \pm i$.

Na Teoria das Probabilidades, temos as variáveis quantitativas e as variáveis qualitativas. As variáveis quantitativas São variáveis numéricas, que variam em um conjunto universo numérico. E as variáveis qualitativas são variáveis categóricas, que variam em um conjunto universo não numérico. Por exemplo, ao estudarmos " peso ao nascer de crianças no Estado do Rio de Janeiro", a variável <u>peso</u> é uma variável quantitativa. E ao estudarmos "cor de olhos da população brasileira", a variável <u>cor de olhos</u> é uma variável qualitativa. Também temos, na Teoria das Probabilidades, as chamadas variáveis aleatórias. Essas variáveis numéricas são assim denominadas, pois elas provém de um certo experimento, cujos resultados ocorrem ao acaso. É sobre isto que trataremos neste capítulo.

3.1 VARIÁVEL ALEATÓRIA: DEFINIÇÃO E EXEMPLOS

Em muitos experimentos, é mais fácil introduzir uma nova variável do que considerar a variável original, principalmente quando esta for qualitativa. Por exemplo, em uma pesquisa de opiniões, podemos decidir perguntar a 30 pessoas se elas concordam ou discordam de uma determinada questão. Se atribuirmos um “1” a concordar e “0” a discordar, o espaço amostral para este experimento tem 230 elementos, cada um deles sendo uma sequência ordenada de 1s e 0s, de tamanho de 30. É possível reduzir o número de elementos deste espaço amostral!

Pode ser que a única quantidade que interessa seja o número de pessoas que concordam (e, equivalentemente, discordam) entre as 30 e, se definirmos uma variável X como sendo o número de 1s registrados entre 30, facilitaremos o problema. Observemos que a variável X assume valores no conjunto de números inteiros {0,1,2,...30}. Este conjunto passará a ser considerado como um novo espaço amostral. Veremos que será mais fácil trabalhar com ele do que com o espaço amostral original.

Ao especificarmos a variável X, definimos uma função, a partir do espaço amostral original para o novo espaço amostral, determinado pelos valores que a variável X pode assumir. Em geral, temos a seguinte definição:

Uma variável aleatória X é uma função de um espaço amostral Ω nos números reais. A tabela a seguir nos apresenta alguns exemplos de variáveis aleatórias:

Exemplos de variáveis aleatórias	
Experimentos	Variáveis Aleatórias
Lançamento de dois dados	X= soma dos números obtidos no lançamento
Lançamento de uma moeda 25 vezes	X= número de caras em 25 lançamentos
Nascimento de 100 bebês	X= número de meninas em 100 nascimentos

Como vimos, ao definirmos uma variável aleatória, também especificamos um novo espaço amostral (o conjunto de valores da variável aleatória). Agora, devemos verificar que a definição de probabilidade em (1), definida no espaço amostral original, pode ser utilizada para a variável aleatória.

Suponhamos que temos um espaço amostral finito equiprovável Ω={a1,...an}. Consideremos uma variável aleatória X com valores reais em X ={x1,...,xm}. Dado $A \subset \Omega$, vimos que

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

Como definir $P(X = x_i)$, ou seja, como definir a probabilidade da variável aleatória X assumir valor real x_i? Primeiramente, devemos notar que $X = x_i$ se, e somente se, o resultado do experimento aleatório for um $a_j \in \Omega$ de modo que $X(a_j) = x_i$. Podemos, então, definir

$$P(X = x_i) = P(A), \qquad (10)$$

onde $A = X^{-1}(\{x_i\}) = \{a \in \Omega : X(a) = x_i\}$.

Uma observação sobre a notação: variáveis aleatórias sempre serão denotadas com letras maiúsculas e os valores assumidos pelas variáveis serão denotados pelas letras minúsculas correspondentes. Portanto, a variável aleatória X pode assumir o valor real x.

Exemplo 12. Consideremos o experimento de lançar uma moeda três vezes. Definamos a variável aleatória X como sendo o número de vezes que saiu cara. Uma enumeração do valor $x_i (0 \le i \le 3)$ de X para cada ponto a_j $(1 \le j \le 8)$ do espaço amostral Ω é dada na tabela abaixo:

a_j	KKK	KKC	KCK	CKK	CCK	CKC	KCC	CCC
x_i	3	2	2	2	1	1	1	0

O conjunto de valores para a variável aleatória X é {0,1,2,3}. Este conjunto é o novo espaço amostral Ω’, obtido a partir de X. Assumindo que todos os oito pontos de Ω têm a probabilidade $\frac{1}{8}$, por (10) basta simplesmente fazer a contagem, na tabela anterior, para termos a probabilidade nos pontos de Ω’, como mostra a seguinte tabela:

x	0	1	2	3
$P(X = x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Por exemplo, $P(X = 1) = P(\{\text{KCC}, \text{CKC}, \text{CCK}\}) = \frac{3}{8}$.

Exemplo 13. Consideremos o experimento de observar o nascimento de 4 bebês. Definamos a variável aleatória X como sendo o número possível de meninos nos quatros nascimentos. Uma enumeração do valor x_i $(0 \leq i \leq 4)$ de X para cada ponto a_j $(1 \leq j \leq 16)$ do espaço amostral Ω é dada na tabela abaixo, onde h denota menino e m denota menina:

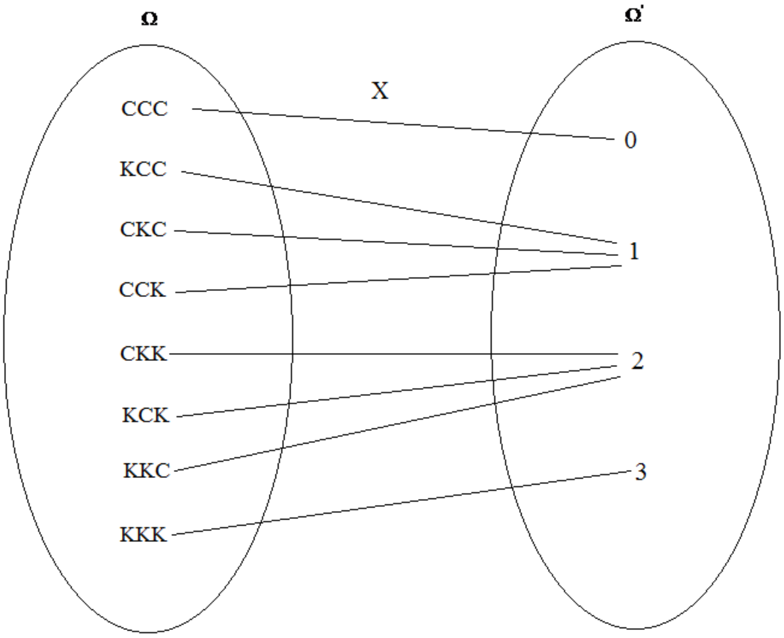
a_j	Hhhh	hhhm, hhmh, hmhh,mhhh	hhmm,hmm, hmmh,mmh, mhmh, mhhm	hmmm, mhmm,mmhm , mmmh	mmmm
x_i	4	3	2	1	0

O conjunto de valores para a variável aleatória X é {0,1,2,3,4}. Este conjunto é o novo espaço amostral Ω’, obtido a partir de X . Por (10) obtemos a probabilidade de cada elemento de Ω’, como mostra a seguinte tabela:

x	0	1	2	3	4
$P(X = x)$	$\frac{1}{16}$	$\frac{4}{16}$	$\frac{6}{16}$	$\frac{4}{16}$	$\frac{1}{16}$

Por exemplo, $P(X = 1) = P(\{hmmm, \text{mmhm, mhmm, mmmh}\}) = \frac{4}{16} = \frac{1}{4}$.

A figura a seguir ilustra a definição (10) para a variável aleatória definida no Exemplo 12.



Temos, por (1) e (10), $P(X=1)=P(\{a\in\Omega : X(a)=1\})=P(\{KCC,CKC,CCK\})=\dfrac{3}{8}$.

3.2 DISTRIBUIÇÃO DE PROBABILIDADES

Como vimos na seção anterior, os valores de uma variável aleatória são determinados pelo resultado de um experimento e, dessa forma, associamos uma probabilidade aos valores possíveis da variável aleatória.

As variáveis aleatórias podem ser classificadas como discretas ou contínuas. Uma variável aleatória é dita discreta se ela assume apenas valores inteiros. E uma variável aleatória é dita contínua quando ela assume valores quaisquer em um intervalo real. Os Exemplos 12 e 13 apresentam variáveis aleatórias discretas. Um exemplo de variável aleatória contínua pode ser o peso de bebês em quatro nascimentos. Notemos que o peso de um bebê pode assumir um valor real. Em geral, ao nascer , um bebê pesa em torno de 2,950 quilogramas.

Na Teoria das Probabilidades, estamos interessados em prever o valor que a variável aleatória X irá assumir, embora essa predição envolva um grau de incerteza. O que se faz, então, é relacionar os valores de uma variável aleatória e a probabilidade de suas ocorrências. No caso das variáveis aleatórias discretas, a função distribuição de probabilidade, ou simplesmente a distribuição de probabilidade, tem essa finalidade. Vamos à definição:

Seja X uma variável aleatória discreta com valores em Ω' . A função distribuição de probabilidade de X denotada por fx , é definida por

$$fx(x) = P(X = x),$$

para todo $x \in \Omega'$.

Consideremos os Exemplos 12 e 13. No Exemplo 12, temos $f_x : \Omega' \to [0,1]$ dada por

$$\begin{cases} f_x(0) = \dfrac{1}{8}, \\ f_x(1) = f_x(2) = \dfrac{3}{8}, \\ f_x(3) = \dfrac{1}{8}. \end{cases}$$

É usual apresentarmos uma função de distribuição de probabilidade por meio de uma tabela. No caso da função f_x acima uma tal tabela é

X = x	P(X=x)
0	1/8
1	3/8
2	3/8
3	1/8

No Exemplo 13, temos $f_x : \Omega' \to [0,1]$ dada por

$$\begin{cases} f_x(0) = f_x(4) = \dfrac{1}{16}, \\ f_x(1) = f_x(3) = \dfrac{4}{16}, \\ f_x(3) = \dfrac{6}{16}, \end{cases}$$

podendo ser representada por meio da tabela

X = x	P(X=x)
0	1/16
1	4/16
2	6/16
3	4/16
4	1/16

Ambas as distribuições de probabilidade acima têm características comuns. Observamos que as probabilidades vão se repetindo. Nas figuras abaixo, temos os gráficos de P(x) :

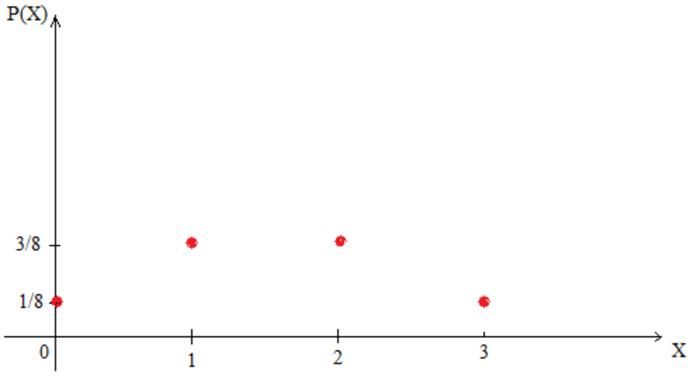


Figura 1: gráfico de P(x) do Exemplo 12

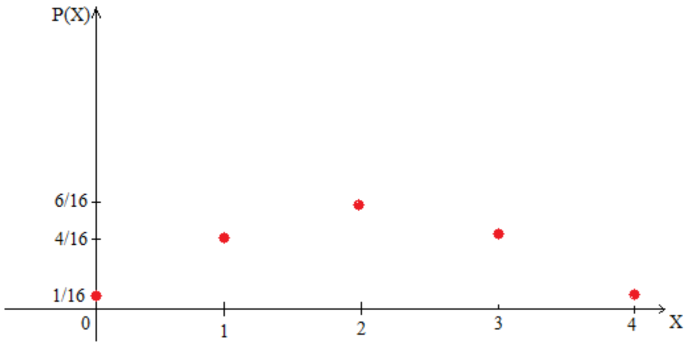


Figura 2: gráfico de P(x) do Exemplo 13

Na Figura 2, vemos que o gráfico de P(x) assume uma forma triangular. Como veremos no próximo capítulo, tais distribuições de probabilidade são distribuições binomiais. As distribuições binomiais se baseiam no Binômio de Newton, cujos coeficientes estão relacionados ao Triângulo de Pascal. Para os interessados em ler sobre o Triângulo de Pascal e o Binômio de Newton, indicamos [1].

4. DISTRIBUIÇÃO BINOMIAL

Neste capítulo, vamos apresentar a distribuição de probabilidade binomial. Esta é uma distribuição para variáveis aleatórias discretas e é largamente usada nas áreas de ciências biológicas e ciências médicas. Como já mencionamos, a distribuição binomial se baseia no Binômio de Newton, que será apresentado a seguir.

4.1 SOBRE O BINÔMIO DE NEWTON

O primeiro contato dos alunos com o Binômio de Newton ocorre no Ensino Fundamental II, com o ensino dos chamados produtos notáveis: $(x+a)^2$ e $(x+a)^3$. Esses produtos são casos particulares do Binômio de Newton. Porém, o que podemos dizer de $(x+a)^4$? Ou, mais geralmente, de $(x+a)^n$ onde $n \in \mathbb{N}$ com $n \geq 4$?

Pela propriedade distributiva da multiplicação em relação à adição, podemos escrever

$$(x+a)^2=(x+a).(x+a) = x^2 + 2ax + a^2$$

e

$$(x+a)^3 = (x+a)^2.(x+a) = (x^2+2ax+a^2).(x+a) =$$
$$= x^3+3ax^2+3a^2x+a^3.$$

A generalização dos casos particulares n = 2 e n = 3 é conhecida como Teorema do Binômio de Newton, que pode ser enunciado da seguinte forma:

Para todo $x, a \in R$ *e* $n \in \mathrm{N}^*$, *temos*

$$(x+a)^n = \binom{n}{0}x^n a^0 + \binom{n}{1}x^{n-1}a^1 + \binom{n}{2}x^{n-2}a^2 + \ldots + \binom{n}{p}x^{n-p}a^p + \ldots + \binom{n}{n-1}x^1 a^{a-1} + \binom{n}{n}x^0 a^n$$

ou seja,

$$(x+a)^n = \sum_{p=0}^{n}\binom{n}{p}x^{n-p}a^p.$$

Lembremos que $\binom{n}{p} = \dfrac{n!}{p!(n-p)!}$, que é chamado de combinação simples de n elementos tomados p a p, sendo n $\geq$ p com $n\ e\ p \in N$.

Usaremos o Princípio da Indução Matemática para demonstrar o Binômio de Newton. Claramente a fórmula é válida para n =1. Suponhamos a fórmula válida para n = k e mostraremos que ela é válida para n = k +1.

Ora,

$$(x+a)^{k+1} = (x+a)^k.(x+a)^1 = \left[\sum_{p=0}^{k}\binom{k}{p}x^{k-p}.a^p\right].(x+a),$$

pela hipótese de indução . Pela distributividade da multiplicação em relação à adição em R , temos

$$\left[\sum_{p=0}^{k}\binom{k}{p}x^{k-p}.a^{p}\right].(x+a)=x\left[\sum_{p=0}^{k}\binom{k}{p}x^{k-p}.a^{p}\right]+a\left[\sum_{p=0}^{k}\binom{k}{p}x^{k-p}.a^{p}\right]=$$

$$=\sum_{p=0}^{k}\binom{k}{p}x^{k-p+1}a^{p}+\sum_{p=0}^{k}\binom{k}{p}x^{k-p}a^{p+1}=\sum_{p=0}^{k}\binom{k}{p}x^{k+1-p}a^{p}+\sum_{p=0}^{k}\binom{k}{p}x^{(k+1)-(p+1)}a^{p+1}=$$

$$=\sum_{p=0}^{k}\binom{k}{p}x^{(k+1)-p}a^{p}+\sum_{p=1}^{k+1}\binom{k}{p-1}x^{(k+1)-p}a^{p}=$$

$$=\binom{k}{0}x^{(k+1)-0}a^{0}+\sum_{p=1}^{k}\left[\binom{k}{p}+\binom{k}{p-1}\right]x^{(k+1)-p}.a^{p}+\binom{k}{k}x^{(k+1)-(k+1)}a^{k+1}=$$

$$=\binom{k+1}{0}x^{(k+1)-0}a^{0}+\sum_{p=1}^{k}\binom{k+1}{p}x^{(k+1)-p}.a^{p}+\binom{k+1}{k+1}x^{(k+1)-(k+1)}.a^{k+1}=$$

$$=\sum_{p=0}^{k+1}\binom{k+1}{p}x^{(k+1)-p}.a^{p},$$

como queríamos demonstrar .

Terminamos esta seção observando que a demonstração acima usa a propriedade distributiva da multiplicação em relação à adição em R e um argumento indutivo, formalizado pelo chamado Princípio da Indução Matemática. Não iremos tratar desse assunto aqui. Para os leitores interessados em ler sobre o Princípio da Indução Matemática indicamos [8].

4.2 DISTRIBUIÇÃO DE PROBABILIDADE BINOMIAL

Consideremos um experimento repetido n vezes em condições idênticas. Suponhamos que em cada vez que o experimento é repetido somente dois eventos são possíveis e eles são mutuamente exclusivos. Além disso, a probabilidade dos eventos é a mesma em cada vez que o experimento é repetido.

Para calcularmos a probabilidade de um dos eventos ocorrer nenhuma vez, ou uma vez ou todas as vezes em que o experimento se repetir, utilizamos a distribuição de probabilidade binomial ou, simplesmente, a distribuição binomial.

A distribuição binomial se baseia no Binômio de Newton. Antes de apresentarmos a definição de uma distribuição binomial, voltemos ao Exemplo 12.

No Exemplo 12, o experimento “lançamento de uma moeda três vezes” pode ser reformulado como a repetição do experimento “lançamento de uma moeda” três vezes. Observemos que, em cada lançamento, podem ocorrer dois eventos mutuamente exclusivos: cara ou coroa. Também observemos que a probabilidade de ocorrer cara, assim como a probabilidade de ocorrer coroa, permanece constante em cada repetição do experimento. Chamando de sucesso a ocorrência de cara e de fracasso a ocorrência de coroa (no Exemplo 12, a variável aleatória X foi definida como sendo o número de vezes que saiu cara), nossa pergunta é: qual é o valor de P(X = k)? No Exemplo 12, calculamos P(X = k) para k = 0, 1, 2 e 3 e, assim, determinarmos a distribuição de probabilidade. Porém, se tivermos que repetir este experimento n vezes, n = 10 por exemplo, o método de calcularmos P(X = k) usando as definições (1) e (10) se torna laborioso.

Suponhamos que um experimento seja repetido n vezes. A probabilidade de nessas n repetições do experimento obtermos k sucessos e, em consequência, n-k fracassos em uma ordem predeterminada, por exemplo, os sucessos nas k primeiras repetições e os fracassos nas demais,

$$\underbrace{SSS...S}_{k\ vezes}\underbrace{FFF...FF}_{n-k\ vezes},$$

é

$$\underbrace{ppp...p}_{k\ fatores}.\underbrace{(1-p)...(1-p)}_{n-k\ fatores} = p^k(1-p)^{n-k},$$

pela regra do produto, vista no Capítulo 2. É claro que, em outra ordem, a probabilidade seria a mesma, pois apenas a ordem dos fatores se alteraria. A probabilidade de obtermos k sucessos e n-k fracassos em qualquer ordem é $p^k(1-p)^{n-k}$ multiplicado pelo número de ordens possíveis que é $\binom{n}{k}$. Para escolher uma ordem basta escolher em quais das n provas ocorrerão os k sucessos. Determinamos assim, a distribuição de probabilidade binomial: A probabilidade de ocorrerem exatamente k sucessos em uma sequência de n repetições independentes, de um certo experimento, na qual a probabilidade de sucesso em cada repetição é p, é igual a

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k},$$

onde X é a variável aleatória definida pelo número de sucessos nas n repetições do experimento.

Chamando q = 1- p, temos que $\binom{n}{k} p^k (q)^{n-k}$ é o k-ésimo termo da expansão de $(p+q)^n$, isto é, é o k-ésimo termo de um Binômio de Newton. Como já mencionamos, os coeficientes de um Binômio de Newton podem ser dispostos em uma forma triangular, conhecida como Triângulo de Pascal. A tabela abaixo exibe os coeficientes binomias, sendo n o número de vezes em que o experimento que gera a variável aleatória X, de distribuição binomial, é repetido:

n	X= 0	1	2	3	4	5	6	7	8	9	10	Soma dos Coefientes
1	1	1										2 = 21
2	1	2	1									4 = 22
3	1	3	3	1								8 = 23
4	1	4	6	4	1							16 = 24
5	1	5	10	10	5	1						32 = 25
6	1	6	15	20	15	6	1					64 = 26
7	1	7	21	35	35	21	7	1				128 = 27
8	1	8	28	56	70	56	28	8	1			256 = 28
9	1	9	36	84	126	126	84	36	9	1		512 = 29
10	1	10	45	120	210	252	210	120	45	10	1	1024 = 210

No Ensino Médio, geralmente o professor não aborda esse assunto da maneira apresentada anteriomente. O que geralmente se faz é que o número de ordens possível também pode ser calculado fazendo a quantidade de anagramas dos códigos desejáveis, através da permutação com repetição.

Para ilustrar o uso da distribuição binomial, voltemos ao Exemplo 13. No experimento “nascimento de um bebê” podem ocorrer menino ou menina. Em quatro nascimentos, temos que, em cada nascimento, a probabilidade de nascer menino, permanece constante. Qual é a probabilidade de nascerem apenas 2 meninos em 4 nascimentos?

Pela distribuição binomial,

$$P(X = 2) = \binom{4}{2}\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^2 = 6.\frac{1}{4}.\frac{1}{4} = \frac{3}{8},$$

o que está de acordo com o que já calculamos no Exemplo 13, usando (1) e (10). Observando a linha n = 4 e a coluna X = 2, na tabela acima, vemos imediatamente o resultado.

5. ALGUMAS APLICAÇÕES DA PROBABILIDADE BINOMIAL NAS ÁREAS DE CIÊNCIAS BIOLÓGICAS E CIÊNCIAS MÉDICAS

No século XIX, as regularidades observadas na contagem dos eventos surgem como a possibilidade de alguma previsibilidade do que parecia até então imprevisível, principalmente no campo das ciências humanas e da medicina. Os dados recolhidos serviram fartamente à aplicação do cálculo de probabilidades. As leis probabilísticas tornam-se uma alternativa para leis estritamente causais; leis humanas tão poderosas quanto as naturais. A natureza humana é substituída pelo homem normal (Hacking, 1995). A descrição de regularidade na distribuição dos caracteres ganha estatuto de lei biológica e social. Assistimos, em meados do século XIX, à polêmica entre Claude Bernard (1813-1878), fisiologista e defensor da medicina experimental, crítico do uso da estatística e da ideia de média como expressão da norma biológica, e os partidários da estatística e do uso do cálculo de probabilidades em medicina. Dentre estes últimos havia duas correntes distintas: os que privilegiavam a estatística e o estudo da frequência de doença nas populações, no sentido de observar a associação de eventos , entre eles Pinel (1745-1826), e os que sugeriam o uso da probabilidade como forma de avaliar e controlar a confiabilidade da correspondência entre sinal, sintoma e lesão (Canguilhem, 1994).

Na segunda metade do século XVII, na Inglaterra e na Holanda, a probabilidade ganha um sentido diferente. Desde o ano de 1592, após terem sido atingidos por uma epidemia, os ingleses mantiveram o hábito de sistematizar obituários; as paróquias publicavam pouco antes do Natal uma lista de todos os óbitos ocorridos naquele ano. Baseando-se nessas listas, John Graunt (1620-1674) propôs um método para determinar a mortalidade provável em qualquer idade. Trabalhos na mesma linha foram realizados, na mesma época, por Jean Hudde e Edmond Halley. Todavia, mais de meio século se passaria antes que o cálculo de probabilidade tivesse o seu valor reconhecido. No século XVIII, seus postulados passam a ser aplicados sobre os dados estatísticos com a finalidade de revelar a existência de causas regulares que explicassem a ocorrência e a frequência dos eventos. A probabilidade torna-se, então, um método universal, uma "lógica geral da incerteza" (Du Pasquier, 1926).

Laplace teve um papel significativo na ampliação do campo de uso da probabilidade (Du Pasquier, 1926). Ele teria demonstrado no Essai Philosophique sur les Probabilités que a probabilidade aplica-se a questões fundamentais relativas à vida dos indivíduos e das nações:

> "As questões mais importantes da vida são em sua maioria problemas de probabilidade! Nós podemos mesmo dizer, falando rigorosamente, que quase todos os nossos conhecimentos só são prováveis; e no pequeno número das coisas que nós podemos saber com certeza, nas próprias ciências matemáticas, os principais meios de chegar à verdade, à indução e à analogia são fundados sobre as probabilidades, de sorte que o sistema inteiro dos conhecimentos humanos se liga a esta teoria" (Laplace, 1814, apud Du Pasquier, 1926:23).

A seguir, vamos apresentar algumas aplicações da distribuição binomial nas áreas das ciências biológicas e das ciências médicas. Vimos que o uso das definições (1) e (10) pode se tornar laborioso em certos cálculos de probabilidade. Os matemáticos criaram vários modelos probabilísticos, que são usados em substituição de (1) e (10). No caso particular de um experimento que gera uma variável aleatória X que se distribui binomialmente, o uso da expressão,

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k},$$

vista no Capítulo 4, facilita os cálculos.

Vejamos a seguir alguns problemas, que podem ser resolvidos com a aplicação da distribuição binomial:

Problema 1. O Traço Falciforme nada mais é que uma alteração genética herdada de um dos pais que não é suficiente para se manifestar como doença, e é bastante comum, a partir disso, analisaremos o problema.Um casal descobre que cada um deles é portador do traço falciforme. A probabilidade de que qualquer criança nascida deles tenha anemia falciforme AF é 0,25. Se o casal tivesse 4 filhos, qual seria a probabilidade de todos os 4 filhos terem AF?

Problema 2. Assim como os contraceptivos, há medicamentos onde a eficácia não chega a 100%, como verificaremos nesta outra atividade. Um determinado medicamento usado para o diagnóstico precoce da gravidez é capaz de confirmar casos positivos em apenas 90% das gestantes muito jovens. Isto porque, em 10% das gestantes muito jovens, ocorre uma escamação do epitélio do útero, que é confundida com a menstruação. Nestas condições, qual é a probabilidade de 2, de 3 gestantes muito jovens que fizeram uso desse medicamento, não terem confirmado a gravidez?

Problema 3. Agora temos mais uma atividade onde o cálculo probabilístico é preponderante. .Considerando-se os leucócitos da circulação sanguínea periférica de um ser humano, a proporção de linfócitos é igual a 0,36. Contando-se 20 leucócitos, qual é a probabilidade de se obter 5 linfócitos?

Os três problemas acima envolvem o cálculo de uma probabilidade. São problemas comuns da área médica, que podem ser resolvidos com o uso da distribuição de probabilidade binomial. A maioria dos livros do ensino médio que abordam este conteúdo, foca mais as questões voltadas para os jogos de azar, no entanto só alguns fazem um elo entre a área médica e a teoria da probabilidade. Poucos livros falam do valor histórico deste estudo, uma vez que esta teoria surgiu a partir destes jogos. Estes trabalham muito com exercícios que praticamente só exigem que o aluno decore uma fórmula e aplique, ou seja, exercícios meramente mecânicos.

Os três problemas apresentados anteriormente podem ser abordados no Ensino Médio, através da seguinte estratégia. Fazendo-se pequenos ensaios como um número reduzido de amostras (como por exemplo, no problema 1, fazer com 1filho, 2 filhos, 3 filhos e assim sucessivamente, em todos os problemas). Isso se dará por meio de permutação com repetição e probabilidade do que se espera e probabilidade do que não se espera, finalmente depois dessas abordagens, mostrar ao aluno que o cálculo que foi feito, pode ser efetuado a partir da aplicação da probabilidade binomial.

Resolução do problema 1. Em cada nascimento, um dos dois eventos mutuamente exclusivos é possível: ter AF ou não ter AF. A probabilidade de uma criança ter AF é constante em cada nascimento: 0,25. Em 4 nascimentos, temos n = 4. Definindo a variável aleatória X como sendo o número de filhos com AF em 4 nascimentos, queremos determinar P(X = 4). Usando, então, a distribuição de probabilidade binomial, temos

$$P(X = 4) = \frac{4!}{4!.0!}.(0,25)^4.(0,75)^0 = (0,25)^4 = 0,00390625.$$

Portanto, a probabilidade é aproximadamente 0,004 de , em 4 nascimentos, todos os filhos apresentarem AF.

Resolução do problema 2. Em cada gestante muito jovem que fizeram uso do medicamento, um dos dois eventos mutuamente exclusivos é possível: ter escamação do epitélio do útero ou não ter a escamação. A probabilidade de uma gestante muito jovem não ter escamação do epitélio do útero é constante em cada uma: 90%. Em 3 gestantes muito jovens, temos n = 3. Definindo a variável aleatória X como sendo o número de gestantes muito jovens que fizeram uso do medicamento, que não tiveram a gravidez confirmada igual a 2, com isso queremos determinar P(X = 2). Usando, então, a distribuição de probabilidade binomial, temos

$$P(X=2)=\frac{3!}{2!.1!}.(0{,}90)^2.(0{,}10)^1=3.0{,}81.0{,}1\cong 0{,}243.$$

Portanto, a probabilidade é 0,243 de , em 2 de 3 gestantes muito jovens que fizeram uso desse medicamento, não terem confirmado a gravidez.

Resolução do problema 3. A cada leucócito contado, ele pode ser um linfócito ou não. A probabilidade de um leucócito ser linfócito é constante: 0,36. Em 20 leucócitos, temos n = 20. Definindo a variável aleatória X como sendo o número de linfócitos contados entre 20 leucócitos, queremos determinar P(X = 5). Usando, então, a distribuição de probabilidade binomial, temos

$$P(X=5)=\frac{20!}{5!.15!}.(0{,}36)^5.(0{,}64)^{15}\cong 0{,}1161.$$

Portanto, a probabilidade de se obter 5 linfócitos em 20 células leucocitárias é aproximadamente de 0,11.

FINALIZAMOS ESTE CAPÍTULO, APRESENTANDO O SOFTWARE PENCALC E O TRABALHO DO MATEMÁTICO STEVEN SMALE, REFERENTE AO ESTUDO DE PROTEÍNAS

A probabilidade está presente em diversas situações que envolvem resultados possíveis (espaço amostral) e resultados favoráveis (eventos). Os jogos de azar, como o dado, as cartas e as loterias, necessitam dos cálculos probabilísticos na determinação das chances de um jogador ganhar ou perder. A Genética é outra área que utiliza a Teoria das Probabilidades, pois os acontecimentos nesse ramo da Biologia envolvem eventos aleatórios, como o encontro dos gametas masculinos e femininos com determinados genes na fecundação.

Acredita-se que um dos motivos para as ideias de Mendel[4] permanecerem incompreendidas durante mais de 3 décadas foi o raciocínio matemático que continham.

---

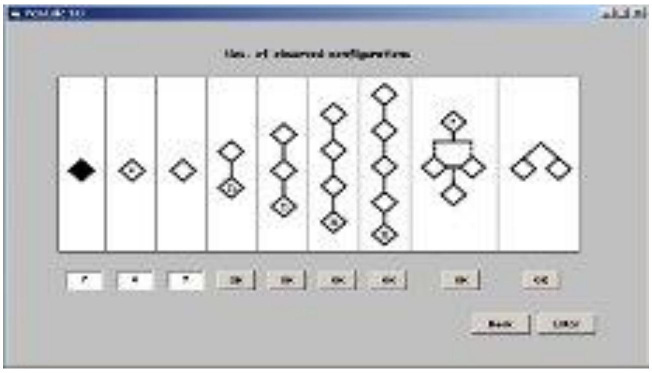
[4]Gregor Johann Mendel (1822-1884) foi um monge agostiniano, botânico e meteorologista austríaco. Durante a sua vida, Mendel publicou dois grandes trabalhos agora clássicos: "Ensaios com plantas híbridas" (*Versuche über Pflanzen-hybriden*), que não abrangia mais de trinta páginas impressas e "Hierácias obtidas pela fecundação artificial" .Em 1865, formula e apresenta em dois encontros da Sociedade de História Natural de Brno as leis da hereditariedade, hoje chamadas *Leis de Mendel*, que regem a transmissão dos caracteres hereditários. Após 1868, as tarefas administrativas

Mendel partiu do princípio que a formação dos gametas seguia as leis da probabilidade, no tocante a distribuição dos fatores.

Tentou-se aqui, neste trabalho, estabelecer a significância de um assunto estudado pelos alunos do Ensino Médio. Este assunto, em alguns casos, não é aprofundado como devia, tendendo assim a ser negligenciado, infelizmente.

Outra abordagem de suma importância do estudo da Teoria das Probabilidades é a sua aplicação na área da saúde, onde este estudo a cada dia ganha mais relevância entre os estudiosos tanto da Matemática quando da Medicina. Um bom exemplo disso são as reportagens a seguir:

SOFTWARE CALCULA PROBABILIDADE DE DOENÇAS HEREDITÁRIAS



Tela do PenCalc, que calcula taxa de penetrância de doenças hereditárias

Pesquisa do Instituto de Biociências (IB) da USP desenvolveu um software, denominado PenCalc, que calcula a taxa de penetrância de doenças autossômicas dominantes, que são doenças hereditárias, passadas de geração para geração. Os afetados são heterozigotos Aa, em que A é o gene que determina o defeito. Em muitas doenças autossômicas dominantes, nem todos os heterozigotos manifestam a doença. A porcentagem de casos Aa com a doença é o que se entende por taxa de penetrância. Por exemplo, no

mantiveram-no tão ocupado que não pode dar continuidade às suas pesquisas, vivendo o resto da sua vida em relativa obscuridade. É conhecido como "Pai da Genética" atualmente.

retinoblastoma, um tumor maligno que afeta a retina, apenas 80% dos heterozigotos Aa tem o tumor. Os outros 20% são normais. Portanto, a taxa de penetrância do gene retinoblastoma, ou a probabilidade de uma pessoa Aa manifestar a doença é de 80%.

A estimativa dessa taxa é feita por meio de cálculos aplicados à estrutura de heredogramas, gráficos que representam a herança genética de determinada característica, que mostram como o defeito é transmitido. "Os cálculos de estimativa da taxa de penetrância são muito complexos. O software automatiza essas operações, agilizando a obtenção dos dados", explica a pesquisadora Andréa Horimoto, autora do estudo.

O PenCalc é destinado para uso de profissionais que trabalham com aconselhamento genético. Além de calcular, por meio da análise de genealogias (árvores genealógicas) que mostram a segregação da doença, o valor apropriado da taxa de penetrância, o programa determina o intervalo de credibilidade a 95% da estimativa e fornece a probabilidade associada ao heredograma em estudo. O intervalo de credibilidade é uma medida de confiabilidade da estimativa da taxa de penetrância. "Usando um exemplo fictício, se tivermos um intervalo de credibilidade de 95% de 0.20 a 0.57 para uma determinada estimativa significa que o verdadeiro valor da taxa de penetrância está contido nesse intervalo com uma probabilidade de 95%", explica a pesquisadora.

Existem duas versões do programa. A originalmente desenvolvida no trabalho de Andréa é a versão disponível para download. Posteriormente, foi desenvolvida uma versão que funciona online. "A versão Web tem ferramentas a mais, que fornecem a probabilidade de heterozigose e o cálculo de risco de repetição da doença na prole de qualquer componente de uma família com casos da doença", conta Andréa.

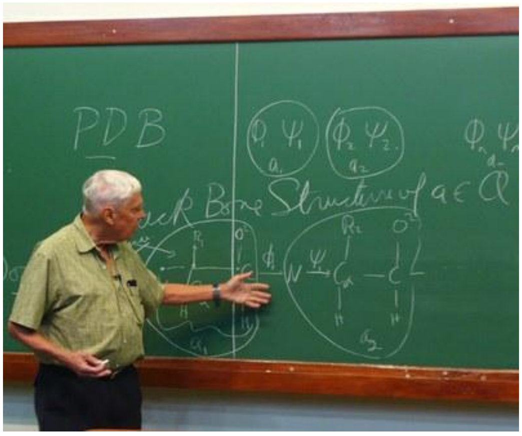
O trabalho foi motivado pelo caso de uma família acompanhada pelo serviço de aconselhamento genético do IB. "Uma síndrome autossômica dominante afetou 53 pessoas de uma família. Analisamos outras 21 genealogias de famílias com a mesma síndrome e testamos o software a partir desse caso", diz a pesquisadora. Ela também conta que a precisão do programa era testada em comparação aos resultados dos cálculos empregando metodologias alternativas.

O PenCalc está disponível para download, desde 2009, na página do professor Paulo Alberto Otto, orientador do trabalho, no site do IB. Por ser gratuito e não exigir cadastro, é difícil mensurar o alcance do software em seu público-alvo.

BIOLOGIA E MATEMÁTICA, APLICADAS E COMBINADAS

Aproximação entre as duas áreas pode dar origem a novos campos. Tendência foi apresentada por matemático norte-americano em palestras que comemoram os 60 anos do IMPA.

Matemáticos preocupados com a estrutura de proteínas e biólogos com a análise dos ângulos formados entre seus aminoácidos. Talvez pareça estranho, mas a aproximação entre matemática e biologia tem se tornado mais comum nas últimas décadas e pode até mesmo dar origem a novos ramos da ciência. Em um ciclo de palestras realizado no contexto da comemoração dos 60 anos do Instituto de Matemática Pura e Aplicada (IMPA), o premiado matemático norte-americano Stephen Smale, da Universidade da Califórnia, nos Estados Unidos, abordou o trabalho de seu grupo e as diversas possibilidades que surgem a partir dessa interação.



matemático norte-americano Stephen Smale

Fonte : http://cienciahoje.uol.com.br/noticias/2012/10/biologia-e-matematica-aplicadas-e-combinadas/image

As possibilidades passam pelo entendimento dos mecanismos de dobras das proteínas. Se descobrirmos esse segredo, vamos responder a uma das questões mais fundamentais da biologia. Na ocasião, Smale apresentou os fundamentos matemáticos dessa nova abordagem e seu potencial para o estudo das estruturas moleculares e das formas de interação entre algumas moléculas básicas para a vida, como os peptídeos. “As possibilidades são muitas, como o aprimoramento de vacinas e o entendimento dos mecanismos por trás das dobras e do enovelamento das proteínas", avalia. "Se descobrirmos esse segredo, vamos responder a uma das questões mais fundamentais da biologia."

O matemático Marcelo Viana, coordenador de atividades científicas do Impa, lembra que uma das principais inspirações da matemática sempre foi a resolução dos problemas colocados pelas ciências experimentais. "Há algumas décadas, com os avanços na genética e na bioinformática, a biologia já vem obtendo bons resultados com a matemática tradicional em áreas como a epidemiologia e a dinâmica evolutiva", explicou. "O trabalho de grupos como o de Smale, no entanto, busca mais; procura elaborar fundamentos matemáticos inovadores para as novas questões experimentais."

Viana compara os esforços desses matemáticos aos de cientistas como Isaac Newton e Gottfried Leibniz, fundamentais para o desenvolvimento do cálculo infinitesimal há quase 400 anos. "Um dos principais motivadores do desenvolvimento do cálculo foi a necessidade de lidar com questões experimentais relacionadas ao estudo da astronomia", avaliou. "Hoje os problemas e desafios apresentados pela biologia podem estimular uma nova visão da matemática, que também poderá beneficiar outras áreas."

Entre as bases dessas novas abordagens, chamadas por Smale de 'novas ciências', estão áreas como a análise combinatória e a topologia – ramo da matemática que estuda as propriedades dos objetos que não se alteram quando eles são torcidos, esticados ou deformados.

No IMPA, Smale aproveitou para lamentar o pouco envolvimento dos matemáticos com ramos como a imunologia. "É uma pena, pois áreas como essa ainda podem se beneficiar muito da estatística e da ciência da computação, por exemplo", ponderou.

O objetivo principal aqui foi deixar claro que a Matemática tem um papel preponderante no estudo das Ciências Biológicas, principalmente no tocante da probabilidade sendo utilizada na genética.

Acredito que, se o professor expor de forma clara o conteúdo sobre probabilidade explorada na Biologia, exemplificando em contextos onde o aluno possa observar essa interação, através de problemáticas que ocorrem tanto na Natureza, quanto no que se diz respeito à saúde de um indivíduo, tal conteúdo será encarado, de modo distinto, mediante ao olhar do aluno no estudo do mesmo.

E isso pode ser feito através da análise da importância dos cálculos probabilísticos voltados à resolução de problemas envolvendo cruzamentos, trabalhando principalmente no

Ensino Médio com as Leis de Mendel; conhecimento dos mecanismos de transmissão dos caracteres, através de aulas expositivas e dialogadas.

Validando desta forma, a importância de se efetuar cálculos de probabilidade para ter conhecimento das possibilidades da existência ou surgimento de determinado genótipo nos descendentes.

Finalizamos, esperando que o presente trabalho sirva de bibliografia complementar ao professor de Matemática do Ensino Médio, ilustrando a aplicabilidade da Matemática em carreiras fora da área de Ciências Exatas e da Terra.

6. **CONSIDERAÇÕES FINAIS**

O trabalho aqui explorado teve como motivação mostrar ao docente um pouco além daquela abordagem feita por ele em sala de aula, fazendo com que ele possa enriquecer seu aprendizado e ensinamentos. Isto fica evidente, por exemplo, quando se trabalha com variável aleatória termo este, que não é utilizado no Ensino Médio, mas que é dever do professor conhecer.

Buscou-se aqui oferecer uma abordagem complementar do estudo da probabilidade, evidenciando casos que o professor pode e deve executar com os seus alunos em sala de aula, não obstante, deve-se salientar que os exemplos usados, de certo modo, destacam a importância da interdisciplinaridade, fato este que é colocado neste trabalho a partir de exemplos onde a matemática e a biologia são coirmãs.

Neste trabalho tentou-se mostrar que o estudo da Teoria das Probabilidades feito de modo contextualizado e bem apresentado, fará com que o aluno tenha mais facilidade ao tratar com tal assunto. Sendo assim, o estudo do Teorema do Binômio de Newton não pode ser negligenciado como ele é, considerando que a Distribuição Binomial Discreta é um assunto evidenciado nos concursos em geral e que depende do ensinamento do Binômio de Newton.

Neste trabalho também tentou-se demonstrar a importância vital que existe entre as disciplinas, onde o docente precisa desbravar isto com o seu aluno. É hesitante observar que a matemática pode e deve auxiliar a área da saúde, entre outras, isto fica claro nas reportagens destacadas no Capítulo 6, onde um programa de computador e um cientista matemático nos mostram tal ligação. Esperamos com isso, fazer com que este trabalho seja mais uma ferramenta auxiliadora dos docentes em sala de aula.

7. BIBLIOGRAFIA

[1] ALEXANDER, Affonso. O Triângulo de Pascal e O Binômio de Newton- Dissertação de Mestrado do PROFMAT. Rio de Janeiro, 2014.

[2] BOYER, Carl B. História da matemática / Carl B. Boyer, revista por Uta C. Merzbach; tradução Elza F. Gomide – 2ª ed. – São Paulo: Edgard Blucher, 1996.

[3] BUSSAB, Wilton de O. MORETTIN, Pedro A. Estatística Básica. 5ª edição. São Paulo: Saraiva, 2006.

[4] CANGUILHEM, G. Etudes d'Histoire et de Philosophie des Sciences. Paris: Vrin, 1994.

[5] DU PASQUIER, G. Le Calcul des Probabilités: Son Évolution Mathématique et Philosophique. Paris: Librairie Scientifique J. Hermann, 1926.

[6] HACKING, I. The Taming of Chance. Cambridge: University Press, 1995.

[7] HAZZAN, Samuel. Fundamentos de Matemática. Vol.5. 6a ed. São Paulo: Atual, 1995.

[8] HEFEZ, Abramo. Elementos de Aritmética. 2ª ed. Coleção Textos Universitários, RJ:SBM, 2011.

[9] HOGG, Robert V. & CRAIG, Allen T.Introduction to Mathematical Statistics.. Fourth Edition. Macmillan Publiching Co., Inc., 1978.

[10] IEZZI, Gelson. Matemática: ciência e aplicações, vol. 2,7 ed.São Paulo: Saraiva, 2013.

[11] MARTINS, Gilberto de A. Estatística Geral e Aplicada. 3ª ed. São Paulo: Atlas, 2005.

[12] MORETTIN, Luiz Gonzaga. Estatística Básica – Volume 1 – Probabilidade – 7ª edição. São Paulo: Pearson Makron Books, 1999.

[13] MORGADO, Augusto César; DE CARVALHO, João B. Pitombeira; CARVALHO, Paulo C. Pinto; FERNANDES, Pedro. Análise Combinatória e Probabilidade, 9º ed., Rio de Janeiro: SBM, 2006. Spiegel, Murray R. Spiegel; tradução e revisão técnica Pedro Consentino-3º Ed.-São Paulo: Makron Books, 1993.

[14] PAIVA,Manoel Rodrigues; MATEMÁTICA,VOL. 2,1° ed. São Paulo: Moderna, 1995.

[15] < http://www.usp.br/agen/p=79810> Acesso em 21/10/2014.

[16] <http://cienciahoje.uol.com.br/noticias/2012/10/biologia-e-matematica-aplicadas-e-combinadas> Acesso em 7/03/2015.