UNIVERSIDADE FEDERAL DO TRIÂNGULO MINEIRO

MESTRADO PROFISSIONAL EM MATEMÁTICA EM REDE NACIONAL

NEILON JOSÉ DE OLIVEIRA

NOÇÕES DE PROBABILIDADE E APLICAÇÕES AO CASO DISCRETO

Uberaba-MG
2013

NEILON JOSÉ DE OLIVEIRA

NOÇÕES DE PROBABILIDADE E APLICAÇÕES AO CASO DISCRETO

Dissertação apresentada ao curso de Mestrado Profissional em Matemática em Rede Nacional-PROFMAT, como parte das atividades para obtenção do título de Mestre em Matemática da Universidade Federal do Triângulo Mineiro - UFTM, Departamento de Matemática.

Uberaba

2013

NEILON JOSÉ DE OLIVEIRA

NOÇÕES DE PROBABILIDADE E APLICAÇÕES AO CASO DISCRETO

Dissertação apresentada ao curso de Mestrado Profissional em Matemática em Rede Nacional-PROFMAT, da Universidade Federal do Triângulo Mineiro, como parte das atividades para obtenção do título de Mestre em Matemática.

____ de ____________ 2013.

Banca Examinadora

____
Prof. Dr. Osmar Aléssio
Orientador
Unviversidade Federal do Triângulo Mineiro

____
Prof. Dr. Ailton Paulo de Oliveira Júnior
Universidade Federal do Triângulo Mineiro

____
Prof. Dr. Ednaldo Carvalho Guimarães
Universidade Federal de Uberlândia

À minha família pelo apoio e carinho que sempre me deu ...

e em especial a meus dois filhos, Breno e Fernanda.

Agradecimentos

Ao término deste trabalho, deixo aqui meus sinceros agradecimentos:

– Agradeço a Deus por ter me dado saúde e paz durante estes dois anos de luta.

– Agradeço a minha família e em especial aos meus pais por estar sempre do meu lado me apoiando.

– Agradeço a CAPES pelo apoio financeiro durante todo o curso.

– Agradeço a todos(as) envolvidos(as) direta ou indiretamente neste grandioso projeto de aperfeiçoamento e melhoria do ensino de matemática em todo o Brasil - PROFMAT.

– Agradeço, também, aos (as) meus (minhas) colegas de curso pela amizade e companheirismo.

– Por fim, agradeço ao meu orientador Prof. Dr. Osmar Aléssio, pela competência e principalmente pelo incentivo e paciência que teve comigo durante o desenvolvimento deste trabalho.

Somente no incerto confio;
Na certeza, apenas, tenho dúvida
E é no cego acaso que busco meu saber.
François Villon

Resumo

O objetivo deste estudo é oferecer aos alunos e professores dos cursos de graduação que possuem uma carga horária reduzida, como Administração, Ciências Contábeis, Psicologia, Sistemas de Informação e outros, um texto que facilite o entendimento intuitivo da Teoria das Probabilidades. O texto aborda as três definições de probabilidade existentes com vários exemplos de aplicação e cita várias propriedades com algumas demonstrações. Inicia-se com a noção intuitiva de probabilidade que surgiu com a preocupação de estudar as possibilidades de ganhar ou perder um jogo, conhecida como teoria do azar. Esta noção intuitiva tornou-se ciência a partir da definição de probabilidade dada por Laplace, conhecida como modelo equiprobabilístico, que reduz todos os acontecimentos do mesmo gênero a certo número de casos igualmente possíveis. Interessa-se calcular as chances de ocorrência de um evento a partir da definição dada pela razão entre o número de casos favoráveis ao experimento (evento) e o número de casos possíveis (espaço amostral). Em seguida mostra a limitação deste modelo, que está em contar os elementos do espaço amostral. Nos casos em que o uso de técnicas de contagem não são suficientes, utiliza-se o modelo frequencial que consiste em repetir, em condições semelhantes, uma experiência n vezes e verificar se o evento A ocorreu em n_A dessas experiências. A probabilidade do evento A é assim definida como sendo o limite da razão $\frac{n_A}{n}$ quando $n \to \infty$. Mas nem sempre é possível repetir uma experiência nas mesmas condições. Daí surge a definição subjetiva, na qual cabe ao interessado usar seu conhecimento "*a priori*" para que seja possível calcular a probabilidade de um evento ocorrer, probabilidade "*a posteriori*". Mas como esta definição subjetiva considera a avaliação da crença do observador na ocorrência de um experimento deveria estipular algumas diretrizes a serem seguidas. Com objetivo de estabelecer algumas regras gerais e de comportamentos racionais, Kolmogorov criou uma definição de probabilidade baseada em três axiomas. A partir desta definição foi possível criar modelos probabilísticos variados com algumas hipóteses bem definidas. Como diversas situações reais se aproximam destas hipóteses, estes modelos são muito úteis no estudo, com a finalidade de solucionar problemas que envolvem variáveis aleatórias discretas ou contínuas. Daí a importância das funções e distribuições de probabilidades neste contexto. Finaliza-se o texto citando os modelos de distribuição de probabilidades discreta com alguns exemplos de aplicação.

Palavras-chave: Probabilidade, Definição, Clássica, Frequentista, Axiomática.

Abstract

The objective of this study is to offer the students and teachers of graduation courses that own a reduced workload, like Business, Accounting, Psychology, Information Systems and others, a text which facilitates the intuitive understanding of the Theory of Probabilities. The text addresses the three existent definitions of probability with several examples of application and states some properties with a few demonstrations. It begins with the notion of wining or losing a game, known as the Theory of Chance. This intuitive notion became science starting the definition of probability given by Laplace, known as equiprobabilistic model, which reduces all the facts of the same gender to certain number of cases equally possible. The interest is to calculate the chances of occurrence of an event from the definition given by the reason between the number of favorable cases to the experiment (event) and the number of possible cases (sampling space). Then to show the limitation of this model, which is on counting the elements of sampling space. In the cases which the use of counting techniques are not enough, it is used the fequencial model that consists in repeating, in similar conditions, one experience n times and verify if the event A occurred in n_A of these experiences. The probability of A event is thus defined as being the limit of reason $\frac{n_A}{n}$ when $n \to \infty$. But it is not always possible to repeat one experience in the same conditions. Hence there is the subjective definition, in which is up to the interested use his/her knowledgement "a priori" making it possible to calculate the probability of an event occurs, probability "a posteriori". However as this subjective definition considers the evaluation of the observer's belief in the occurrence of one experiment, it should stipulate some guidelines to be followed. With the objective of establishing some general rules and of rational behaviors, Kolmogorov created a definition of probability based on three axioms. From this definition it was possible to create varied probabilistic models with some well defined hypothesis. As several real situations approach to these hypothesis, these models are very useful in the study, with the purpose of solving problems that involve several discrete or continuous random variables. Thus the importance of functions and distribution of discrete probabilities with some examples of application.

Keywords: Probability, Definition, Classic, Frequentist, Axiomatic.

Sumário

Lista de Tabelas

Lista de Figuras

1 INTRODUÇÃO E OBJETIVOS

Este trabalho inicia-se com um breve histórico do surgimento da noção de probabilidade citando os principais estudiosos que fizeram parte na construção do vasto campo de conhecimentos, atualmente denominado Teoria das Probabilidades. Serão apresentadas as três definições de probabilidade: a definição clássica, a frequentista e a subjetiva, sistematizada pelos axiomas de Kolmogorov. Além dessas definições, virão inseridas no contexto outras definições e teoremas de suma importância, que são: a definição de probabilidade condicional e suas propriedades, o teorema de Bayes, o teorema da probabilidade total, a definição de Variável Aleatória, a definição do Valor Médio de uma Variável Aleatória conhecida como Esperança Matemática, as definições de variância e desvio padrão.

Em seguida será dado o conceito de Distribuição de Probabilidades, Função de uma Distribuição de Probabilidade e Função de Distribuição Acumulada, enfatizando que o objetivo de uma distribuição é descrever o comportamento de uma Variável Aleatória. Finalmente serão apresentados alguns modelos probabilísticos para Variáveis Aleatórias Discretas, como a Distribuição Uniforme, Distribuição Binomial, Distribuição Multinomial, Distribuição Geométrica, Distribuição Hipergeométrica, Distribuição de Poisson e a Distribuição Binomial Negativa conhecida também por Distribuição de Pascal. Também terão no decorrer do texto vários exemplos resolvidos que possam relacionar a teoria com a prática a partir de situações problemas do cotidiano.

O objetivo deste estudo é oferecer aos alunos e professores dos cursos de graduação que possuem uma carga horária reduzida, como Administração, Ciências Contábeis, Psicologia, Sistemas de Informação e outros, um texto que facilite o entendimento intuitivo da Teoria das Probabilidades.

2 BREVE HISTÓRICO

As primeiras ideias de probabilidade surgiram com Girolamo Cardano (1501-1576). Girolamo, veja Figura 1, foi um cientista a moda de seu tempo, era médico, professor de Geometria, Astrologia e Filosofia. Seus estudos sobre probabilidade limitou-se aos jogos de azar, cujo interesse estava voltado em planejar estratégias de apostas (MORGADO, et al., 2006).

Figura 1: Girolamo Cardano, médico, filósofo e matemático, nasceu em 24 de setembro de 1501 em Pavia e morreu em 21 de setembro de 1576 em Roma. Fonte: http://www-history.mcs.st-andrews.ac.uk/Biographies/Cardan.html

A limitação no estudo da teoria das probabilidades retardou por muito tempo como disciplina no campo da Matemática. Grandes nomes da história da Matemática são responsáveis pelo corpo de conhecimentos que constitui hoje a Teoria das Probabilidades. Depois de Cardano, entre outros, vieram as ideias do Italiano Galileu Galilei (1564 - 1642), os franceses Blaise Pascal (1623 - 1662) e Pierre de Fermat (1601 - 1665), o neerlandês Christiaan Huygens (1629 - 1695), o inglês Isaac Newton (1642 - 1727), o suíço Jacob Bernoulli (1654 - 1705) também conhecido como James Bernoulli, que estudou um dos processos mais simples e mais utilizados em Probabilidade, conhecido atualmente como Os Ensaios de Bernoulli. O inglês Thomas Bayes (1702 - 1761) merece ser citado, pois foi o matemático conhecido por ter formulado o caso especial do teorema de Bayes que deu origem a um importante campo de pesquisa conhecida atualmente como Inferência Bayesiana. Destacando também, o francês Pierre-Simon de Laplace (1749 - 1827) que publicou, em 1812, o livro Theorie Analytique des Probabilités, no qual aborda a definição clássica de probabilidade que marcou um nova era nos estudos das Teorias das Probabilidades. E mais recentemente, tivemos a participação importantíssima de Andrei Nikolaevich Kolmogorov (1903 - 1987). Kolmogorov, veja Figura 2, participou das principais descobertas científicas do século XX nas áreas de probabilidade e estatística e teoria da informação. Graças a Kolmogorov a probabilidade passou a fazer parte de um contexto formal, permitindo assim o seu desenvolvimento enquanto ramo específico da ciência e permitindo sua utilização nas ciências aplicadas. Um de seus principais trabalhos publicados foi "Grundbegriffe der Wahrscheinlichkeitsrechnung" ("Fundamentos de Teoria das Probabilidades"), em que ele lança as bases da axiomatização da Teoria das Probabilidades e esboça o que seria a teoria da medida. As ideias de Kolmogorov fez com que a probabilidade, que vinha sendo tratada de maneira informal desde o final da idade média, se revestisse de um caráter científico rigoroso tornando-se assim, uma área específica da ciência aplicada (MORGADO, et al., 2006; COSTA, 2005; GONGALVES, 2004).

Figura 2: Andrei Nikolaevich Kolmogorov, matemático, nasceu em 25 de abril de 1903 em Tambov e morreu em 20 de outubro de 1987 em Moscou. Fonte: http://pt.wikipedia.org/wiki/Andrei Kolmogorov

A partir de Kolmogorov o progresso dessa teoria não parou. Com o surgimento de novas tecnologias, computadores com planilhas eletrônicas e softwares que facilitam e agilizam os cálculos, novos estudos foram e estão sendo realizados com maior eficiência, proporcionando aos estudiosos a aplicação da probabilidade na solução de diversos problemas presentes no cotidiano das pessoas.

Atualmente, os meios de comunicação de massa e os meios de comunicação em mídias, entre eles os jornais, as revistas, o rádio, a televisão e, mais recentemente a internet, popularizaram o conhecimento de forma geral, em particular, os conceitos e noções da Teoria das Probabilidades, desmistificando assim a associação inicial de probabilidade e jogos de azar. Mesmo assim, em todo livro que trata de Teoria das Probabilidades faz uso de jogos para introduzir os primeiros conceitos, contudo estudam-se um ato de jogar uma moeda (ou um dado), ou o ato de retirar cartas de um baralho como introdução à probabilidade. Obviamente, há poucas ocasiões em que uma decisão administrativa, ou um estudo de um cientista em um laboratório dependa do resultado da jogada de uma moeda (ou lançamento de um dado), mas o estudo das probabilidades associadas à jogada de moedas (ou lançamento de um dado) propicia prática valiosa para alguns conceitos mais difíceis que um administrador ou um cientista poderá encontrar em estudos adiante.

A Estatística, em épocas recentes, era considerada uma parte da Matemática. Atualmente, é uma ciência paralela à Matemática, é um campo de trabalho e é também uma profissão. Há cursos de Estatística de diversos níveis: ensino médio (como tópico da matemática), superior (bacharel) e pós-graduação (Latu Sensu, Strictu Sensu e Doutorado). Esta ciência faz uso intenso da Teoria das Probabilidades. Existem linhas de pesquisas atualmente limitando-se exclusivamente a este estudo, o que justifica atualmente a separação do matemático, do estatístico e do probabilístico. Pode-se encontrar diversos exemplos que ilustram a aplicação das probabilidades: a previsão da inflação acumulada do período, das tendências do eleitorado, da preferência da população por um determinado programa de televisão ou por um determinado tipo de sabão e os resultados em pesquisas de mercado. Estas probabilidades podem ser obtidas a partir de coletas e análises de dados tomados de uma parte da população denominado amostra estatística. Neste caso são contratados pesquisadores que tenha este conhecimento para que estes cálculos sejam feitos com objetivo de saber quanto a inflação morderá o bolso do consumidor, qual a popularidade dos políticos serem eleitos, qual o programa tem maior audiência, qual o sabão é o preferido pelas donas de casas brasileiras e se um produto poderá ou não ser lançado no mercado.

3 TIPOS DE EXPERIMENTOS

O experimento determinístico é um acontecimento previsível, que sempre repetido em condições semelhantes dará o mesmo resultado. Já o experimento aleatório é algo que não se sabe ao certo se ocorrerá, mas é possível, a partir de informações prévias, verificar quais são as chances de ocorrência (DANTAS, 2004).

3.1 Experimentos determinísticos

Sabe-se que a água pura, sob condições normais, atinge seu ponto de ebulição a uma temperatura de 100°C, veja Figura 3.

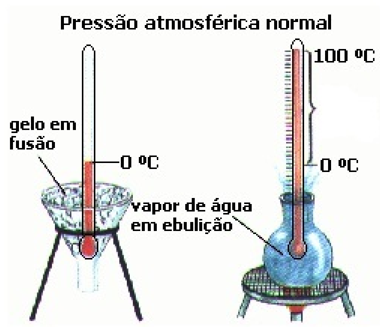

Figura 3: Experimento determinístico referente ao ponto de ebulição da água (à pressão normal) é 100^oC.

Este é um experimento que pode ser repetido várias vezes em condições quase idênticas que se obtém o mesmo resultado, a água irá ferver ao atingir a temperatura de 100°C.

Abandonando-se um objeto a uma altura de 20 metros em um local onde a aceleração da gravidade é $9,8m^2$, pela equação horária do movimento $h = \frac{1}{2}.g.t^2$, pode-se calcular o tempo de queda deste objeto, que será de aproximadamente 2 segundos. A Figura 4 ilustra este experimento. Do mesmo modo, este é um experimento que pode ser repetido várias vezes em condições quase idênticas que se obtém o mesmo resultado, tempo aproximado de queda igual a 2 segundos. A expressão "quase idêntica" foi utilizada por que nem sempre podemos repetir um experimento nas mesmas condições que ocorreu a primeira vez. Experimentos como estes, em que o resultado é previsível, ou seja, sempre

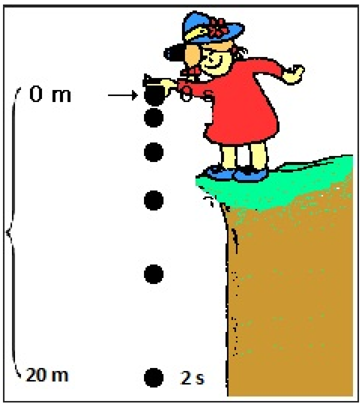

Figura 4: Experimento determinístico referenre ao tempo de queda livre de um corpo é calculado com precisão a partir das equações da mecânica.

que o experimento é repetido várias vezes, em condições quase idênticas, conduzir a um mesmo resultado, são denominados **experimentos determinísticos** (MORGADO, et al., 2006)..

3.2 Experimentos aleatórios

Por outro lado, há experimentos cujos resultados são impossíveis de serem previstos com precisão. A Figura 5 ilustra um experimento aleatório. O resultado de um jogo do campeonato meneiro é imprevisível. Em quase tudo, com maior ou menor grau, dá-se importância ao acaso. Assim, da afirmação "é provável que o Cruzeiro ganhe a partida hoje do Atlético" pode resultar:

Figura 5: Experimento aleatório referente ao jogo do compeonato mineiro.

- que, apesar de que o Cruzeiro apresenta uma melhor equipe, ele perca;
- que, como afirma-se, o Cruzeiro ganhe;

- que o jogo termine empatado.

Como se pode ver, nesta afirmação, o resultado final depende do acaso. Cita-se várias situações em que não se pode dizer com precisão o que ocorrerá. Qual o número de ganhadores do próximo concurso da mega sena? Qual país será o campeão de futebol da copa de 2014? Estes experimentos que se repetidos sob condições quase idênticas produzem resultados geralmente diferentes são denominados **experimentos aleatórios** ou **experimentos probabilísticos**. Outros exemplos de experimentos aleatórios:

1. Conta-se o número de chamadas telefônicas que chegam a uma central das 12:00h às 14:00h;

2. Conta-se o número de veículos que passam por um posto de pedágio das 24:00h às 8:00h. A Figura 6 mostra um posto de pedágio em que chegam aleatoriamente vários veículos em todo instante;

Figura 6: Experimento aleatório que consiste em contar o número de veículos que passam em um posto de pedágio em um determinado período;

3. Contar o número de partículas (de certo tipo) emitidas por uma massa radioativa durante 1 minuto;

4. Verificar o tempo de vida de uma lâmpada de certa marca;

5. Um indivíduo é selecionado de uma população cujos elementos são hipertensos ou não. Observar se o indivíduo é hipertenso ou não. Considera-se como “sucesso”se o indivíduo for hipertenso e “fracasso”caso contrário;

6. Selecionar ao acaso (aleatoriamente) um número real no intervalo (0,1);

7. Um indivíduo chega ao seu serviço, ao acaso, entre 8:00h e 10:00h. Observar o instante em que ele chega;

8. Seja um lote de um item produzido por uma fábrica. Selecionar, ao acaso, um item e observa-se sua qualidade: defeituoso (sucesso) e bom (fracasso);

9. De uma população de diabéticos ou não, relacionar (um de cada vez) ao acaso, sem reposição, até selecionar um diabético. Observa-se:

   (a) A sequência obtida;

   (b) O número de indivíduos não diabéticos;

   (c) O número de seleções.

10. O número de nascimentos registrados em uma região durante cinco anos.

3.3 Espaço amostral e Eventos

A importância dos números naturais provém do fato de que eles constituem o modelo matemático que torna possível o processo de contagem, ou seja, eles respondem perguntas do tipo: "Quantos elementos tem esse conjunto?". Para contar os elementos de um cojunto é necessário usar a noção de correspondência biunívoca que há entre os elementos desse cojunto com os números naturais. Mas esta correspondência não é possível ser feita para alguns conjuntos, como por exemplo, $\mathbb{N}$ e $\mathbb{R}$. Um conjunto é dito **enumerável** se seus elementos podem corresponder biunivocamente com os números naturais. Caso contrário, diz-se que o conjunto é **não enumerável**.

3.3.1 Espaço amostral

Definição 3.3.1. *O conjunto que possui todos os resultados possíveis de um experimento aleatório é denominado* **Espaço Amostral** *e representado por* $\boldsymbol{\Omega}$.

Se $\boldsymbol{\Omega}$ for finito ou infinito enumerável diz-se que $\boldsymbol{\Omega}$ é um **Espaço Amostral Discreto**. Para um espaço amostral discreto, $|\boldsymbol{\Omega}|$ indica o número de elementos de $\boldsymbol{\Omega}$, isto é, a cardinalidade de $\boldsymbol{\Omega}$ (DANTAS, 2004). Os espaços amostrais associados aos experimentos dos exemplos anteriores (seção 3.2) são:

1. Neste experimento pode-se ter zero chamada, ou uma chamada, ou um número muito grande de chamadas, o que leva a concluir que $\Omega = \{0,\ 1,\ 2,\ 3,\ \ldots,\ n,\ \ldots\}$ ou $\Omega = \mathbb{N}$.

2. $\Omega = \mathbb{N}$.

3. $\Omega = \mathbb{N}$.

4. O tempo de vida que uma lâmpada venha assumir pode ser qualquer número real positivo, o que leva a concluir que $\Omega = R_+$.

5. Neste caso o espaço amostral é formado por dois elementos, $\Omega = \{S, F\}$, ou como é comum usar $S = 1$ e $F = 0$, tem-se , $\Omega = \{0, 1\}$, logo $|\Omega| = 2$.

6. Neste experimento o resultado poderá assumir qualquer número real no intervalo, logo o espaço amostral será $\Omega = (0, 1)$.

7. $\Omega = (8, 10)$.

8. $\Omega = \{0, 1\}$ e $|\Omega| = 2$.

9. Neste experimento tem-se:

   (a) $\Omega = \{(S), (FS), (FFS), (FFFS), (FFFFS), \ldots\}$
   ou $\Omega = \{(1), (01), (001), (0001), (00001), \ldots\}$.

   (b) $\Omega = \mathbb{N}$.

   (c) $\Omega = \mathbb{N}^*$

10. $\Omega = \mathbb{N}$.

Nestes 10 exemplos pode-se verificar que somente os exemplos 5 e 8 possuem espaços amostrais finitos, os outros, os espaços amostrais são infinitos. Claro que no exemplo 1 pode-se imaginar um número muito grande de chamadas que seja impossível de ocorrer no período citado, como 2^{2013} chamadas telefônicas. Sendo este número um número natural pode-se dizer que os naturais não poderia ser o espaço amostral deste experimento. No entanto, a escolha do conjunto para o espaço amostral do experimento deve ser feito de modo que ele seja o melhor conjunto que o represente. Também se observa que há espaços amostrais contínuos (os resultados podem ser um número qualquer em um intervalo real) como é o caso do exemplo 6, e casos de espaços amostrais discretos (os resultados podem ser somente números inteiros) como é o caso do exemplo 10. Veja agora outros exemplos de experimentos aleatórios cujos espaços amostrais são finitos:

i) Lançamento de uma moeda e ler a face voltada para cima. O espaço amostral deste experimento é:
$$\Omega = \{cara, coroa\} \quad ou \quad \Omega = \{K, C\},$$

logo $|\Omega| = 2$. No início de um jogo de futebol, veja Figura 7, o árbitro lança uma moeda e pede que um dos representantes dos times da competição escolha "cara"ou "coroa", com o objetivo de decidirem com quem ficará a posse inicial da bola.

Figura 7: Experimento aleatório referente ao início de um jogo de futebol em que o árbitro lança uma moeda para decidir quem inicia com posse da bola.

ii) Um casal quer sortear dois destinos, para fazer uma viagem. A agência de turismo oferece como opção Fortaleza (FO), São Luís (SL), Manaus (MA) e Recife (RE). O espaço amostral que representa os resultados possíveis deste sorteio é dado pelo conjunto:

$$\Omega = \left\{ \begin{array}{ccc} (FO,\ SL), & (FO,\ MA), & (FO,\ RE), \\ (SL,\ FO), & (SL,\ MA), & (SL,\ RE), \\ (MA,\ FO), & (MA,\ SL), & (MA,\ RE), \\ (RE,\ FO), & (RE,\ SL), & (RE,\ MA) \end{array} \right\},$$

logo:

$$|\Omega| = 12$$

se a ordem de visita nas cidades interferir na escolha. Caso contrário, o espaço amostral será:

$$\Omega = \{(FO,\ SL),\ (FO,\ MA),\ (FO,\ RE),\ (SL,\ MA),\ (SL,\ RE),\ (MA,\ RE)\}$$

sendo $|\Omega| = 6$.

iii) Lançamento de duas moedas diferentes e ler a face voltada para cima. O espaço amostral é:

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

iv) Lançamento de um dado e uma moeda e verificar os resultados possíveis. O espaço amostral é:

$$\Omega = \{1K,\ 2K,\ 3K,\ 4K,\ 5K,\ 6K,\ 1C,\ 2C,\ 3C,\ 4C,\ 5C,\ 6C\}\,.$$

v) Lançamento de dois dados diferentes, e ler os números da face superior de cada dado (veja figuara 8).

Figura 8: Experimento aleatório referente ao lançamento de dois dados e verificar os números das faces voltadas para cima.

O espaço amostral consiste de todos os pares ordenados (i, j) onde i e j são números inteiros positivos pertencentes ao intervalo [1, 6]. Assim, o espaço amostral é:

$$\Omega = \left\{ \begin{array}{cccccc} (1,1), & (2,1), & (3,1), & (4,1), & (5,1), & (6,1), \\ (1,2), & (2,2), & (3,2), & (4,2), & (5,2), & (6,2), \\ (1,3), & (2,3), & (3,3), & (4,3), & (5,3), & (6,3), \\ (1,4), & (2,4), & (3,4), & (4,4), & (5,4), & (6,4), \\ (1,5), & (2,5), & (3,5), & (4,5), & (5,5), & (6,5), \\ (1,6), & (2,6), & (3,6), & (4,6), & (5,6), & (6,6) \end{array} \right\}.$$

vi) Escolher, aleatoriamente, duas cartas de um baralho de 52 cartas. Neste caso o espaço amostral possui 1326 elementos. De fato, para retirar a primeira carta tem-se 52 opções e para retirar a segunda carta há 51 opções, veja Figura 9. Como a ordem das cartas não muda o resultado (ás e dois = dois e ás) logo, usando um argumento combinatório, tem-se $(52 \times 51) \div 2 = 1326$. Para calcular o número de elementos do espaço amostral pode-se utilizar das técnicas de contagens citadas no Anexo II, já que há 6 opções de escolha para o lançamento de cada dado, então, pelo princípio multiplicativo, tem-se 6x6=36, que é a cardinalidade de Ω. No caso em que o espaço amostral possuir um número grande de elementos é mais cômodo calcular somente o número de elementos desse conjunto, não sendo necessário relacionar seus elementos.

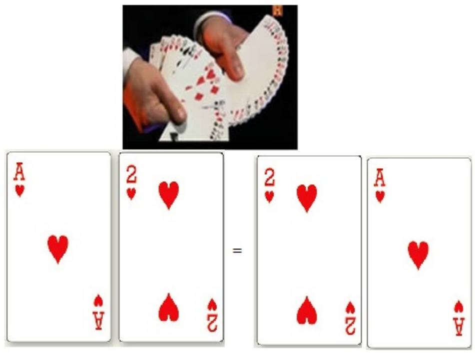

Figura 9: Experimento aleatório referente a escolher duas cartas de um baralho.

Note que neste caso é trabalhoso relacionar todos esses elementos como foi feito nos exemplos anteriores.

vii) O Super Mercado Bretas premiará duas famílias na cidade de Patrocínio-MG que fizeram compras no dia da inauguração. Suponha que em Patrocínio 2.500 famílias fizeram compras neste dia. Assim, por um argumento combinatório (veja ANEXO II) o espaço amostral desse experimento terá $(2.500 \times 2.499) \div 2 = 3.123.750$.

viii) Serão escolhidos dois alunos e quatro alunas do IV período do curso de Administração-UNICERP para participarem de um congresso. Se a sala possui 18 alunos e 12 alunas matriculados, tem-se assim um espaço amostral com 75.735 elementos. De fato, para escolha dos dois alunos tem-se: $(18 \times 17) \div 2 = 153$. Agora, para a escolha das quatro alunas tem-se $(12 \times 11 \times 10 \times 9) \div 4! = 495$. Assim, conclui-se que $|\Omega| = 153 \times 495 = 75.735$.

Cada elemento do conjunto Ω que corresponde a um resultado possível do experimento aleatório de interesse recebe o nome de **ponto amostral**. Assim, no exemplo ii), (FO, SL)$\in \Omega$ e diz-se que (FO, SL) é um ponto amostral de Ω. Note que neste caso Ω possui 12 pontos amostrais. Enquanto que, no exemplo 6 da seção 3.2, o espaço amostral Ω possui infinitos pontos amostrais.

Observa-se que em alguns casos é possível calcular o número de elementos do espaço amostral e relacioná-los, pois tais conjuntos possuem um número pequeno de elementos, como é o caso dos exemplos i), ii), iii), iv) e v), o que não ocorre nos exemplos vi), vii) e viii), nos quais seria muito trabalhoso relacionar os elementos de tais conjuntos. Verifica-se que a Análise Combinatória é uma ferramenta poderosa para o cálculo dos elementos dos espaços amostrais finitos. Esta parte não será tratada neste texto, veja ANEXO II, mas, no texto, será utilizado o princípio multiplicativo e alguns argumentos combinatórios para a resolução de alguns problemas.

3.3.2 Eventos

Dado um experimento aleatório com espaço amostralΩ, um subconjunto de Ω é chamado de **evento** e será representado por letras maiúsculas do nosso alfabeto. Assim, evento é qualquer subconjunto do espaço amostral de um experimento. Realizando um experimento diz-se que o evento ocorre se ocorrer um resultado do experimento (ou ponto amostral) que pertença a este evento. Considere o experimento do exemplo v), lançar dois dados e observar a face voltada para cima. Seja o eventos A = "ocorrer números iguais nos dois dados"e B = "ocorrer máximo faces menores que 3". Temos então que

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

e

$$B = \{(1,1),(1,2),(2,1),(2,2)\}.$$

No experimento do exemplo 8 da seção 3.2, o espaço amostral é

$$\Omega = \{\{1,0\},\ \{\ \ \},\ \{0\},\ \{1\}\ \ e\ \ \{1,0\}\}.$$

Note que são todos os subconjuntos (ou eventos) de Ω. O conjunto destes subconjuntos é denominado **conjuntos das partes de Ω** e é representado por $\boldsymbol{\mathcal{P}(\Omega)}$. O evento $\{\ \ \}$ que também pode ser representado por ϕ é o evento impossível e $\{1,0\}$, que é o próprio espaço amostral, é o evento certo. Para o evento $\{0\}$, 0 é o seu único ponto amostral.

Definição 3.3.2. Conjunto das Partes do Espaço Amostral $\boldsymbol{\Omega}$:
Dizemos que o conjunto $\mathcal{P}$ é o conjunto das partes do espaço amostral Ω se:

$$\mathcal{P} = \{A; A \subset \Omega\}.$$

Exemplo 3.3.1. *Considere $\Omega = \{1,2,3\}$ o conjunto das partes de Ω é a coleção de todos os seus subconjuntos. Então*

$$\mathcal{P} = \{\emptyset, \Omega, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}\}.$$

Obs.: Se o conjunto for finito com n elementos, então o conjunto das partes tem 2^n elementos (veja anexo 1).

Definição 3.3.3. Partição do Espaço Amostral Ω:
Dizemos que os eventos $A_1, A_2, A_3, \cdots, \ A_n$ formam uma partição do espaço amostral Ω se:

a) $A_i \neq \emptyset$, $i = 1, 2, \cdots, n$

b) $A_i \bigcap A_j = \phi$, *para* $i \neq j$

c) *Se* $\bigcup_i^n A_i = \Omega$.

Isto significa que os conjuntos $A_1, A_2, A_3, \cdots, \ A_n$ são disjuntos dois a dois e sua união é o conjunto Ω (MORGADO, et al., 2006).

Definição 3.3.4. Álgebra*: Uma classe de subconjuntos de Ω, representada por $\mathcal{A}$ é denominada uma álgebra se satisfaz as seguintes propriedades(MAGALHÃES, 2006):*

(A_1) $\Omega \subset \mathcal{A}$

(A_2) *Se* $A \in \mathcal{A}$, *então* $A^c \in \mathcal{A}$

(A_3) *Se* $A_i \in \mathcal{A}$, $i \geq 1$, *então* $\bigcup\limits_{i}^{n} A_i \in \mathcal{A}$

Definição 3.3.5. σ-Álgebra*: Uma classe de subconjuntos de Ω, representada por $\mathcal{F}$ é denominada uma σ-álgebra se satisfaz as seguintes propriedades (MAGALHÃES, 2006):*

(A_1) $\Omega \subset \mathcal{F}$

(A_2) *Se* $A \in \mathcal{F}$, *então* $A^c \in \mathcal{F}$

(A_3) *Se* $A_i \in \mathcal{F}$, $i \geq 1$, *então* $\bigcup\limits_{i}^{\infty} A_i \in \mathcal{F}$

Obs: Uma σ-Álgebra é sempre uma Álgebra.

Exemplo 3.3.2. *Considere $\Omega = \{1, 2, 3\}$ e as seguintes coleções de subconjuntos:*

$$\mathcal{F}_0 = \{\emptyset, \Omega\}$$

$$\mathcal{F}_1 = \{\emptyset, \Omega, \{1\}, \{2, 3\}\}$$

$$\mathcal{F}_2 = \{\emptyset, \Omega, \{1\}, \{2\}, \{1, 3\}, \{2, 3\}\}$$

$$\mathcal{F}_3 = \{\emptyset, \Omega, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}\} = \mathcal{P}.$$

Seriam todas σ-Álgebra?

A coleção $\mathcal{F}_0$ é a menor σ-Álgebra formada a partir dos subconjuntos de Ω denominada σ-Álgebra trivial. A coleção $\mathcal{F}_3$ é o conjunto das partes de $\Omega = \{1, 2, 3\}$, então esta coleção satisfaz os três axiomas.

As coleções $\mathcal{F}_1$ e $\mathcal{F}_2$ contém o espaço amostral e, assim (A_1) está satisfeita para ambas. Vamos verificar (A_2). Observe que os complementos dos elementos de $\mathcal{F}_1$ estão todos em $\mathcal{F}_1$. O mesmo acontece para $\mathcal{F}_2$. Para verificar (A_3) observamos que, como o número de elementos em cada coleção é finito, basta verificar se todas as uniões possíveis com seus elementos pertencem a coleçao. A união com o vazio é inócua e a com Ω dá o próprio Ω, logo vamos nos ocupar das demais uniões que poderiam não pertencer a coleção.

Para $\mathcal{F}_1$ temos

$$\{1\} \bigcup \{2, 3\} = \Omega \in \mathcal{F}_1,$$

e, portanto, a propriedade (A_3) está satisfeita e a coleção $\mathcal{F}_1$ é uma σ-álgebra.

Agora isto não ocorre para a coleção $\mathcal{F}_2$, pois

$$\{1\} \bigcup \{2\} = \{1, 2\} \notin \mathcal{F}_2,$$

logo (A_3) não vale e a coleção $\mathcal{F}_2$ não forma uma σ-álgebra.

Exemplo 3.3.3. *Considere $\Omega = \{A, B, C, \cdots\}$, a coleção $\mathcal{F} = \{\emptyset, A, A^c, \Omega\}$ é uma σ-Álgebra?*

De fato, ela satisfaz a três condições. A coleção $\mathcal{F}$ contém o espaço amostral e, assim (A_1) está satisfeita. Vamos verificar (A_2). Observe que os complementos dos elementos de $\mathcal{F}$ estão todos em $\mathcal{F}$. Para verificar (A_3) observamos que, como o número de elementos em cada coleção é finito, basta verificar se todas as uniões possíveis com seus elementos pertencem a coleção. A união com o vazio é inócua e união com Ω dá o próprio Ω, logo vamos nos ocupar das demais uniões que poderiam não pertencer a coleção, como $A \bigcup A^c = \Omega$ temos que a condição (A_3) é satisfeita.

Definição 3.3.6. A Menor $\boldsymbol{\sigma}$-álgebra*: Em várias situações, nosso interesse é construir uma σ-álgebra que tenha, entre seus elementos, um particular subconjunto $A \subset \Omega$. A coleção $\mathcal{F} = \{\phi,\ A,\ A^c,\ \Omega\}$ é uma σ-álgebra sendo A um dos seus elementos (MAGALHÃES, 2006).*

Obs Qualquer outra σ-álgebra que também contiver A será "maior", isto é, terá os elementos de $\mathcal{F}$ mais alguns. No exemplo 3.3.2 F_1 e F_3 são σ-álgebra, a F_1 é a menor.

A σ-álgebra de Borel em $\mathbb{R}$, denotada por $\mathcal{B}(\mathbb{R})$, é a menor σ-álgebra que contém todos os intervalos abertos reais. É possível verificar que ela pode ser gerada pelos intevalos $(-\infty,\ x)$ com $x \in \mathbb{R}$. Existem outras escolhas para o intervalo gerador, mas o importante é que, qualquer tipo de intevalo dos reais, pode ser obtido através de um número enumerável, finito ou infinito de operações de uniões e interseções com o intervalo acima.

Para ilustrar essas ideias, vamos verificar que o intervalo $[a,\ b]$, com $a < b$, a, b, $\in \mathbb{R}$, está na σ-álgebra gerada por intervalos do tipo $(-\infty,\ x)$ com $x \in \mathbb{R}$.

Observe que, para qualquer $b \in \mathbb{R}$, existe uma sequência $b_n \to b$ de modo que $(-\infty,\ b] = \bigcap_{n}^{\infty}(-\infty, b_n)$. Logo $(-\infty,\ b]$ está em $\mathcal{B}(\mathbb{R})$, pois ele é obtido da interseção infinita enumerável de intervalos do tipo $(-\infty,\ x)$, todos em $\mathbb{R}$. Note que, sendo $(-\infty,\ a)^c = [a,\ \infty)$ este intervalo também está em $\mathcal{B}(\mathbb{R})$. Finalmente, $[a,\ b]$ está na σ-álgebra de Borel, pois é interseção de dois de seus elementos, isto é, $[a,\ b] = [a, \infty) \cap (-\infty, b]$.

A Teoria das Probabilidades estuda as várias formas de calcular as chances de o evento ocorrer antes de realizar o experimento. Há três definições distintas de probabilidade. A definição **clássica** ou Laplaciana, a definição **frequentista** e a definição **subjetiva** ou axiomática, devido a Kolmogorov. Na seção 4 será feito uma abordagem dessas três concepções de probabilidade e quando estas poderão ser ou não utilizadas.

4 PROBABILIDADE

A probabilidade é uma parte da matemática que trata dos processos aleatórios, tentando descrevê-los e colocá-los ao alcance da análise. Tal ciência é de vital importância, pois os processos aleatórios aparecem tanto na atividade científica como na vida cotidiana com relativa frequência. Jogos de azar, prognósticos de resultados esportivos, chances em loterias e sorteios, cálculos de prêmios de seguros, preços de mensalidades de planos de saúde e as chances que um político tem de ser ou não eleito são exemplos de situações nas quais os processos aleatórios cruzam a vida do homem comum. Uma das grandes ironias da vida: viver em um mundo cheio de possibilidades e dúvidas. Pensam nisso todos os dias, porém, poucos entendem a sorte. Se é que alguém entende. Por exemplo: se uma moeda é jogada para cima e dá “cara”cinco vezes seguidas qual é o resultado mais provável na próxima vez? Se você disse “coroa”, pense novamente. “Cara”é tão provável quanto “coroa”, porque eventos aleatórios como cara ou coroa não têm memória do que aconteceu antes. Mas não se cobre muito se você errou essa. Até mesmo matemáticos mais brilhantes já tropeçaram nas leis da probabilidade. Eles levaram séculos para descobrir as regras básicas e aprender como usá-las.

A Teoria das Probabilidades, como atualmente trata-se esta parte da matemática, fornece explicações científicas satisfatórias para várias questões simples, mas intrigantes. Suponha que num jogo de dados que utiliza um dado com faces que tenham a mesma probabilidade de sair (não seja um dado fraudulento) um jogador tenha obtido o número 6 nos últimos seis lançamentos. Será que é sensato supor que no próximo lançamento o número 6 não vai sair novamente? A Teoria das Probabilidades nos diz que não. A probabilidade de sair 6 novamente é exatamente igual a de sair qualquer outro resultado. Analogamente, quando alguém diz que é muito mais difícil de ocorrer em um sorteio da mega-sena as dezenas 20, 21, 22, 23, 24 e 25, do que qualquer outra combinação que não seja uma sequência de números. Curiosamente, as pessoas evitam jogar este tipo de combinação julgando-se equivocadamente ser menos provável.

4.1 Probabilidade - Definição Clássica (Laplace)

Esta noção popular de probabilidade deu lugar na matemática por uma definição formal de probabilidade como sendo o número de "casos favoráveis"pelo "número de casos possíveis"e apareceu pela primeira vez de forma clara na obra *Liber de Ludo Aleae* de Girolamo Cardano (1501-1576). Mais recentemente, veja Figura 10, o francês Pierre-Simon de Laplace (1749 - 1827) publica, em 1812, o livro *Theorie Analytique des Probabilités*, no qual aborda a definição clássica de probabilidade que marca um nova era nos estudos das Teorias das Probabilidades (MORGADO, et al., 2006).

Figura 10: Pierre Simon Laplace (1749 - 1827), matemático, astrônomo e físico frances.

A partir desta definição, poder-se-ia dizer que a probabilidade de retirar uma carta desejada de um baralho de 52 cartas seria $\frac{1}{52}$ e que no lançamento de duas moedas idênticas (não fraudulentas) a probabilidade de obter resultados iguais (cara-cara ou coroa-coroa) é de $\frac{1}{2}$.

4.1.1 Definição Clássica - Pierre Simon - Marquês de Laplace - (1812).

Suponha $\Omega = \{w_1, w_2, w_3, \ldots, w_N\}$ e a cardinalidade de Ω dada por $|\Omega| = N$. Suponha também que todos os pontos amostrais são igualmente possíveis de ocorrerem quando da realização do experimento. Dado um conjunto A, defini-se a probabilidade do evento A por:

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

ou seja,

$$P(A) = \frac{(n^o\ de\ casos\ favoráveis\ a\ A)}{(n^o\ de\ casos\ possíveis)}.$$

Propriedades:

1. A probabilidade de cada ponto amostral é dada por: $P(\{w_j\}) = \frac{|\{w_j\}|}{|\Omega|} = \frac{1}{N}$, com $j \in \{1, 2, \ldots, N\}$, que é denominada probabilidade elementar.

2. $P(\phi) = \frac{|\phi|}{|\Omega|} = \frac{0}{N} = 0$.

3. $P(\Omega) = \frac{|\Omega|}{|\Omega|} = 1$.

4. Dado qualquer evento A tem-se $0 \leq P(A) \leq 1$. De fato, tem-se $0 = |\phi| \leq |A| \leq 1 = |\Omega|$, logo $\frac{0}{N} \leq \frac{|A|}{N} \leq \frac{1}{N} \Leftrightarrow 0 \leq P(A) \leq 1$.

5. Seja A e B eventos mutuamente exclusivos. Então $P(A \cup B) = P(A) + P(B)$. De fato, de $P(A \cup B) = P(A) + P(B) - P(A \cap B)$, veja ANEXO I, e por hipótese $P(A \cap B) = 0$, daí, segue o resultado.

Para o exemplo iv) da seção 3.3.1 citado anteriormente, lançar um dado e uma moeda e verificar os resultados possíveis. Na Figura 11 mostra-se o espaço amostra para o lançamento de um dado e o espaço amostra para o lançamento de uma moeda separadamente.

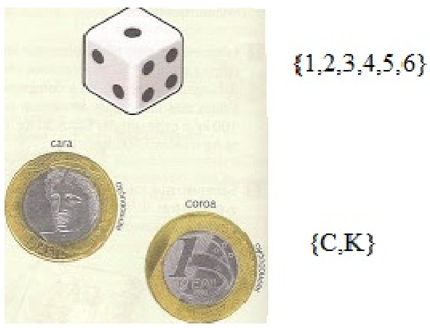

Figura 11: Experimento referente ao lançamento de uma moeda e um dado equilibrados.

O espaço amostral para o lançamente de um dado e uma moeda simultaneamente é

$$\Omega = \{1K, 2K, 3K, 4K, 5K, 6K, 1C, 2C, 3C, 4C, 5C, 6C\}.$$

Seja o evento A = "obter número par e cara". A cardinalidade de A é $|A| = 3$, pois

$$A = \{2K,\ 4K,\ 6K\}.$$

Logo:

$$P(A) = \frac{|A|}{|\Omega|} = \frac{3}{12} = 0,25 \quad ou \quad 25\%.$$

Note que

$$P(\{1K\}) = P(\{2K\}) = \ldots = P(\{6K\}) = \frac{1}{12}.$$

Isto significa dizer que os pontos amostrais de Ω são igualmente prováveis, ou seja, são **equiprováveis**.

De um modo geral se A for um evento qualquer de Ω, então $P(A) = \sum_j P(w_j)$, onde a soma é estendida a todos os pontos amostrais de A.

Exemplo 4.1.1. *Uma fábrica produz determinado artigo. Da linha produção são retirados quatro artigos aleatoriamente, e cada um é classificado como perfeito (P) ou defeituoso (D). O espaço amostral desse experimento é:*

$$\Omega = \left\{ \begin{array}{ccccc} & & PPDD & & \\ & & PDPD & & \\ & DPPP & PDDP & DDDP & \\ & PDPP & DPDP & DDPD & \\ & PPDP & DPPD & DPDD & \\ PPPP & PPPD & DDPP & PDDD & DDDD \end{array} \right\}.$$

Seja A = "obter exatamente dois artigos defeituosos". Então:

$$P(A) = \sum_j P(w_j)$$

$$= P(PPDD) + P(PDPD) + P(DPPD) + P(DPDP) + P(PDDP) + P(DDPP)$$
$$= \frac{1}{16} + \frac{1}{16} + \frac{1}{16} + \frac{1}{16} + \frac{1}{16} + \frac{1}{16}$$

Portanto:

$$P(A) = \frac{6}{16}$$

ou

$$P(A) = 37,5\%.$$

É importante notar que esta definição de probabilidade, denominada **definição Clássica ou Laplaciana**, representa a proporção do número de resultados favoráveis ao evento em relação ao número de resultados possíveis do fenômeno, quando todos são considerados **equiprováveis**, o que significa igualmente provável, ou seja, não preferir alguns resultados em detrimento de outros. Isso é fácil de observar quando ocorre algum tipo de simetria no fenômeno estudado. Veja os espaços amostrais para alguns exemplos.

Exemplo 4.1.2. *Um casal quer ter dois filhos. Isto pode ocorrer da seguinte forma (considere M = mulher e H = homem), $0H2M$, $1H1M$ e $2H0M$, ou seja:*

$$\Omega = \left\{ \begin{array}{ccc} & HM & \\ MM & MH & HH \end{array} \right\}.$$

Exemplo 4.1.3. *Um casal quer ter três filhos. Neste caso tem-se 0H3M, 1H2M, 2H1M e 3H0M, ou seja:*

$$\Omega = \left\{ \begin{array}{cccc} & MMH & HHM & \\ & MHM & HMH & \\ MMM & HMM & MHH & HHH \end{array} \right\}.$$

Assim, nestes dois exemplos tem-se a probabilidade de cada evento elementar igual a $\frac{1}{N}$, ou seja, são **equiprováveis**.

Exemplo 4.1.4. *De um grupo de 3 homens e 5 mulheres uma pessoa será sorteada para participar de um evento. Quer saber a probabilidade da pessoa escolhida ser do sexo feminino. Veja que agora o evento "a pessoa escolhida é do sexo feminino"tem probabilidade $\frac{5}{8}$ e a probabilidade do evento "a pessoa escolhida é do sexo masculino"é $\frac{3}{8}$ são diferentes. Logo, estes dois eventos* **não são equiprováveis**.

Exemplo 4.1.5. *Suponha que um dado seja lançado com*

$$\Omega = \{1, 2, 3, 4, 5, 6\} \quad e$$

$$P(\{j\}) = kj,$$

para k inteiro, ou seja, a probabilidade é proporcional ao ponto amostral. Assim tem-se:

$$P(\{1\}) = k \neq P(\{3\}) = 3k,$$

o que mostra serem eventos não equiprováveis.

Assim, pode-se verificar que esta definição é muito restrita, pois ela não resolve todos os problemas em que se necessita calcular probabilidades. Encontra-se dados de várias formas e com probabilidades diferentes para cada face, o que vai depender de como ele foi fabricado, daí, cada ponto amostral desse experimento pode ter uma probabilidade diferente.

Verifica-se também que a definição clássica só realiza o cálculo de probabilidade de alguns eventos bem simples, utilizando, para isso, a contagem dos elementos dos conjuntos.

Na próxima seção a probabilidade será vista como a frequência relativa em que evento ocorre.

4.2 Probabilidade Frequentista

Suponha um pesquisador interessado em saber que idade predomina entre as pessoas economicamente ativa de determinada cidade. Para isto, a partir de uma amostra bem representativa da população, ele entrevistou 80 pessoas e obteve os seguintes dados da Tabela 1 abaixo, denominados dados brutos, que são dados não organizados numericamente, ou seja, aqueles que não se encontram preparados para análise (dados originados da pesquisa).

28	27	31	33	30	33	27	31
30	33	33	29	32	27	34	37
37	31	30	30	26	29	29	34
30	27	32	24	30	27	31	30
31	31	30	30	27	30	27	27
30	28	33	28	36	29	32	27
33	27	27	30	33	30	33	33
30	39	27	27	31	31	36	28
33	31	31	30	28	23	32	30
29	24	33	30	33	21	35	36

Tabela 1: Dados Brutos.

A partir destes dados brutos obtém-se o Rol, veja Tabela 2, que é o arranjo dos dados brutos em ordem crescente ou decrescente.

21	23	24	24	27	27	27	27
27	27	27	27	27	27	27	27
27	27	28	28	28	28	28	29
29	29	29	29	30	30	30	30
30	30	30	30	30	30	30	30
30	30	30	30	30	31	31	31
31	31	31	31	31	31	31	32
32	32	32	32	33	33	33	33
33	33	33	33	33	33	33	34
34	35	36	36	36	37	37	39

Tabela 2: Rol.

A partir deste rol pode-se construir a Tabela 3 seguinte, denominada Distribuição de

Frequências (DF).

IDADE	FREQUÊNCIA ABSOLUTA	FREQUÊNCIA RELATIVA (em %)
21	1	1,25%
23	1	1,25%
24	2	2,50%
26	1	1,25%
27	13	16,25%
28	5	6,25%
29	5	6,25%
30	17	21,25%
31	10	12,50%
32	4	5,00%
33	12	15,00%
34	2	2,50%
35	1	1,25%
36	3	3,75%
37	2	2,50%
39	1	1,25%
Total	80	100,00%

Tabela 3: Distribuição de Frequência (DF). Fonte: dados hipotéticos.

A **frequência absoluta** é o número de pessoas que apresenta a idade correspondente e a **frequência relativa** é o percentual de pessoas que apresenta esta idade. Essas frequências são consideradas estimativas de quantidades desconhecidas das quais foram extraídas de uma amostra de 80 pessoas de uma população, neste caso, a população de uma cidade. Segundo Bussab (2006, p.103) as **frequências relativas** são estimativas de probabilidades de ocorrências de certos eventos de interesse. Portanto, a probabilidade estimada de que uma pessoa economicamente ativa desta cidade tenha 30 anos é de $21,25\%$.

Esta probabilidade estimada não representa a probabilidade real da população, ela é um valor aproximado do real. Para obter uma probabilidade mais próxima do real teríamos que aumentar o tamanho da amostra ou, quando possível, consultar toda a população. Segundo Bussab (2006, p.103), com suposições adequadas, e sem observar diretamente o fenômeno aleatório de interesse, pode-se criar um modelo teórico que reproduza de maneira razoável a distribuição das frequências, quando o fenômeno é observado diretamente. Tais modelos são denominados **modelos probabilísticos**.

4.2.1 Definição de Probabilidade Frequentista.

A concepção frequentista de probabilidade baseia-se em repetir um experimento um número grande de vezes e fazer uma estimativa do número de vezes que ocorre determinado evento.

Seja A um evento associado a um experimento aleatório. Realiza-se n repetições do experimento e seja n_A o número em que A ocorreu nas n repetições do experimento. Seja

$$f_A(n) = \frac{n_A}{n}$$

a frequência relativa de ocorrência do evento A. A probabilidade frequentista do evento A é dada por

$$P(A) = p \cong f_A(n)$$

para n suficientemente grande, ou seja,

$$P(A) = p = \lim_{n \to \infty} f_A(n).$$

Quando o experimento for repetido um número suficientemente grande de vezes espera-se que surja uma regularidade, isto é, haverá uma estabilidade da fração $f_A(n) = \frac{n_A}{n}$ (frequência relativa).

O gráfico da Figura 12 ilustra esta ideia, para o lançamento de uma moeda equilibrada, que espera-se a probabilidade de ocorrer cara tender a $0,5$.

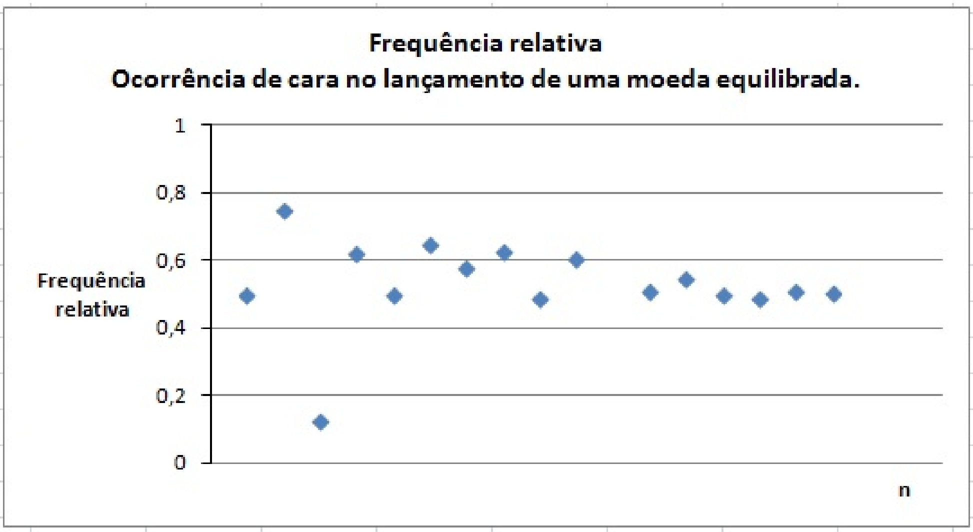

Figura 12: Frequência relativa $f_A(n) = \frac{n_A}{n}$. Fonte: Dados hipotéticos.

Suponha um experimento em que se joga um percevejo, usado para afixar painéis de aviso, sobre uma superfície lisa. Qual seria a probabilidade do evento A = "o percevejo cair com a ponta virada para cima"?

Primeiro deve-se entender que, neste caso, não se pode recorrer a propriedades de simetria, pois, no caso do percevejo, elas não existem. A ideia é aproximar a probabilidade p pela estimativa da probabilidade de ocorrência do evento, ou seja, jogar o percevejo um número n suficientemente grande de vezes, procurando manter as mesmas condições. Assim:

$$P(A) = P(o\ \ percevejo\ \ cair\ \ com\ \ a\ \ ponta\ \ virada\ \ para\ \ cima) = p \cong f_A(n)$$

ou

$$p \cong \frac{n^o\ \ de\ \ vezes\ \ que\ \ cair\ \ com\ \ a\ \ ponta\ \ voltada\ \ para\ \ cima}{n},$$

onde n é o número total de observações, ou seja, número de repetições do experimento. Desta forma, esta razão tende a se estabilizar, isto é, aproximar de um limite.

Para criar um modelo probabilístico de um experimento, deve-se verificar que esse experimento revela (FONSECA; MARTINS, 2006):

1. Cada experimento poderá ser repetido indefinidamente sob condições semelhantes (quase idênticas).

2. Não se conhece um particular valor do experimento *"a priori"*, porém podem-se conhecer todos os possíveis resultados, o espaço amostral.

3. Quando o experimento é repetido um número finito de vezes haverá uma estabilidade das frequências relativas.

Propriedades:

1. $P(\phi) = \frac{n_\phi}{n} = \frac{0}{n} = 0.$

2. $P(\Omega) = \frac{n_\Omega}{n} = 1.$

3. Dado qualquer evento A tem-se $0 \leq n_A \leq n \Leftrightarrow 0 \leq P(A) = \frac{n_A}{n} \leq 1.$

4. Seja A e B eventos mutuamente exclusivos. Então:

$$P(A \cup B) = \frac{n(A \cup B)}{n} = \frac{(n_A + n_B)}{n} = \frac{n_A}{n} + \frac{n_B}{n} = P(A) + P(B).$$

O modelo probabilístico para o lançamento de uma moeda equilibrada é dado por:

Face	Frequência teórica
K (coroa)	$\frac{1}{2}$
C (coroa)	$\frac{1}{2}$
Total	1

Tabela 4: Distribuição de Frequência (modelo probabilístico teórico) para o lançamento de uma moeda.

O modelo probabilístico para o lançamento de um dado equilibrado é dado por:

Face	Frequência teórica
1	$\frac{1}{6}$
2	$\frac{1}{6}$
3	$\frac{1}{6}$
4	$\frac{1}{6}$
5	$\frac{1}{6}$
6	$\frac{1}{6}$
Total	1

Tabela 5: Distribuição de Frequência (modelo probabilístico teórico) para o lançamento de um dado.

O modelo probabilístico para o exemplo 4.1.4 é dado por:

Sexo	Frequência teórica
Feminino	$\frac{5}{8} = 0,625$
Masculino	$\frac{3}{8} = 0,375$
Total	1

Tabela 6: Distribuição de Frequência (modelo probabilístico teórico) para o exemplo 4.1.4.

Em eventos que não são possíveis de serem repetidos, nas mesmas condições, não se aplica a definição de probabilidade frequentista . Veja alguns contra exemplos para esta definição.

Exemplo 4.2.1. *Em um hospital estão fazendo um novo tipo de cirurgia. Aplica-se a cirurgia em um paciente e obseva: sucesso (sobrevive) e fracasso (morre). Como repetir* *esse experimento várias vezes, em condições semelhantes, para calcular a probabilidade de ocorrer sucesso ou fracasso?*

Exemplo 4.2.2. *Resultado do jogo Cruzeiro e Atlético no final do campeonato mineiro. Este experimento é impossível de ser repetido em condições semelhantes à de uma situação ocorrida anteriormente. Um atleta, veja Figura 13, pode querer saber qual a probabilidade dele ganhar ou perder o jogo.*

Figura 13: Um jogo de futebol não pode ser repetido em condições semelhantes.

Exemplo 4.2.3. *Suponha que se queira atribuir uma probabilidade de aceitação de um novo perfume recém-fabricado para mulheres. O químico que produziu o perfume poderá ter atribuído uma probabilidade de seu produto bastante diferente da probabilidade do comerciante junto a um grupo de mulheres que irão comprá-lo.*

Exemplo 4.2.4. *Qual a probabilidade de um enfermo se recuperar completamente.*

Estes últimos exemplos são situações em que não se aplica nenhum dos conceitos (clássico e frequentista) vistos até aqui. Na próxima unidade será discutida a terceira definição de probabilidade, **a probabilidade subjetiva**, que explica como iremos calcular a probabilidade quando se considera a avaliação da crença do observador na ocorrência de um evento.

4.3 Probabilidade Subjetiva - Kolmogorov (1933).

Como já foi citado nos últimos exemplos da seção anterior, **a Probabilidade Subjetiva** é aquela em que cada indivíduo, baseado em informações prévias e na sua própria opinião a respeito do evento de interesse, pode ter uma resposta para a probabilidade desse evento. Assim, a definição subjetiva considera a avaliação da crença do observador na ocorrência de um evento (BATANERO, 2005). Daí surge a pergunta: mas não existe, na definição de probabilidade subjetiva uma regra a ser seguida? Vamos a um exemplo. Qual a probabilidade de chover no próximo fim de semana? Se levar em conta que há duas possibilidades, chove ou não chove, esta probabilidade é de 50%. Mas pode ocorrer que a pessoa interessada em saber esta probabilidade se informou sobre a previsão do tempo e que na cidade onde ela mora a previsão é de tempo nublado e possível chuva para o fim de semana em questão. Assim esta probabilidade passa ser maior que 50%. Para que a teoria construída sobre estas ou outras questões subjetivas (opiniões pessoais) tenha consistência, seja coerente, algumas regras gerais e de comportamentos racionais foram estabelecidas (BATANERO, 2005). Estas regras são baseadas em alguns axiomas que serão apresentados observando o caso em que o espaço amostral é finito e o caso em que é infinito .

4.3.1 Caso Finito

Definição 4.3.1. *Uma função P, definida na álgebra $\mathcal{A}$ de subconjuntos de Ω e com valores em $[0,1]$, é uma probabilidade se satisfaz os Axiomas de Kolmogorov (MAGALHÃES, 2006):*

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

(ii) $0 \leq P(A_j) \leq 1, \ \forall A_j \subset \Omega$.

(iii) Para toda sequência finita de eventos $A_1, A_2, A_3, \ldots, A_n$ *dois a dois disjuntos (mutuamente exclusivos) tem se* $P(\cup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i)$.

A trinca $(\Omega, \mathcal{A}, P)$ é denominada espaço de probabilidade. A função P é definida do seguinte modo:

Se $\Omega = \{w_1, w_2, w_3, \ldots, w_N\}$ finito, defini-se a função $P(w_j) = p_j$ tais que $p_j \geq 0$ e $\sum_{j=1}^{N} p_j = 1$. Assim, a função P é tal que $P(w_1) = p_1$, $P(w_2) = p_2$, $P(w_3) = p_3$, $\ldots$, $P(w_N) = p_N$, sendo $p_1 \geq 0$, $p_2 \geq 0$, $p_3 \geq 0$, $\ldots$, $p_N \geq 0$ e $\sum_{j=1}^{N} p_j = p_1 + p_2 + p_3 + \ldots + p_N = 1$.

Também, dado um evento A qualquer com $A \subset \Omega$, define-se $P(A) = \sum_{j:w_j \in A} [P(w_j)]$, que é a soma das probabilidades dos pontos amostrais de A.

Exemplo 4.3.1. *Suponha o experimento lançamento de um dado qualquer e observar a face voltada para cima. Assim:* $\Omega = \{1, 2, 3, 4, 5, 6\}$. *Podemos ter* $P(1) = 0$, $P(2) = 0,8$, $P(3) = 0,2$, $P(4) = 0$, $P(5) = 0$ *e* $P(6) = 0$. *Logo a probabilidade do evento* $A =$ *"obter um número par no lançamento de um dado"é* $P(2, 4, 6) = P(2) + P(4) + P(6) = 0,8 + 0 + 0 = 0,8$. *Já a probabilidade do evento* $B =$ *"obter um número ímpar no lançamento de um dado"é* $P(1, 3, 5) = P(1) + P(3) + P(5) = 0 + 0,2 + 0 = 0,2$.

Exemplo 4.3.2. *Suponha o experimento lançamento de uma moeda e observa a face voltada para cima. Tem-se* $\Omega = \{K, C\}$. *A função* P *pode ser definida como sendo* $P(K) = p$ *e* $P(C) = 1 - p$, *onde* $0 \leq p \leq 1$. *Análogo, pode-se definir a probabilidade para* $\Omega = \{0, 1\}$, *onde* $1 =$ *"sucesso"* $0 =$ *"$fracasso$" e , daí pode-se definir* $P(1) = p$ *e* $P(0) = 1 - p$, *onde* $0 \leq p \leq 1$.

Exemplo 4.3.3. *Seja* $\Omega = \{w_1, w_2, w_3\}$. P *pode ser definida como sendo:*

$$P(\{w_j\}) = \begin{cases} P(\{w_1\}) = p_1 \\ P(\{w_2\}) = p_2 \\ P(\{w_3\}) = p_3 = 1 - (p_1 + p_2) \end{cases}$$

com $p_1 \geq 0$, $p_2 \geq 0$ *e* $p_3 \geq 0$. *Assim podemos ter:*

$$P(\{w_1\}) = P(\{w_2\}) = P(\{w_3\}) = \frac{1}{3}.$$

ou

$$P(\{w_1\}) = \frac{1}{12},\ P(\{w_2\}) = \frac{3}{12} \quad e \quad P(\{w_3\}) = \frac{8}{12}.$$

Exemplo 4.3.4. *A partir dos axiomas da definição de probabilidade é possível, neste momento, construir um espaço de probabilidade para o exemplo 4.1.5. Um dado é lançado tal que a chance de ocorrer a face de número j, com* $j \in \{1,\ 2,\ 3,\ 4,\ 5,\ 6\}$, *é proporcional a j. Para que se tenha um espaço de probabilidade coerente com esta informação deve-se ter* $\lambda = \frac{p_1}{1} = \frac{p_2}{2} = \frac{p_3}{3} = \frac{p_4}{4} = \frac{p_5}{5} = \frac{p_6}{6}$, *com* $\lambda \in \mathbb{R}$, *sendo* p_j *a probabilidade de j ocorrer. Então* $p_1 = \lambda$, $p_2 = 2\lambda$, $p_3 = 3\lambda$, $p_4 = 4\lambda$, $p_5 = 5\lambda$ *e* $p_6 = 6\lambda$. *Como* $\sum_{j=1}^{6} p_j = 1$, $p_1 + p_2 + p_3 + p_4 + p_5 + p_6 = 1 \Leftrightarrow 21\lambda = 1 \Leftrightarrow \lambda = \frac{1}{21}$. *Portanto:*

$$p_1 = \frac{1}{21},\ p_2 = \frac{2}{21},\ p_3 = \frac{3}{21},\ p_4 = \frac{4}{21},\ p_5 = \frac{5}{21} \ e\ p_6 = \frac{6}{21}.$$

4.3.2 Caso Infinito

O sistema de axiomas (i, ii, iii) da definição de Probabilidade é completamente adequado para tratar questões quando o espaço amostral Ω é finito. Quando Ω é infinito, mesmo que enumerável, o modelo matemático decorrente daquele sistema deixa de servir para a construção de uma teoria que se deseja ser mais eficiente (JAMES, 2004).

Exemplo 4.3.5. *Considere-se o experimento aleatório em que se lança uma moeda até que apareça a "face". Convencionando que o inteiro n significa que a "face"apareceu a 1ª vez no n-ésimo lançamento, o espaço de resultados pode representar-se pela sucessão infinita $\Omega = \{1, 2, 3, 4, 5, \cdots\}$. Considere-se em Ω a classe $\mathcal{F}$ composta pelo conjunto $\emptyset$, Ω, por todos os conjuntos finitos de Ω e pelos respectivos complementares. Esta classe é uma álgebra.*

De fato, o axioma (i) é satisfeito, pois $\Omega \in \mathcal{F}$. O axioma (ii) também é satisfeito, pois todo conjunto tomado na classe seu complementar também está na classe $\mathcal{F}$. O axioma (iii) também é satisfeito, pois a união finita de conjuntos finitoS é finita, então está na classe $\mathcal{F}$.

No entanto pode pensar-se em acontecimentos (subconjuntos de Ω) que não pertencem a $\mathcal{F}$. Por exemplo, o acontecimento "a face sai pela primeira vez num lançamento de ordem ímpar", isto é, $A = \{1, 3, 5, \cdots\}$. Este acontecimento não pertence a $\mathcal{F}$, pois não é finito e nem complementar de conjunto finito. No entanto, A pode exprimir-se como a união infinita de elementos de Ω,

$$A = \bigcup_{k=1}^{\infty} 2k - 1$$

.

Exemplo 4.3.6. *Considere-se o experimento aleatório em que se seleciona, ao acaso, um ponto do intervalo $[0, 1]$. Aqui, $\Omega = [0, 1]$ e a álgebra $\mathcal{A}$ são todos os subconjuntos cujos comprimentos estejam bem definidos. Quem são esses conjuntos? Considerando primeiro uniões finitas de intervalos e seja $\mathcal{A}_0 = \{A \subset [0, 1]$: A é união finita de intervalos $\}$. Notemos que $\mathcal{A}_0$ é álgebra, pois $\Omega \in \mathcal{A}_0$, se $A \in \mathcal{A}_0$ então A^c também é união finita de intervalos, (iii) é trivial. O conjunto $\emptyset$ e o evento elementar $\{\omega\}$, onde $\omega \in [0, 1]$, serão interpretados como intervalos degenerados de comprimento 0, portanto são elementos de $\mathcal{A}_0$.*

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

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

é uma união enumerável de intervalos mas obviamente não é uma união finita, isto é, $A \notin \mathcal{A}_0$.

Os dois exemplos veicula a ideia de que no caso Ω infinito há acontecimentos interessantes que se exprimem pela união infinita de outros acontecimentos ou de acontecimentos elementares. Se o domínio da função de conjunto $P(.)$, probabilidade de um ponto, deve conter tais acontecimentos então em vez de uma álgebra deve optar-se por uma σ-álgebra.

Definição 4.3.2. *Uma função P, definida na σ-álgebra $\mathcal{F}$ de subconjuntos de Ω e com valores em $[0, 1]$, é uma probabilidade se satisfaz os Axiomas de Kolmogorov:*

(i)* $P(\Omega) = 1$.

(ii)* $0 \leq P(A_j) \leq 1,\ \forall A_j \subset \Omega$.

(iii) Para toda sequência infinita de eventos $A_1, A_2, A_3, \ldots$ dois a dois disjuntos (mutuamente exclusivos) tem se* $P(\cup_{i=1}^{\infty} A_i) = \displaystyle\sum_{i=1}^{\infty} P(A_i)$.

A trinca $(\Omega, \mathcal{F}, P)$ é denominada espaço de probabilidade.

Exemplo 4.3.7. *Agora, é possível definir uma probabilidade para o exemplo 1 citado no início da seção 3. Conta-se o número de chamadas telefônicas que chegam a uma central das 12:00h às 14:00h.*

Tem-se $\Omega = \{0,\ 1,\ 2,\ 3,\ \ldots,\ n\} = \mathbb{N}$. *Pode-se definir* $P(\{0\}) = \frac{1}{2}$, $P(\{1\}) = \frac{1}{2^2}$, $P(\{2\}) = \frac{1}{2^3}, \ldots$. *Assim*

$$P(\{0\}) + P(\{1\})) + P(\{2\})\ + \ldots\ = \frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \ldots.$$

Como esta sequência é uma P.G. infinita com $|q| < 1$, então a soma de seus termos é dada por $S = \frac{a_1}{1-q}$. Logo:

$$P(\{0\}) + P(\{1\})) + P(\{2\})\ + \ldots\ = \frac{\frac{1}{2}}{1 - \frac{1}{2}} = 1,$$

odedecendo assim, os três axiomas da definição de probabilidade. A Figura 14 mostra o primeira central de telefone do mundo.

Figura 14: A primeira central telefônica do mundo entrou em funcionamento no dia 25 de janeiro de 1878, em Connecticut, nos Estados Unidos. Central ajudou a popularizar o uso do telefone. Fonte: http://inovabrasil.blogspot. com.br/2010/01/surge-profissao-de-telefonista-em-1878.html.

Este problema é comum nos livros de estatística e, a partir dos axiomas de Kolmogorov, pode-se definir outras probabilidades para a central de telefone. Veja o próximo exemplo.

Exemplo 4.3.8. *Análogo ao exemplo 4.3.7 define-se uma probabilidade para as chamadas telefônicas do seguinte modo: para* $e = 2,718\ldots$ *e* $n \in \mathbb{N}$,

$$
\begin{aligned}
P(\{0\}) &= \frac{e^{-1}}{0!}\\
P(\{1\}) &= \frac{e^{-1}}{1!}\\
P(\{2\}) &= \frac{e^{-1}}{2!}\\
P(\{3\}) &= \frac{e^{-1}}{3!}\\
&\vdots\\
P(\{n\}) &= \frac{e^{-1}}{n!}\\
&\vdots
\end{aligned}
$$

Assim,

$$P(\{0\}) \ + \ P(\{1\}) \ + \ P(\{2\}) \ + \ \ldots \ = \ e^{-1}.(1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \ \ldots\ ).$$

Mas como

$$e^x = \sum_{n=o}^{\infty} \frac{x^n}{n!}, \ \forall x \in \mathbb{R},$$

então

$$e^1 = \left(1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \ \dots \right) \quad e \quad e^{-1} = \left(1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \ \dots \right).$$

Logo

$$P(\{0\}) + P(\{1\}) + P(\{2\}) + \dots \ = \ e^1 \ \times \ e^{-1} = 1,$$

satisfazendo assim, os axiomas i^*), ii^*) e iii^*). Esta probabilidade é conhecida como Probabilidade de Poisson ou probabilidade realista para as chamadas telefônicas.

4.4 Propriedades de uma Probabilidade

Dados Ω, $\mathcal{F}$ e P, denomina-se $(\Omega,\ \mathcal{F},\ P)$ Espaço de Probabilidade. Para o espaço de probabilidade $(\Omega,\ \mathcal{F},\ P)$ vale as seguintes propriedades:

P_1) Seja ϕ o evento impossível. Então $P(\phi) = 0$.

Demonstração:

Considere a sequência de eventos $\Omega,\ \phi,\ \phi,\ \phi,\ \dots$. Esta é uma sequência de eventos dois a dois mutuamente exclusivos, logo obedece ao axioma iii), em que: $P(\cup_{n=1}^{w} A_n) = \displaystyle\sum_{n=1}^{w} P(A_n)$, com $A_i \cap A_j = \phi$. Então tem-se:

$$P(\cup_{n=1}^{w} A_n) = \sum_{n=1}^{w} P(A_n) \Leftrightarrow$$

$$P(\Omega \cup \ \phi \ \cup \ \phi \ \cup \ \phi \ \cup \ \dots) = P(\Omega) + P(\phi) + P(\phi) + P(\phi) + \dots$$

$$P(\Omega) = P(\Omega) + P(\phi) + P(\phi) + P(\phi) + \dots \ = \ P(\Omega) + \sum_{n=1}^{\infty} P(\phi)$$

Pelo axioma ii), tem-se $P(\Omega) = 1$, então $1 = 1 + \displaystyle\sum_{n=1}^{\infty} P(\phi)$, daí, $\displaystyle\sum_{n=1}^{\infty} P(\phi) = 0$. E como $P(\phi) \geq 0$ conclui-se que $P(\phi) = 0$.

P_2) Se $A_1,\ A_2,\ A_3,\ \dots\ A_n$ forem n eventos 2 a 2 mutuamente exclusivos, então:

$$P(\cup_{j=1}^{n} A_j) = \sum_{j=1}^{n} P(A_j)$$

Demonstração:

Sejam $A_1,\ A_2,\ A_3,\ \dots,\ A_n,\ \phi,\ \phi,\ \phi,\ \dots,$ uma sequência de eventos 2 a 2 mutuamente exclusivos. Então:

$P(A_1 \cup A_2 \cup \ldots\ A_n \cup \phi \cup \phi \cup \phi\ \ \ldots) = P(A_1)+P(A_2)+\ldots+P(A_n)+P(\phi)+P(\phi)+\ldots$

Como $P(\phi = 0)$, conclui-se que:

$$P(A_1\ \cup A_2\ \cup\ \ldots\ A_n \cup\ \phi\ \cup\ \phi\ \cup\ \phi\ \ldots) = P(A_1) + P(A_2) + \ldots + P(A_n)$$

Isto significa que a probabilidade da união dos eventos é a soma das probabilidades de cada evento.

P_3) Dados dois eventos A e B, então $P(A - B) = P(A) - P(A \cap B)$.

Demonstração:

Veja na Figura 15 que $(A-B)\cup(A\cap B) = A$, sendo $(A-B)$ e $(A\cap B)$ mutuamente exclusivos.

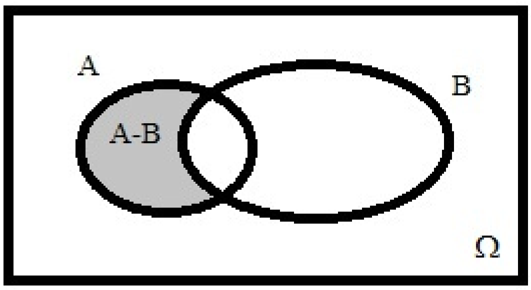

Figura 15: Diferença entre dois eventos $A - B$. Fonte: Próprio autor.

Então:

$$P\left[(A - B) \cup (A \cap B)\right] = P(A) \Leftrightarrow P(A - B) + P(A \cap B) = P(A)$$

Assim, conclui-se que: $P(A - B) = P(A) - P(A \cap B)$.

P_4) Suponha que os eventos A e B são tais que $B \subset A$. Então:

$$P(A - B) = P(A) - P(B) \quad e \quad P(B) \leq P(A)$$

Demonstração:

Por P_3 , $P(A - B) = P(A) - P(A \cap B)$. Como $P(A \cap B) = B$ (veja na Figura 16), então:

$$P(A - B) = P(A) - P(B)$$

e como $P(A-B) \geq 0$ tem-se

$$P(A) - P(B) \geq 0,$$

daí conclui-se que:

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

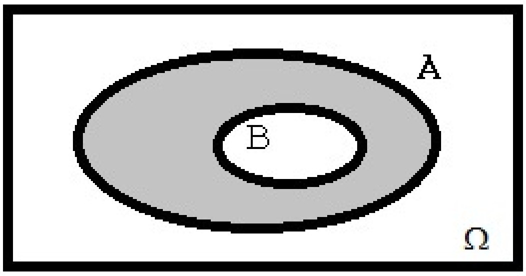

Figura 16: Diferença entre dois eventos $A - B$. Fonte: Próprio autor.

P_5) Para todo evento A tem-se $P(A^c) = 1 - P(A)$, sendo A^c o evento complementar de A.

Demonstração:

Como A^c e A são mutuamente exclusivos tem-se $P(A \cup A^c) = P(A) + P(A^c)$. Mas,

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

Assim conclui-se que $P(A) + P(A^c) = 1$, ou ainda, $P(A^c) = 1 - P(A)$.

P_6) Dados os eventos quaisquer A e B, então:

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

Demonstração:

Como A e $(B-A)$ são mutuamente exclusivos e $A \cup B = A \cup (B-A)$, veja ANEXO I, tem-se:

$$P(A \cup B) = P\left[(A \cup (B-A)\right] \Leftrightarrow P(A \cup B) = P(A) + P(B-A)$$

Por P_3 tem-se:

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

Daí, substituindo, temos:

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

donde conclui-se que:

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

Da mesma forma, sabendo que A e B podem ser escritos como uma reunião de dois eventos mutuamente exclusivos, então:

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

e

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

Logo,

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

Outra demonstração seria usando a propriedade (veja ANEXO I),

$$|A \cup B| = |A| + |B| - |A \cap B|.$$

Dividindo-se ambos os membros por $|\Omega|$ tem-se:

$$\frac{|A \cup B|}{|\Omega|} = \frac{|A|}{|\Omega|} + \frac{|B|}{|\Omega|} - \frac{|A \cap B|}{|\Omega|}$$

que resulta em:

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

P_7) Dados os eventos quaisquer A, B e C, então:

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

Demonstração:

Por P_6 tem-se:

$$P(A \cup B \cup C) = P((A \cup B) \cup C)$$

$$= P(A \cup B) + P(C) - P\left[(A \cup B) \cap C\right]$$

Mas,

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C).$$

Daí, tem-se:

$$P(A \cup B \cup C) = P((A \cup B) \cup C) = P(A \cup B) + P(C) - P\left[(A \cap C) \cup (B \cap C)\right]$$

$$= P(A) + P(B) - P(A \cap B) + P(C) - P(A \cap C) - P(B \cap C) + P\left[(A \cap C) \cap (B \cap C)\right]$$

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

Assim conclui-se que:

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

P_8) Suponha n eventos quaisquer $A_1,\ A_2,\ A_3,\ \ldots\ ,\ A_n$. Então:

$$\begin{aligned} P\left(\bigcup_{j=1}^{n} A_j\right) &= \sum_{j=1}^{n} P(A_j) \\ &- \sum_{(1 \leq j_1 < j_2 \leq n)}^{n} P\left[A_{j_1} \cap A_{j_2}\right] \\ &+ \sum_{1 \leq j_1 < j_2 < j_3 \leq n}^{n} P\left[A_{j_1} \cap A_{j_2} \cap A_{j_3}\right] \\ &- \sum_{1 \leq j_1 < j_2 < j_3 < j_4 \leq n}^{n} P\left[A_{j_1} \cap A_{j_2} \cap A_{j_3} \cap A_{j_4}\right] \\ &+\ \ldots\ +\ (-1)^{(n-1)} \sum P\left[A_{j_1} \cap A_{j_2} \cap A_{j_3} \cap A_{j_4} \cap \ldots \cap A_{j_n}\right]. \end{aligned}$$

Veja a demonstração desta propriedade em (MORGADO, et al., 2006, p. 191).

4.4.1 Exemplos de aplicação.

Os exemplos seguintes relaciona-se situações em que o espaço amostral poderá ser finito ou infinito enumerável.

Exemplo 4.4.1. *Num prédio residencial há 2 blocos I e II. No bloco I, há 80 apartamentos, dos quais 15% estão em atraso com o condomínio. No bloco II, há 50 apartamentos, 10% dos quais com taxas atrasadas. As fichas de todos os moradores estão reunidas, e uma delas é escolhida ao acaso. Pergunta-se:*

a) *Qual é probabilidade de que a ficha escolhida seja do bloco I e esteja quite com o condomínio?*

b) *Qual é probabilidade de que a ficha escolhida seja do bloco I ou esteja quite com o condomínio?*

Para responder a essas perguntas, primeiro vamos analisar o problema a partir do seguinte quadro.

	I	II	Totais
Em atraso	15% de 80 = 12	10% de 50 = 5	17
Quite com o condomínio	68	45	113
Totais	80	50	130

Tabela 7: As probabilidades: exemplo 4.4.1. Fonte: dados hipotéticos

Para o **item (*a*)**, *a probabilidade de que a ficha escolhida seja do bloco I* **e** *esteja com o condomínio em atraso é dada pela* **interseção** *desses eventos, A = "ser do bloco I"e Q = "esteja quite com o condomínio", ou seja,* $P(A \cap Q) = \frac{68}{130}$. *Note que há 68 condôminos que moram no bloco I e estão quites com seus condomínios.*

Para o **item (*b*)**, *a probabilidade de que a ficha escolhida seja do bloco I* **ou** *esteja quite com seus condomínios é dada pela* **união** *desses eventos. Assim, pela propriedade* P_6, *tem-se:*

$$P(A \cup Q) = P(A) + P(Q) - P(A \cap Q)$$
$$P(A \cup Q) = \frac{80}{130} + \frac{113}{130} - \frac{68}{130}$$
$$P(A \cup Q) = \frac{125}{130} \cong 96,15\%$$

Exemplo 4.4.2. *(DANTAS, 2004) Uma moeda balanceada é lançada dez vezes. Sabe-se que o espaço amostral desse experimento tem* 2^{10} *pontos amostrais. Para calcular a probabilidade de ocorrência de cara (K) em pelo menos um dos lançamentos de número 2, 4, 6 e 8, denota-se por* A_{2k} *o evento cara ocorra no lançamento 2k, para* $k \in \{1, 2, 3, 4\}$. *Assim, o evento de interesse é:* $\bigcup_{k=1}^{4} A_{2k}$. *Note que o evento* $\bigcup_{j=1}^{4} A_{2k}^c$ *corresponde a ocorrência de coroa nos lançamentos 2, 4, 6 e 8. O número de pontos amostrais desse evento é* 2^6, *pois, fixa-se coroa nesses quatro lançamentos e nos outros seis tem-se duas possibilidades em cada um deles. Assim, da propriedade 5, tem-se:*

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

Daí decorre que:

$$\bigcup_{j=1}^{4} A_{2k} = 1 - \bigcup_{k=1}^{4} A_{2k}^{c} = 1 - \frac{2^6}{2^{10}} = 0,0625$$

Exemplo 4.4.3. *Seja uma urna contendo três bolas distintas, sendo duas bolas brancas (b_1 e b_2) e 1 vermelha (v). Seleciona-se, ao acaso, duas bolas, uma de cada vez e* **com reposição**.

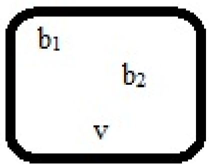

Figura 17: Urna com 2 bolas brancas e 1 vermelhra. Fonte: Próprio autor.

Considere os eventos:

- B_1 = *"a primeira bola selecionada é branca"'.*
- B_2 = *"a segunda bola selecionada é branca"'.*
- V_1 = *"a primeira bola selecionada é vermelha"'.*
- V_2 = *"a segunda bola selecionada é vermelha"'.*

A partir destes dados é possível construir um modelo de probabilidade para este experimento. O espaço amostral é dado por:

$$\Omega = \left\{ \begin{array}{ccc} b_1b_1 & b_2b_2 & vv \\ b_1b_2 & b_2b_1 & vb_1 \\ b_1v & b_2v & vb_2 \end{array} \right\}.$$

Obs. *Aqui foi definido que b_1 é a bola branca 1 e B_1 é o evento que consiste em retirar a primeira bola e esta ser branca. Portanto $b_1 \neq B_1$.*

Como as bolas são idênticas a probabilidade de cada ponto amostral é a mesma e igual a $\frac{1}{N}$, onde $N = |\Omega|$. Logo, para o evento B_1 tem-se:

$$P(B_1) = \frac{n}{N},$$

onde n é o número de casos favoráves a B_1. *Este é um modelo de probabilidade. Se o interesse fosse somente a ocorrência de* B_1 *e* V_1 *em uma única retirada, teria se* $\Omega = \{b_1, b_2, v\}$. *Daí,* $P(B_1) = \frac{2}{3}$ *e* $P(V_1) = P(B_1^c) = \frac{1}{3}$. *Note que* $B_1 = V_1^c$ *e* $V_1 = B_1^c$. *Mas como o interesse é considerar duas retiradas, uma após a outra, pelo princípio multiplicativo, tem-se* $|\Omega| = 3 \times 3 = 9$, *daí decorre que:*

$$|B_1| = 2 \times 3 = 6 \Rightarrow P(B_1) = \frac{6}{9} = \frac{2}{3}$$

$$|V_1| = 1 \times 3 = 3 \Rightarrow P(V_1) = \frac{3}{9} = \frac{1}{3}$$

$$|B_1 \cap V_2| = 2 \times 1 = 2 \Rightarrow P(B_1 \cap V_2) = \frac{2}{9}$$

$$|B_1 \cap B_2| = 2 \times 2 = 4 \Rightarrow P(B_1 \cap B_2) = \frac{4}{9}$$

$$|V_1 \cap B_2| = 1 \times 2 = 2 \Rightarrow P(V_1 \cap B_2) = \frac{2}{9}$$

$$|V_2| = 3 \times 1 = 3 \Rightarrow P(V_1) = \frac{3}{9} = \frac{1}{3}$$

$$|B_2| = 3 \times 2 = 6 \Rightarrow P(B_2) = \frac{6}{9} = \frac{2}{3}$$

Agora, pode-se fazer a análise para este mesmo problema, considerando que a primeira bola retirada não será colocada de volta na caixa, **sem reposição**. *O espaço amostral passa a ser:*

$$\Omega = \left\{ \begin{array}{ccc} b_1b_2 & b_2b_1 & vb_1 \\ b_1v & b_2v & vb_2 \end{array} \right\}.$$

Assim $|\Omega| = 3 \times 2 = 6$, *daí, decorre que:*

$$|B_1| = 2 \times 2 = 4 \Rightarrow P(B_1) = \frac{4}{6} = \frac{2}{3}$$

$$|B_2| = (2 \times 1) + (1 \times 2) = 4 \Rightarrow P(B_2) = \frac{4}{6} = \frac{2}{3}$$

Observe que $P(B_1) = P(B_2)$, *ou seja, a probabilidade de selecionar uma bola branca na primeira tentativa é a mesma de sair bola branca na segunda tentativa. Analogamente para* V_1 *e* V_2 *tem-se* $P(V_1) = P(V_2) = \frac{1}{3}$. *Também tem-se:*

$$|B_1 \cap B_2| = 2 \times 1 = 2 \Rightarrow P(B_1 \cap B_2) = \frac{2}{6} = \frac{1}{3}$$

$$|B_1 \cap V_2| = 2 \times 1 = 2 \Rightarrow P(B_1 \cap B_2) = \frac{2}{6} = \frac{1}{3}$$

$$|V_1 \cap B_2| = 1 \times 2 = 2 \Rightarrow P(B_1 \cap B_2) = \frac{2}{6} = \frac{1}{3}$$

Pode-se também mostrar que, em três retiradas, **sem reposição***, a probabilidade de sair bola branca em qualquer posição é sempre a mesma e igual a* $\frac{2}{3}$*. O espaço amostral, neste caso, é*

$$\Omega = \left\{ \begin{array}{ccc} b_1b_2v & b_2b_1v & vb_1b_2 \\ b_1vb_2 & b_2vb_1 & vb_2b_1 \end{array} \right\}.$$

Daí, decorre que:

$$P(B_1) = 2 \times 2 \times 1 = 4$$

$$P(B_2) = 2 \times 2 \times 1 = 4$$

$$P(B_3) = 1 \times 2 \times 2 = 4$$

Logo,

$$P(B_1) = P(B_2) = P(B_3) = \frac{4}{6} = \frac{2}{3}.$$

Finalmente, considere agora três retiradas, **com reposição***. Assim tem-se:* $|\Omega| = 3 \times 3 \times 3 = 27$*. Daí, decorre que:*

$$|B_1| = 2 \times 3 \times 3 = 18 \Rightarrow P(B_1) = \frac{18}{27} = \frac{2}{3}$$

$$|B_2| = 3 \times 2 \times 3 = 18 \Rightarrow P(B_2) = \frac{18}{27} = \frac{2}{3}$$

$$|B_3| = 3 \times 3 \times 2 = 18 \Rightarrow P(B_1) = \frac{18}{27} = \frac{2}{3}$$

Exemplo 4.4.4. *Suponha uma população de n elementos* $\{a_1, a_2, a_3, \ldots, a_n\}, n \in \mathbb{N}$*. Selecione r indivíduos, ao acaso,* **com reposição***, um de cada vez, sendo* $r \leq n$*. Considere o evento A= "um elemento fixado aparece na amostra pelo menos uma vez". Determinar P(A). Em cada seleção há n opções de escolha, já que a seleção será feita com reposição do elemento observado. Se for feita r seleções, o número de pontos do espaço amostral desse experimento é, pelo princípio multiplicativo, dado por* $|\Omega| = n^r$*. Para calcular P(A) tem-se que obter número de casos favoráveis a A. Assim:*

$$P(A) = \frac{?}{n^r}.$$

O evento complementar de A é A^c *= "o elemento fixado não aparece na amostra". Daí tem-se* $|A^c| = (n-1)^r$*. Assim,*

$$P(A) = 1 - P(A^c) = 1 - \frac{(n-1)^r}{n^r} = 1 - \left(1 - \frac{1}{n}\right)^r.$$

Note que para uma população grande, ou seja, $n \longrightarrow \infty$*, tem-se* $P(A) = 0$*. Do contrário, se fixado n e fazendo um número muito grande de escolha, ou seja,* $r \longrightarrow \infty$*, tem-se*

$(n-1)^r/n^r \longrightarrow 0$, daí $P(A)=1$. Logo, para n fixo e r muito grande a probabilidade de encontrar a_1 é de 100%, isto é claro, pois n fixo e aumentando o número de escolha, em algum momento a_1 aparecerá. Note também que se $n=r$, tem-se

$$P(A) = 1 - \left(1 - \frac{1}{n}\right)^n$$

e fazendo $n \longrightarrow \infty$, tem-se

$$P(A) = 1 - e^{-1} = 1 - 0,367879\ldots = 0,632121\ldots \ .$$

Isto significa que numa população de 5 elementos, a probabilidade de que um determinado elemento apareça em uma amostra de tamanho 5, sendo feita 5 seleções com reposição, é de aproximadamente 63,21%.

Exemplo 4.4.5. *Suponha uma população de n elementos $\{a_1, a_2, a_3, \ldots, a_n\}$, $n \in \mathbb{N}$. Selecione r indivíduos, ao acaso,* **com reposição**, *um de cada vez, sendo $r \leq n$. Considere o evento A="não há elementos repetidos na amostra". Para calcular P(A) tem-se:*

$$P(A) = \frac{(n^o \ de \ casos \ favoráveis \ a \ A)}{(n^o \ de \ casos \ possíveis)}$$

$$P(A) = \frac{n.(n-1).(n-2). \ \ldots \ .(n-(r-1))}{n^r}.$$

Assim:

$$\begin{aligned} P(A) &= \frac{n}{n} \times \frac{n-1}{n} \times \frac{n-2}{n} \ \ldots \frac{n-(r-1)}{n} \\ &= \left(1 - \frac{1}{n}\right) \times \left(1 - \frac{2}{n}\right) \ \ldots \ \left(1 - \frac{(r-1)}{n}\right). \end{aligned}$$

Note que se $x \longrightarrow \infty$, as funções $f(x) = 1-x$ e $g(x) = e^{-x}$ se aproximam, ou seja,

$$x \cong 0 \Longrightarrow 1 - x \cong e^{-x}.$$

Assim pode-se dizer que para r razoavelmente menor que n tem-se:

$$1 - \frac{1}{n} \cong e^{\frac{-1}{n}}$$

$$1 - \frac{2}{n} \cong e^{\frac{-2}{n}}$$

$\vdots$

$$1 - \frac{(r-1)}{n} \cong e^{\frac{-(r-1)}{n}}.$$

Daí,

$$P(A) = (1 - \frac{1}{n}) \times (1 - \frac{2}{n}) \times \ \ldots \times (1 - \frac{(r-1)}{n})$$

$$
\begin{aligned}
P(A) &= e^{\frac{-1}{n}} \times e^{\frac{-2}{n}} \times \ldots \times e^{\frac{-(r-1)}{n}} \\
&= e^{\frac{-1}{n} \times (1+2+3+ \ \ldots \ +(r-1))} \\
&= e^{\frac{-1}{n} \times \frac{r.(r-1)}{2}} \\
&= e^{\frac{-r.(r-1)}{2n}}, \quad com \ n \in \mathbb{N}.
\end{aligned}
$$

Uma aplicação deste exemplo

Em um grupo de r pessoas, qual a probabilidade de haver pelo menos duas que fazem aniversário no mesmo dia?

Seja o evento A = "aparecer pelo menos duas pessoas que fazem aniversário no mesmo dia", então A^c = "não haver pessoas que fazem aniversário no mesmo dia".

Assim, tome $n = 365$ dias, e da propriedade do complementar, tem-se:

$$
\begin{aligned}
P(A) = 1 - P(A^c) &= 1 - e^{\frac{-r.(r-1)}{2.365}} \\
&= 1 - e^{\frac{-r.(r-1)}{730}}.
\end{aligned}
$$

Assim, variando r obtemos a Tabela 8:

r	P(A)
1	0,0000
2	0,0027
3	0,0082
4	0,0163
5	0,0270
6	0,0403
7	0,0559
8	0,0738
9	0,0939
10	0,1160
20	0,4058
30	0,6963
60	0,9922
80	0,9998

Tabela 8: Probabilidade de haver pelo menos duas pessoas que fazem aniversário no mesmo dia em um grupo de r pessoas

Pela tabela verifica-se que P(A) aumenta surpreendentemente. Note que para $r = 1$, é óbvio que $P(A) = 0\%$. Mas para $r > 30$ esta probabilidade já é bem grande, maior que $69,63\%$. Agora, em um grupo de 80 pessoas, é quase certo que sempre encontrará duas delas que fazem aniversário no mesmo dia, com probabilidade de não ocorrer igual a $0,02\%$.

Exemplo 4.4.6. *Suponha uma população com 10 elementos* $\{a_1, a_2, a_3, \cdots, a_{10}\}$. *Selecione 4 destes elementos, um de cada vez, ao acaso,* **sem reposição**. *Seja o evento A = "a_4 e a_5 estar sempre presente". A probabilidade de A é $\frac{2}{15}$. De fato, o número de elementos do espaço amostral deste experimento é A_{10}^4, pois a escolha é feita retirando-se um elemento de cada vez e* **sem reposição**. *Por outro lado o número de casos favoráveis a A é $A_4^2 \times A_8^2$, pois, deve-se calcular o número de agrupamentos em que aparece a_4 e a_5, que é A_4^2 e estas aparecem em número de A_8^2, pois deve-se escolher dois dos 8 elementos restante do conjunto. Assim:*

$$P(A) = \frac{A_4^2 \times A_8^2}{A_{10}^4} = \frac{672}{5040} = \frac{2}{15}.$$

Exemplo 4.4.7. *(Generalização do exemplo anterior) Suponha uma população com N elementos* $\{a_1, a_2, a_3, \ldots, a_N\}, N \in \mathbb{N}$. *Seleciona-se r elementos, um de cada vez, ao acaso,* **sem reposição**. *Seja o evento A = "n elementos, $n < r$, fixados na amostra". Calcule P(A).*

O espaço amostral desse experimento é A_N^r, pois a escolha é feita retirando-se um elemento de cada vez e sem reposição. Em um conjunto de r elementos escolhe-se n elementos fixados o que resulta em A_r^n e dos $(N-n)$ restantes escolhe-se $(r-n)$, ou seja, A_{N-N}^{r-n}. Então,

$$|A| = A_r^n \times A_{N-n}^{r-n} = n! \times \binom{r}{n} \times (r-n)! \times \binom{N-n}{r-n} = r! \times \binom{N-n}{r-n}.$$

Portanto:

$$P(A) = \frac{r! \times \binom{N-n}{r-n}}{(A_N^r)} = \frac{r! \times \binom{N-n}{r-n}}{r! \times \binom{N}{r}} = \frac{\binom{N-n}{r-n}}{\binom{N}{r}}.$$

Veja que, para o exemplo 6, tem-se $N = 10, r = 4$ e $n = 2$, então:

$$P(A) = \frac{\binom{10-2}{4-2}}{\binom{10}{4}} = \frac{28}{210} = \frac{2}{15}.$$

Exemplo 4.4.8. *(Problema dos reencontros) Dado uma urna com n bolas numeradas* $\{1, 2, 3, \ldots, n\}$. *Selecionemos n bolas da urna, uma de cada vez, ao acaso,* **sem reposição**. *Dizemos que há um reencontro na seleção j-ésima se a bola de número j for selecionada na seleção j (por exemplo, se a bola 8 for selecionada na oitava seleção, houve um reencontro). Seja $A_j =$ "ocorre um reencontro na seleção j", $j \in \{1, 2, 3, \ldots, n\}$. Calcule e probabilidade de haver pelos menos um reencontro. A probabilidade de haver pelo menos um reencontro é a probabilidade da união dos A_j eventos, ou seja, $P(A_1 \cup A_2 \cup \ldots \cup A_j)$. Então:*

$$P\left(\bigcup_{j=1}^{n} A_j\right) = \sum_{j=1}^{n} P(A_j) - \sum_{1 \leq j_1 < j_2 < n}^{n} P\left(A_{j_1} \cap A_{j_2}\right)$$

$$+ \sum_{1 \leq j_1 < j_2 < j_3 < n}^{n} P\left(A_{j_1} \cap A_{j_2} \cap A_{j_3}\right) + \ldots + (-1)^{n-1} P\left(A_1 \cap A_2 \ldots A_n\right)$$

Como:

$$P(A_j) = \frac{(n-1)!}{n!} = \frac{1}{n}$$

$$P(A_j) = \frac{(n-2)!}{n!} = \frac{1}{n(n-1)}$$

$$P(A_j) = \frac{(n-3)!}{n!} = \frac{1}{n(n-1)(n-2)}$$

Daí,

$$P\left(\bigcup_{j=1}^{n} A_j\right) = \sum_{j=1}^{n} \frac{1}{n} - \sum_{1 \leq j_1 < j_2 < n}^{n} \frac{1}{n(n-1)} + \sum_{1 \leq j_1 < j_2 < j_3 < n}^{n} \frac{1}{n(n-1)(n-2)} - \ldots + (-1)^{n-1} . \frac{1}{n!}$$

$$P\left(\bigcup_{j=1}^{n} A_j\right) = \frac{\begin{pmatrix} n \\ 1 \end{pmatrix}}{n} - \frac{\begin{pmatrix} n \\ 2 \end{pmatrix}}{n(n-1)} + \frac{\begin{pmatrix} n \\ 3 \end{pmatrix}}{n(n-1)(n-2)} - \ldots + (-1)^{n-1} \frac{1}{n!}$$

$$= 1 - \frac{1}{2!} + \frac{1}{3!} - \ldots + (-1)^{n-1} \frac{1}{n!}$$

$$= 1 - \left[\frac{1}{2!} - \frac{1}{3!} - \ldots + (-1)^{n-1} \frac{1}{n!}\right]$$

Como $e^x = \sum_{n=o}^{w} \frac{x^n}{n!}, \forall x \in \mathbb{R}$, *tem-se:*

$$e^{-1} = \sum_{n=o}^{w} \frac{(-1)^n}{n!} = 1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \ldots + (-1)^n . \frac{1}{n!}.$$

Assim, para $n \longrightarrow \infty$ *tem-se:*

$$P\left(\bigcup_{j=1}^{n} A_j\right) = 1 - e^{-1}.$$

Note que a probabilidade de não ocorrer reencontro é dada por $P\left(\bigcup_{j=1}^{n} A_j\right)^c = e^{-1}$. *Os casos favoráveis a este evento são conhecidos por permutações caóticas (casos em que nenhum elemento ocupa a posição inicial), (MORGADO, et al., 2006). Veja que para* $n = 2$, *teríamos* $\{1, 2, 3\}$, *daí as permutações são*

$$\{123, 132, 231, 213, 321, 312\}.$$

Há reencontro somente em

$$\{123,\ 132,\ 213\ e\ 321\},$$

daí:

$$P(A) = \frac{4}{6} \cong 66,67\%,$$

que é igual a

$$1 - \frac{1}{2} + \frac{1}{6} - \frac{2}{3} \cong 66,67\%.$$

Fazendo variar n, obtemos a Tabela 9:

n	P(A)
1	1,0000
2	0,5000
3	0,6667
4	0,6250
5	0,6330
6	0,6314
7	0,6321

Tabela 9: Probabilidade de haver pelos menos um reencontro em um conjunto com n elementos.

4.5 Probabilidade Condicional (Bayes)

Considerando novamente o exemplo 4.4.1, e suponha agora que se queira calcular a probabilidade de que a ficha sorteada seja de um condômino em atraso, com a condição de que, antes do sorteio, já se sabe que a ficha é de um condômino do bloco II. Veja a Tabela 10.

	Bloco I	Bloco II	Totais
Em atraso	12	5	17
Quite com o condomínio	68	45	113
Totais	80	50	130

Tabela 10: Dados do exemplo 4.4.1. Fonte: Dados hipotéticos.

Assim, a probabilidade desejada é igual a $\frac{5}{50}$, pois há 5 casos favoráveis (condôminos do bloco II que estão em atraso) e 50 casos possíveis (condôminos do bloco II). Seja:
A = "estar em atraso"e
B = "ser do bloco II".
A probabilidade desejada é denominada probabilidade condicional de A dado B e representada por $P(A|B)$.

Definição 4.5.1. *Para dois eventos A e B, sendo $P(B) > 0$, define-se* **a probabilidade condicional de A dado B**, *como sendo:*

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

No exemplo citado acima, tem-se $A \cap B$ = "estar em atraso e ser do bloco II". Daí $P(A \cap B) = \frac{5}{130}$ *e* $P(B) = \frac{50}{130}$. Assim,

$$P(A|B) = \frac{\frac{5}{130}}{\frac{50}{130}} = \frac{5}{50},$$

como já havia obtido no cálculo anterior.

A probabilidade $P(A)$ é a probabilidade "*a priori*"de A, e com a informação adicional de que B ocorreu, obtém-se a probabilidade "*a posteriori*" $P(A|B)$. No exemplo $P(A) = \frac{17}{130}$ e $P(A|B) = \frac{5}{50}$. Note que, neste caso, $P(A|B) > P(A)$, o que justifica que a informação de que B ocorreu fez aumentar as chances de A ocorrer. Da definição decorre que: $P(A \cap B) = P(A|B) \times P(B)$, se $P(B) > 0$ ou ainda $P(A \cap B) = P(B|A) \times P(A)$, se $P(A) > 0$, que é conhecida como regra do produto. Isto significa que, para se avaliar a probabilidade de ocorrerem dois (ou mais) eventos simultâneos (ou sucessivos), que é $P(A \cap B)$, basta multiplicar a probabilidade de ocorrer um deles, neste caso $P(B)$, pela probabilidade de ocorrer o outro, sabendo que o primeiro já ocorreu, que é dada por $P(A|B)$. Veja a aplicação desta propriedade nos exemplos seguintes.

Exemplo 4.5.1. *Numa empresa há 10 departamentos, 7 responsáveis pelas vendas de* *um produto X e 3 responsáveis pelas vendas de um produto Y. A probabilidade de um produto X ser vendido para um cliente desconhecido é de* 80% *e, para o produto Y, essa probabilidade é de 60%. Um cliente desconhecido chega a esta empresa a fim de comprar um dos produtos, e dirige-se, ao acaso, a um dos departamentos. Qual a probabilidade deste departamento não conseguir vender o produto para este cliente? Note que para o produto não ser vendido, o cliente tem que escolher um dos (X ou Y) e em seguida não comprar. Assim, devemos calcular a probabilidade de o cliente não comprar o produto dado que ele necessita do produto X ou a probabilidade de o cliente não comprar o produto dado que ele necessita do produto Y, ou seja:*

$$P(\textit{não compra} \mid \textit{necessita de } X) \textit{ ou } P(\textit{não compra} \mid \textit{necessita de } Y).$$

Mas:

$$\begin{aligned} P(\textit{não compra e necessita de } X) &= P(\textit{não compra} \mid \textit{necessita de } X) \times P(\textit{necessita de } X) \\ &= \frac{2}{10} \times \frac{7}{10} = 14\%, \quad e \\ P(\textit{não compra e necessita de } Y) &= P(\textit{não compra} \mid \textit{necessita de } Y) \times P(\textit{necessita de } Y) \\ &= \frac{4}{10} \times \frac{3}{10} = 12\%. \end{aligned}$$

Logo, a probabilidade do departamento não conseguir vender o produto para o cliente é de 26%. Problemas como este são melhores visualizados pelo diagrama abaixo:

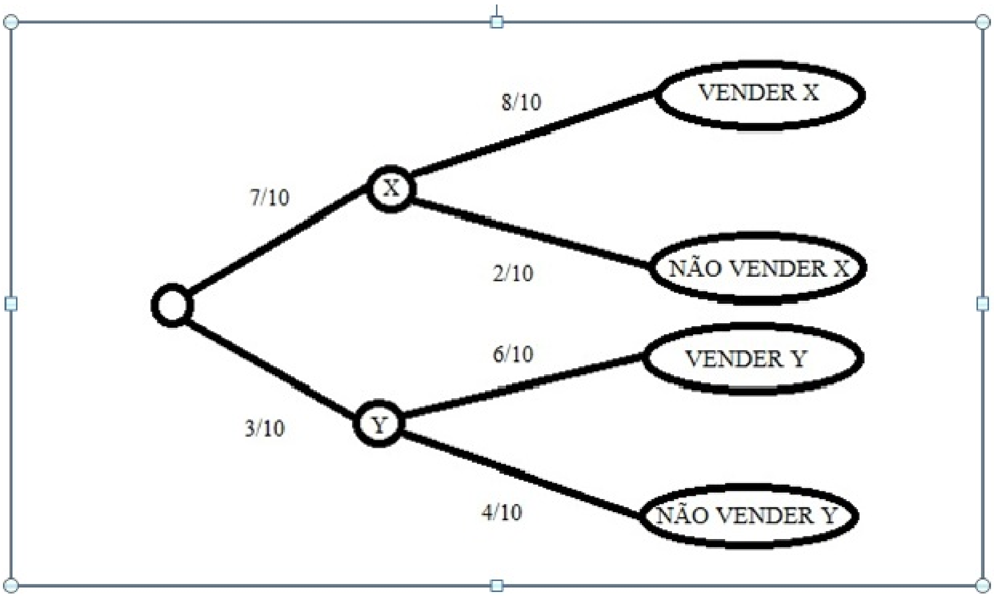

Figura 18: Diagrama das probabilidades condicionais.

Note que:

- *P(vender X no depto. X)=P(vender X|X)*$\times P(X) = \frac{8}{10} \times \frac{7}{10} = 56\%$,

- *P(não vender X no depto. X)=P(não vender X|X)*$\times P(X) = \frac{2}{10} \times \frac{7}{10} = 14\%$,

- *P(vender Y nødepto. Y)=P(vender Y|Y)*$\times P(Y) = \frac{6}{10} \times \frac{3}{10} = 18\%$,

- *P(não vender Y no depto. Y)=P(não vender Y|Y)*$\times P(Y) = \frac{4}{10} \times \frac{3}{10} = 12\%$ *e*

- *P(vender X|X)+P(não vender X|X)+P(vender Y|Y)+P(não vender Y|Y)*$= 1$.

4.5.1 Propriedades

i) $P(\phi|A) = 0$;

ii)) $P(\Omega|A) = 1$;

iii) $0 \leq P(B|A) \leq 1$;

iv) $P\left[(B \cup C)|A\right] = P(B|A) + P(C|A)$ *se* $B \cap C = \phi$;

v) $P(A_1 \cap A_2 \cap A_3) = P\left[A_3|(A_1 \cap A_2)\right].P(A_2|A_1).P(A_1)$.

Demonstração:

i) $P(\phi|A) = \frac{P(\phi \cap A)}{P(A)} = \frac{P(\phi)}{P(A)} = \frac{0}{P(A)} = 0$.

ii) $P(\Omega|A) = \frac{P(\Omega \cap A)}{P(A)} = \frac{P(A)}{P(A)} = 1$.

iii) Como $0 \leq P(A \cap B) \leq P(A)$, tem-se $0 \leq \frac{P(A \cap B)}{P(A)} \leq 1$, logo: $0 \leq P(B|A) \leq 1$.

iv) $P\left[(B \cup C)|A\right] = \frac{P[(B \cup C) \cap A)]}{P(A)} = \frac{P[(B \cap A)\ \cup\ (C \cap A)]}{P(A)}$

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

v) Temos que $P(A_1 \cap A_2 \cap A_3) = P(A_3|(A_1 \cap A_2).P(A_1 \cap A_2)$, mas como, por definição, tem-se $P(A_1 \cap A_2) = P(A_2|A_1).P(A_1)$, segue o resultado.

O item v) é conhecido como **o teorema do produto**, e vale para n eventos. Assim, para os eventos $\{A_1,\ A_2,\ A_3,\ \ldots\ ,\ A_n\}$ com $P(\cap_{i=1}^{n} A_i) > 0$, a regra do produto de probabilidades é dada por:

$$P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_1) \times P(A_2|A_1) \times P\left[A_3|(A_1 \cap A_2)\right] \times \ \ldots$$

$$\times P\left[A_n|(A_1 \cap A_2 \cap A_3 \times \ldots \times \cap A_{n-1})\right].$$

A demonstração será feita por indução. Note que, por hipótese, todos os condicionamentos do lado direito são feitos em eventos com probabilidade positiva, pois todos eles contém $\bigcap_{i=1}^{n} A_i$. Para $n = 2$, pela definição de probabilidade condicional, tem-se: $P(A_2|A_1) = \frac{P(A_1 \cap A_2)}{P(A_1)} \Rightarrow P(A_1 \cap A_2) = P(A_1).P(A_2|A_1)$, pois $P(A_1) > 0$. Suponha o resultado válido para $n = k$, então: $P(A_1 \cap A_2 \cup \ \ldots \cap A_k) = P(A_1).P(A_2|A_1).P(A_3|(A_1 \cap A_2) \ldots P\left[A_k|(A_1 \cap A_2 \cap A_3 \ldots A_{k-1})\right]$. Quer-se provar para $n = k + 1$. Assim,

$$P(A_1 \cap A_2 \cap \ \ldots \cap A_k \cap A_{k+1}) =$$

$$P[(A_1 \cap A_2 \cap \ldots \ \cap A_k) \cap (A_{k+1})] =$$

$$P(A_1 \cap A_2 \cap \ldots \ \cap A_k).P[A_{k+1}|(A_1 \cap A_2 \cap \ldots \cap A_k)].$$

e por hipótese de indução tem-se:

$$P(A_1 \cap A_2 \cap \ \ldots \cap A_k \cap A_{k+1}) =$$

$$P(A_1).P(A_2|A_1). \ldots \ .P(A_k|(A_1 \cap A_2 \cap A_3 \ldots \cap A_{k-1}).P[A_{k+1}|(A_1 \cap A_2 \cap \ldots \cap A_k)]$$

Mostrando assim que a afirmativa é verdadeira.

Como a probabilidade condicional é uma probabilidade, ou seja, a definição de probabilidade condicional satisfaz os axiomas i), ii) e iii) da definição de probabilidade axiomática, então, sejam A, B e C eventos de Ω, valem as propriedades (DANTAS, 2006):

- $P\left[(B \cup C)|A\right] = P(B|A) + P(C|A) - P\left[(B \cap C)|A\right];$
- $P(B^c|A) = 1 - P(B|A);$
- $P\left[(B - C)|A\right] = P(B|A) - P\left[(B \cap C)|A)\right].$

Exemplo 4.5.2. *Há três envelopes idênticos sobre uma mesa. Sabe-se que em um deles contém uma nota de* $R\$100,00$, *em outro contém uma nota de* $R\$5,00$ *e o terceiro contém uma nota de* $R\$2,00$. *Três pessoas, cada uma, na ordem, escolhe um envelope. Seja o* *evento $G_i =$ "o i-ésimo jogador ganha R\$100,00". Quais são as probabilidades dos três jogadores.*

- *O primeiro jogador terá três opções para escolher um dos envelopes. Portanto a probabilidade $P(G_1) = \frac{1}{3}$.*
- *Para que segundo jogador ganhe os cem reias é necessário que o primeiro não pegue o envelope com a nota de R\$100,00. Portanto $P(G_2) = P(G_1^c \cap G_2)$, ou seja, $P(G_2) = P(G_1^c \cap G_2) = P(G_2|G_1^c).P(G_1^c) = \frac{1}{2}.\frac{2}{3} = \frac{1}{3}$.*
- *Para que o terceiro jogador ganhe, os dois primeiros terão que perder. Assim, $P(G_3) = P(G_1^c \cap G_2^c \cap G_3) = P(G_3|G_1^c \cap G_2^c)P(G_2^c|G_1^c).P(G_1^c) = 1 \times \frac{1}{2} \times \frac{2}{3} = \frac{1}{3}$.*

Note que os três jogadores tem a mesma probabilidade de ganhar independente da ordem de escolha.

Exemplo 4.5.3. *(Urna de Polya - 1945) Selecionamos uma bola n vezes de uma urna com b bolas brancas e v bolas vermelhas com $b \geq 1$ e $a \geq 1$. Cada bola selecionada é devolvida a urna juntamente com c bolas da mesma cor da bola selecionada, antes da próxima seleção. Seja: $V_i =$ "a i-ésima bola selecionada é vermelha".*
$B_i =$ "a i-ésima bola selecionada é branca".
Então para V_1 e B_1 tem-se:

$$P(V_1) = \frac{v}{(b+v)}$$

e

$$P(B_1) = \frac{b}{(b+v)}.$$

Agora para V_2, ou seja, a segunda bola selecionada seja vermelha, tem-se:

$$V_2 = (B1 \text{ e } V2) \text{ ou } (V1 \text{ e } V2), \text{ ou seja, } V_2 = (B_1 \cap V_2) \cup (V_1 \cap V_2).$$

Daí:

$$P(V_2) = P(B_1 \cap V_2) + P(V_1 \cap V_2) = P(V_2|B_1).P(B_1) + P(V_2|V_1).P(V_1) =$$

$$P(V_2) = \frac{v}{b+v+c} \times \frac{b}{b+v} + \frac{v+c}{b+v+c} \times \frac{v}{b+v}$$

$$P(V_2) = \frac{v \times (b+v+c)}{(b+v) \times (b+v+c)} = \frac{v}{b+v}.$$

Note que o evento V_2 tem mesma probabilidade que o evento V_1. Agora para o evento $B_2 =$ "a segunda bola seleciona é branca", tem-se:

$$P(B_2) = 1 - P(V_2) = 1 - \frac{v}{(b+v)} = \frac{b}{(b+v)}.$$

Observe também que os eventos B_1 e B_2 tem probabilidades iguais. Mostra-se por indução sobre n que $P(V_n) = \frac{v}{(b+v)}$ e $P(B_n) = \frac{b}{(b+v)}$.

4.5.2 Teorema da probabilidade Total

Seja $\{A_1,\ A_2,\ A_3,\ \ldots, A_n\}$ eventos mutuamente exclusivos, com $P(A_1) > 0,\ P(A_2) > 0,\ P(A_3) > 0,\ \ldots\ , P(A_n) > 0$ e, seja B um conjunto tal que: $B \subseteq \bigcup_{j=1}^{n} A_j$. Então:

$$P(B) = P(B|A_1).P(A_1) + P(B|A_2).P(A_2)\ +\ \ldots +\ P(B|A_n).P(A_n)$$

ou

$$P(B) = \sum_{i=1}^{n} P(B|A_j).P(A_j),$$

(MORGADO, et al., 2006, p. 163).

Demonstração: *Veja na Figura 19 que $B = (A_1 \cap B) \cup (A_2 \cap B) \cup (A_3 \cap B) \cup \ldots \cup (A_n \cap B)$. Como esta é uma união disjunta, tem-se:*

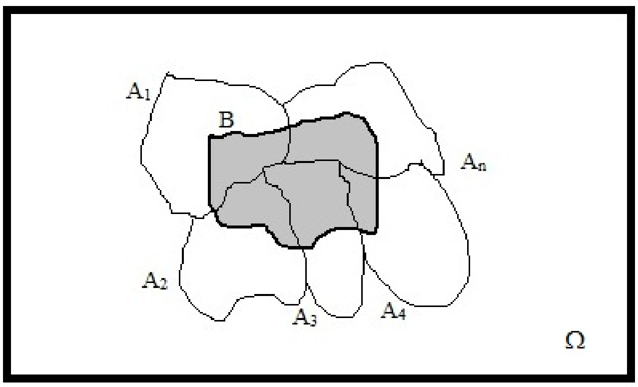

Figura 19: Partição de Ω - União disjunta, $B \subseteq \bigcup_{j=1}^{n} A_j$.

$$P(B) = P(A_1 \cap B) + P(A_2 \cap B) + P(A_3 \cap B)\ + \ldots\ +\ P(A_n \cap B).$$

Por definição de probabilidade condicional decorre que:

$$P(B) = P(B|A_1).P(A_1) + P(B|A_2).P(A_2)\ +\ \ldots\ +\ P(B|A_n).P(A_n).$$

4.5.3 Teorema de Bayes

Nas condições do teorema anterior, se $P(B) > 0$, então, para i, $i \in \{1,\ 2,\ 3,\ \ldots,\ n\}$,

$$P(A_i|B) = \frac{P(B|A_i).P(A_i)}{P(B|A_1).P(A_1) + P(B|A_2).P(A_2)\ + \ldots\ +\ P(B|A_n).P(A_n)},$$

(MORGADO, et al., 2006).

Demonstração:
Da definição de probabilidade condicional tem-se:

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

E pelo teorema anterior, decorre que:

$$P(A_i|B) = \frac{P(B|A_i.P(A_i)}{P(B|A_1).P(A_1) + P(B|A_2).P(A_2)\ + \ldots +\ P(B|A_n).P(A_n)}$$

ou

$$P(A_i|B) = \frac{P(B|A_i).P(A_i)}{\sum_{j=1}^{n} P(A_j)P(B|A_j)}$$

para todo $i\ =\ 1,\ 2,\ 3,\ \ldots,\ n.$

A dedução da fórmula de Bayes é bem simples, porém permite interpretação bastante profunda e que é responsável pelo desenvolvimento de uma área da Estatística que hoje é conhecida como Inferência Bayesiana. Como $A_1,\ A_2,\ A_3,\ \ldots,\ A_n$ é uma partição do espaço amostral Ω, ou seja, $\bigcup_{j=1}^{n} A_j = \Omega$ e $A_i \cap A_j = \phi$, segue que $P(A_1) + P(A_2) + P(A_3) + \ldots +\ P(A_n)\ =\ 1$. Também se verifica que $\sum_{i=1}^{n} P(A_i|B)\ =\ 1$. De fato, isto decorre do axioma iii) da definição de probabilidade na seção 4.3.1.

Exemplo 4.5.4. *Há uma indústria com três máquinas. A porcentagem de peças defeituosas produzidas pelas máquinas I, II e III são* 1%, 3% *e* 2% *respectivamente. Da produção total,* 20% *vem de I,* 30% *vem de II e* 50% *vem de III. Selecione, ao acaso, um item da produção total. Sejam os eventos:*

- B = *"o item selecionado é defeituoso";*

- $A_I =$ *"o item selecionado foi produzido pela máquina I";*
- $A_{II} =$ *"o item selecionado foi produzido pela máquina II";*
- $A_{III} =$ *"o item selecionado foi produzido pela máquina III".*

Qual a probabilidade destes eventos?
Tem-se que: $P(A_I) = 0,2$, $P(A_{II}) = 0,3$ *e* $P(A_{III}) = 0,5$. *Estas são as probabilidades "a priori". O evento B é dado por* $B = (A_I$ *e* $B)$ *ou* $(A_{II}$ *e* $B)$ *ou* $(A_{III}$ *e* $B)$. *Como* $B = (A_I \cap B) \cup (A_{II} \cap B) \cup (A_{III} \cap B)$ *é uma união disjunta, pelo teorema da probabilidade total, tem-se:*

$$P(B) = P(A_I \cap B) + P(A_{II} \cap B) + P(A_{III} \cap B)$$

ou

$$P(B) = P(B|A_I).P(A_I) + P(B|A_{II}).P(A_{II}) + P(B|A_{III}).P(A_{III})$$

$$P(B) = 0,01 \times 0,2 + 0,03 \times 0,3 + 0,02 \times 0,5$$

$$P(B) = 0,021 = 2,1\%$$

Se o desejo é calcular a probabilidade de que o item selecionado foi produzido pela máquina I sabendo-se que este é defeituoso, ou seja, $P(A_I|B)$, *tem-se:*

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

Mas como $P(B) = 0,021$, $P(B|A_I) = 0,01$ *e* $P(A_I) = 0,2$, *decorre que:*

$$P(A_I|B) = \frac{0,01 \times 0,2}{0,021} = 0,095238$$

Esta é a probabilidade "a posteriori". Analogamente tem-se:

$$P(A_{II}|B) = \frac{P(B \cap A_{II})}{P(B)} = \frac{P(B|A_{II}).P(A_{II})}{P(B)} = \frac{(0,03 \times 0,3)}{0,021} = 0,428571$$

$$P(A_{III}|B) = \frac{P(B \cap A_{III})}{P(B)} = \frac{P(B|A_{III}).P(A_{III})}{P(B)} = \frac{(0,02 \times 0,5)}{0,021} = 0,47619$$

Note que $P(A_I|B) + P(A_{II}|B) + P(A_{III}|B) = 1$.

Exemplo 4.5.5. *Suponha que* $0,5\%$ *de uma população tenha determinada doença. Há um teste para essa doença, que não é, entretanto, muito preciso. Além disso, o teste acusa resultado falso positivo para* 1% *das pessoas sadias testadas. Há uma chance de* 95% *do teste dar positivo para uma pessoa que de fato está doente. Seja* D_+ *o evento "ter a doença",* D_- *o evento "não ter a doença",* T_+ *o evento correspondente a "teste positivo"e* T_- *o evento "teste negativo". Deseja-se calcular a probabilidade de alguém, escolhido ao* *acaso, ter a doença* (D_+)*, sabendo que o teste foi positivo* (T_+)*, (DOWNING; CLARK, 2002).*

O problema fornece:

- *P(ter a doença)=* $P(D_+) = 0,005$;

- *P(não ter a doença)=* $P(D_-) = 1 - 0,005 = 0,995$*;*

- *P(teste positivo | não ter a doença)=* $P(T_+|D_-) = 0,01$*;*

- *P(teste positivo | ter a doença)=* $P(T_+|D_+) = 0,95$.

O que se quer calcular é a probabilidade $P(T_+|D+)$. *Como:*

$$P(D_+|T_+) = \frac{P(D_+ \cap T_+)}{P(T_+)},$$

então:

$$P(D_+|T_+) = \frac{P(D_+).P(T_+|D_+)}{P(T_+)}$$

$$P(D_+|T_+) = \frac{0,05 \times 0,95}{?}.$$

Daí, tem-se que calcular $P(T_+)$. *Mas o evento* T_+*, que corresponde a "teste positivo", pode ser dado por:*

(ter a doença e teste positivo)

ou

(não ter a doença e teste positivo),

ou seja:

$$T_+ = (D_+ \cap T_+) \cup (D_- \cap T_+)$$

E como esta é uma união disjunta, tem-se:

$$P(T_+) = P(D_+ \cap T_+) + P(D_- \cap T_+) =$$

$$P(D+).P(T_+|D_+) + P(D-).P(T_+|D_-) =$$

$$P(T_+) = 0,005 \times 0,95 + 0,995 \times 0,01$$

$$P(T_+) = 0,00475 + 0,00995$$

$$P(T_+) = 0,0147.$$

Assim:

$$P(D_+|T_+) = \frac{P(D_+).P(T_+|D_+)}{P(T_+)}$$

Então, a partir dos dados, tem-se:

$$P(D_+|T_+) = \frac{0,05 \times 0,95}{0,0147} = 0,3231$$

Visualiza-se melhor estas probabilidades a partir da Tabela 11.

	Probabilidade de ocorrer D_+	Probabilidade de ocorrer D_-	TOTAL
Probabilidade de ocorrer T_+	95% de 0,5%=0,475%	0,025%	1,47%
Probabilidade de ocorrer T_-	0,025%	98,5%	98,53%
TOTAL	0,5%	99,5%	100%

Tabela 11: Probabilidade Condicional. Fonte: Dados hipotéticos.

Um fato importante é a maneira como se deve enunciar uma probabilidade condicional. A probabilidade de o teste dar positivo, dado que a pessoa tem a doença $P(T_+|D_+)$ não é a mesma coisa que a probabilidadede de a pessoa ter a doença, dado que o teste acusou resultado positivo, $P(D_+|T_+)$.

Para Downing e Clark (2006, p. 75), a regra de Bayes nos diz como achar a probabilidade condicional $P(B|A)$, desde que se conhece $P(A|B)$, ou seja:

$$P(B|A) = \frac{P(A|B).P(B)}{P(A|B).P(B)\ +\ P(A|B^c).P(B^c)}.$$

4.5.4 Independência Estocástica - Eventos independentes.

Seja $(\Omega,\ \mathcal{A},\ P)$ espaço de probabilidade e sejam A e B eventos tais que $P(B|A) = P(B)$. A e B dizem-se **estocasticamente independentes** ou estatisticamente independentes. Assim, da definição de probabilidade condicional $P(A|B) = \frac{(P(A \cap B)}{P(B)}$, decorre que se dois eventos são independentes estocasticamente, então $P(A \cap B) = P(A).P(B)$.

4.5.4.1 Propriedades

i) Ω é independente de qualquer evento A;

ii) ϕ é independente de qualquer evento A;

iii) $P(B|A) = P(B) \Leftrightarrow P(A|B) = P(A)$;

iv) $P(A \cap B) = P(A).P(B) \Leftrightarrow P(B|A) = P(B)$;

v) Se A e B são eventos independentes então:

a) A^c e B são independentes;

b) A e B^c são independentes;

c) A^c e B^c são independentes.

Demonstração:

i) Como $P(\Omega) = 1$ e $A \cap \Omega = A$ tem-se $P(A \cap \Omega) = P(A) \Leftrightarrow P(A \cap \Omega) = P(A).P(\Omega)$.

ii) Como $P(\phi) = 0$ e $A \cap \phi = \phi$ tem-se $P(A \cap \phi) = P(\phi) = 0 = P(A).P(\phi)$.

iii) Da definição de probabilidade condicional tem-se:

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

ou seja, $P(A|B) = P(A)$. Esta propriedade nos diz que se A independe de B, então B independe de A, daí diz-se que A e B **são independentes**, caso contrário diz-se que são **dependentes**.

iv)

($\Rightarrow$) $P(A \cap B) = P(B|A)P(A)$ e como $P(B|A) = P(B)$ tem-se $P(A \cap B) = P(B).P(A)$.

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

v)

a) Da união disjunta $B = (A \cap B) \cup (B - A) = (A \cap B) \cup (B \cap A^c)$, veja a Figura 20, tem-se:

$$P(B) = P(A \cap B) + P(B \cup A^C)$$

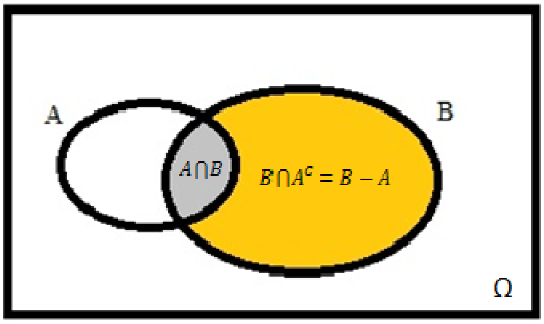

Figura 20: União disjunta $(A \cap B) \cup (B \cap A^c)$

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

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

$$P(A^c \cap B) = [1 - P(A)]\,.P(B)$$

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

b) Da união disjunta $A = (A \cap B) \cup (A - B) = (A \cap B) \cup (A \cap B^c)$, veja a Figura 21, tem:se:

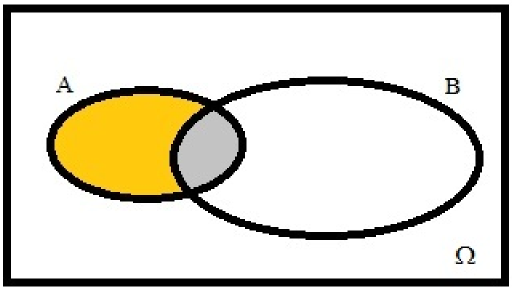

Figura 21: União disjunta $(A \cap B) \cup (A \cap B^c)$

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

$$P(A \cap B^c) = P(A \cap B) - P(A) = P(A) - P(A).P(B) = P(A).[1 - P(B)] = P(A).P(B^c)$$

c) Como $(A^c \cap B^c) = (A \cup B)^c$, veja anexo I, tem-se:

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

$$P(A^c \cap B^c) = 1 - P(A \cup B) = 1 - [P(A) + P(B) - P(A \cap B)]$$

$$P(A^c \cap B^c) = 1 - P(A) - P(B) + P(A).P(B) = 1 - P(A) - P(B).\,[1 - P(A)]$$

$$P(A^c \cap B^c) = P(A^c) - P(B).P(A^c) = P(A^c).\,[1 - P(B)]$$

Portanto,

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

Exemplo 4.5.6. *Uma urna contém 3 bolas vermelhas e duas brancas. Seleciona-se duas bolas, ao acaso, uma de cada vez, e* **com reposição**. *Seja os eventos:*

- B_1 = *"bola branca na 1ª seleção";*
- B_2 = *"bola branca na 2ª seleção";*
- V_1 = *"bola vermelha na 1ª seleção";*
- V_2 = *"bola vermelha na 2ª seleção".*

O espaço amostral deste experimento é $\Omega = \{B_1B_2,\ B_1V_2,\ V_1B_2,\ V_1B_2\}$. *Assim:*

$$P(B_1 \cap B_2) = P(B_2|B_1).P(B_1) = \frac{2}{5}.\frac{2}{5} = \frac{4}{25}$$

e

$$P(B_1) = \frac{2}{5} = P(B_2)$$

Como $P(B_1 \cap B_2) = P(B_1).P(B_2), B_1$ *e* B_2 *são independentes. Análogo verifica-se que* B_1 *e* V_2 *são independentes. Agora, considere o experimento* **sem reposição**. *Daí:*

$$P(B_1 \cap B_2) = P(B_2|B_1).P(B_1) = \frac{1}{4}.\frac{2}{5} = \frac{2}{20}.$$

E como $P(B_1) = \frac{2}{5} = P(B_2)$, *então tem-se:*

$$P(B_1 \cap B_2) \neq P(B_1).P(B_2),$$

logo, B_1 *e* B_2 *não são independentes. Também* V_1 *e* V_2, *para o caso sem reposição, não são independentes. Note que* $P(V_1 \cap V_2) = P(V_2|V_1).P(V_1) = \frac{2}{4}.\frac{3}{5} = \frac{6}{20} = \frac{3}{10}$ *e como* $V_2 = (B_1$ *e* $V_2)$ *ou* $(V_1$ *e* $V_2)$ *que é o mesmo que* $V_2 = (B_1 \cap V_2) \cup (V_1 \cap V_2)$, *e sendo esta uma união disjunta tem-se:* $P(V_2) = P(B_1 \cap V_2) + P(V_1 \cap V_2)$. *Assim:*

$$P(V_2) = P(V_2|B_1).P(B_1) + P(V_2|V_1).P(V_1)$$
$$= \frac{3}{4} \times \frac{2}{5} + \frac{2}{4} \times \frac{3}{5} = \frac{12}{20} = \frac{3}{5}.$$

Daí:

$$P(V_1 \cap V_2) = \frac{3}{10} \neq P(V_1).P(V_2) = \frac{3}{5} \times \frac{3}{5} = \frac{9}{25}$$

4.5.4.2 Eventos independentes - Generalização.

Na seção anterior foi feita uma análise para o caso de dois eventos serem ou não independentes. Agora, considere três eventos A_1, A_2 e A_3. Diz-se que três eventos são independentes se:

i) $P(A_1 \cap A_2) = P(A_1).P(A_2)$

ii) $P(A_2 \cap A_3) = P(A_2).P(A_3)$

iii) $P(A_1 \cap A_3) = P(A_1).P(A_3)$

iv) $P(A_1 \cap A_2 \cap A_3) = P(A_1).P(A_2).P(A_3)$

Para definir independência de três eventos, deve-se impor que essas quatro igualdades valham.

Exemplo 4.5.7. *(DANTAS, 2004, p. 55) Visita-se um casal que tem duas crianças. Sejam os eventos:*

- $M_1 =$ *"o primeiro filho do casal é mulher";*
- $H_1 =$ *"o primeiro filho do casal é homem";*
- $M_2 =$ *"o segundo filho do casal é mulher";*
- $H_2 =$ *"o segundo filho do casal é homem";*
- $A =$ *"os filhos tem o mesmo sexo".*

O espaço amostral é $\Omega = \{M_1M_2,\ M_1H_2,\ H_1M_2,\ H_1H_2\}$. *Note que para os eventos* M_1, M_2 *e* A, *tem-se:* $P(M_1) = P(M_2) = P(A) = \frac{1}{2}$, *e:*

$$P(M_1 \cap M_2) = P(M_1).P(M_2) = \frac{1}{2}.\frac{1}{2} = \frac{1}{4}$$

$$P(M_1 \cap A) = P(M_1).P(A) = \frac{1}{2}.\frac{1}{2} = \frac{1}{4}$$

$$P(M_2 \cap A) = P(M_1).P(H_2) = \frac{1}{2}.\frac{1}{2} = \frac{1}{4}$$

$$P(M_1 \cap M_2 \cap A) = \frac{1}{4} \neq P(M_1).P(M_2).P(A) = \frac{1}{8}$$

Portanto os eventos M_1, M_2 *e* A *são dependentes, pois não satisfazem a igualdade iv).*

A extensão da noção de independência para n eventos $A_1,\ A_2,\ A_3,\ \ldots,\ A_n$ é feita naturalmente a partir do teorema do produto. Diz-se que $A_1,\ A_2,\ A_3,\ \ldots,\ A_n$ são **independentes** se para toda escolha de um número arbitrário deles, a probabilidade da interseção é igual ao produto das probabilidades, ou seja:

$$P(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}) = P(A_{i_1}).P(A_{i_2}).\ \ldots\ .P(A_{i_k}),$$

$\forall k \in \{2,\ 3,\ \ldots,\ n\}$ e $\{i_1,\ i_2,\ \ldots, i_k\} \subset \{1, 2, \ldots, n\}$. Note que o número de relações que terá que ser verificadas para se constatar a independência dos n eventos coincide com o número de subconjuntos com dois ou mais elementos contido em um conjunto com n elementos, ou seja:

$$\sum_{k=2}^{n} \binom{n}{k} = \sum_{k=0}^{n} \binom{n}{k} - \binom{n}{0} - \binom{n}{1} = 2^n - n - 1$$

Portanto, para provar que dois eventos A e B são independentes deve-se verificar somente uma identidade, que é $P(A \cap B) = P(A).P(B)$. Para verificar se 3 eventos são independentes deve-se verificar 4 identidades, como foram citadas no início desta seção e, para verificar se n eventos são independentes deve-se verificar $2^n - n - 1$ identidades (MORGADO, et al., 2006).

Exemplo 4.5.8. *Visita-se um casal que tem três crianças. Sejam os eventos:*

- $M_1 =$ *"o primeiro filho do casal é mulher";*
- $M_2 =$ *"o segundo filho do casal é mulher";*
- $M_1 =$ *"o terceiro filho do casal é mulher"*

O espaço amostral é:

$\Omega = \{M_1M_2M_3,\ M_1M_2H_3,\ M_1H_2M_3,\ H_1M_2M_3,\ H_1H_2M_3,\ H_1M_2H_3,\ M_1H_2H_3,\ H_1H_2H_3\}$.

Note que para os eventos M_1, M_2 e M_3, tem-se: $P(M_1) = P(M_2) = P(M_3) = \frac{4}{8} = \frac{1}{2}$, e:

$$P(M_1 \cap M_2) = P(M_1).P(M_2) = \frac{4}{8} \times \frac{4}{8} = \frac{1}{4},$$

$$P(M_1 \cap M_3) = P(M_1).P(M_3) = \frac{4}{8} \times \frac{4}{8} = \frac{1}{4},$$

$$P(M_2 \cap M_3) = P(M_1).P(H_2) = \frac{4}{8} \times \frac{4}{8} = \frac{1}{4},$$

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

Portanto os eventos M_1, M_2 *e* M_3 *são independentes, pois satisfazem as quatro igualdades.*

5 VARIÁVEL ALEATÓRIA

Um evento aleatório é algo que não se sabe ao certo se ocorrerá, mas cuja probabilidade de ocorrência pode-se calcular. A partir desta ideia define-se **variável aleatória**, representada por **v.a.**, como sendo uma variável tal que não se sabe ao certo que valor ela poderá assumir, mas para a qual se pode calcular a probabilidade de tomar determinado valor.

5.1 Definição

Formalmente variável aleatória é definida como uma função de conjuntos (espaço amostral) para números reais, ou seja, é uma variável que toma um valor especificado quando ocorre determinado evento aleatório (FONSECA; MARTINS, 2006, p. 37). Assim, variável aleatória é uma função cujo valor é um número real determinado por cada elemento em um espaço amostral.

Assim, seja E um experimento e Ω o espaço amostral associado a este experimento. Uma função X, que associe a cada elemento $w \in \Omega$ um número real $X(w)$ é denomida variável aleatória. Veja na Figura 22, a representação por digramas da função X.

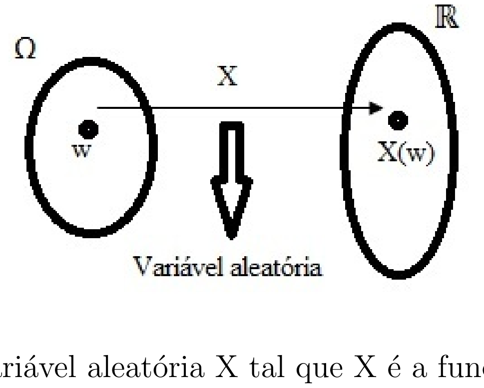

Figura 22: Variável aleatória X tal que X é a função $f : \Omega \to \mathbb{R}$.

Exemplo 5.1.1. *Duas peças são retiradas sucessivamente, sem reposição, de uma caixa que contém quatro peças boas (B) e 3 defeituosas (D). Para esse experimento o espaço* *amostral é dado por* $\Omega = \{BB,\ BD,\ DB,\ DD\}$. *Seja a função que associa o número de peças boas ao conjunto* $\{0,\ 1,\ 2\}$. *Fazendo uma correspondência entre* Ω *e* $\{0,\ 1,\ 2\}$ *tem-se:*

Ponto amostral	X=x
BB	2
BD	1
DB	1
DD	0

Tabela 12: A função $f : \Omega \to \{0,\ 1,\ 2\}$.

Utiliza-se letras maiúsculas latinas (X, Y, Z, etc.) para representar a variável aleatória e a correspondente letra minúscula (x, y, z, etc.) para representar um dos seus valores.

Se um espaço amostral contém um número finito de pontos ou uma sequência infinita enumerável de pontos amostrais como encontrar peças defeituosas em um lote de produção com n peças, o número de filhos do sexo masculino de um casal, o número de lançamentos de um dado até que se encontre o número 6 na face voltada para cima, o número de acidentes que ocorre na BR 365 que liga Patrocínio e Uberaba, etc., ele é denominado **espaço amostral discreto**. A variável aleatória definida sobre este espaço amostral é denominada **variável aleatória discreta (v.a.d)**.

Por outro lado, se um espaço amostral contém pontos amostrais que formam uma continuidade então ele é denominado **espaço amostral contínuo**. A variável definida sobre esse espaço é denominada **variável aleatória contínua (v.a.c.)**. Alturas de um grupo de pessoas, pesos de animais, tempos de vida de um componente mecânico ou eletrônico, temperaturas, podem ser consideradas como v.a.c., mas deve-se tomar cuidado que suas medidas estão limitadas na precisão do aparelho de medida.

Veja na Figura 23 que a representação gráfica de uma variável aleatória discreta é um gráfico de dispersão de pontos por ser uma função cujo domínio é um conjunto finito ou infinito enumerável, enquanto que o gráfico da variável aleatória contínua é um linha, pois seu domínio é um conjunto infinito.

Na maior parte dos problemas práticos as variáveis aleatórias discretas representam

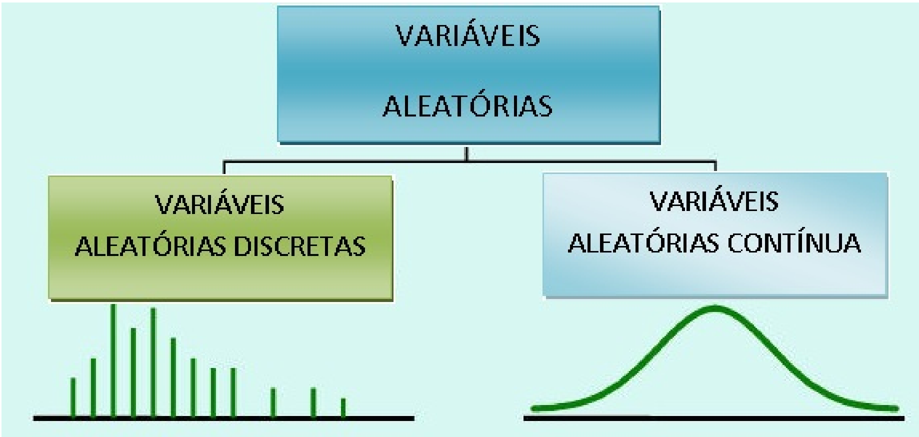

Figura 23: Variável aleatória discreta x Variávela aleatória contínua.

dados originados da contagem de elementos de um conjunto e as variáveis aleatórias contínuas representam dados originados de medidas (DOWNING; CLARK, 2002). A variável aleatória discreta não precisa tomar apenas valores em números inteiros. Por exemplo, no lançamento de dois dados, seja M a variável aleatória que representa a média dos dois números que estão na face voltada para cima. Então M admite os valores do conjunto $\left\{1, \frac{3}{2}, 2, \frac{5}{2}, 3, \frac{7}{2}, 4, \frac{9}{2}, 5, \frac{11}{2}, 6\right\}$.

Exemplo 5.1.2. *Geralmente para definir uma variável aleatória segue-se os dois casos a seguir (DOWNING; CLARK, 2002, p.79):*

- *Seja um conjunto de N elementos. Desses, M tem uma característica de interesse, e os outros N-M não a têm. A estimativa de M deve ser feita a partir de uma amostra aleatória de n desses elementos e seja X o número de elementos que tem essa característica. X é uma v.a., pois seus valores depende dos elementos escolhidos aleatoriamente para formar a amostra.*

- *Seja um conjunto de indivíduos em que cada um destes possui uma carasterística mensurável, tal como altura, idade, tempo de assistência à televisão. Escolhe-se uma pessoa ao acaso e seja X o valor do atributo para a mesma. X é uma variável aleatória.*

5.2 Distribuições de Probabilidade de uma variável aleatória discreta

Para o exemplo 5.1.1 pode-se calcular a probabilidade da variável aleatória assumir um determinado valor, que representa-se por $P(X = x_i)$ para $i = \{1,\ 2,\ \ldots,\ \}$, obtendo assim a Tabela 13 seguinte.

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

Tabela 13: A função $f(x) = P(X = x_i)$

A partir da Tabela 13 verifica-se que é possível estabecer uma relação unívoca entre os valores da variável aleatória X e a probabilidade de ocorrência de um ponto amostral de Ω. Esta correspondência define uma função onde os x_i formam o domínio da função e as probabilidades p_i o seu conjunto imagem. Esta função, assim definida, é denominda **função de probalidade**. Chama-se **função de probabilidade** de uma variável aleatória discreta X que assume os valores x_i para para $i = \{1,\ 2,\ \ldots\}$ a função $f(x) = P(X = x_i)$ que a cada valor x_i associa sua probabilidade de ocorrência.

Exemplo 5.2.1. *Seja o experimento "dois lançamentos consecutivos de uma moeda", e seja a variável aleatória X = "ganha-se 1 se der cara (C) em qualquer dos lançamentos e perde-se 1 se der coroa em qualquer dos lançamentos". Pode-se representar por "sucesso = C = 1"e "fracasso = K = 0". Assim* $\Omega = \{(0,0);(1,0);(0,1);(1,1)\}$ *. e*

$$X = \begin{cases} -2, \ \textit{se ocorrer}\ (0,0); \\ 0, \ \textit{se ocorrer}\ (1,0)\ \textit{ou}\ (0,1); \\ +2, \ \textit{se ocorrer}\ (1,1). \end{cases}$$

A variável aleatória X é uma função definida em Ω *a valores na reta real, pois:*

- $(X = -2) = \{w \in \Omega; X(w) = -2\} = \{(0,0)\}$;
- $(X = 0) = \{w \in \Omega; X(w) = -2\} = \{(0,1);(1,0)\}$;
- $(X = +2) = \{w \in \Omega; X(w) = -2\} = \{(1,1)\}$;

Assim, se P(0)=q e P(1)=p então:

- $P(X=-2) = P\left[(0,0)\right] = P(0).P(0) = q.q = q^2.$
- $P(X=0) = P\left[(0,1);(1,0)\right] = P(0).P(1) + P(1).P(0) + = p.q + q.p = 2p.q$
- $P(X=+2) = P\left[(1,1)\right] = P(1).P(1) = p.p = p^2$

Note que

$$P(X=-2) + P(X=0) + P(X=+2) = 1,$$

então

$$q^2 + 2p.q + p^2 = 1 \Leftrightarrow$$

$$(p+q)^2 = 1 \Leftrightarrow p+q = 1.$$

Se $p = q = \frac{1}{2}$ *tem-se:*

$$X = \begin{cases} -2, & P(X=-2) = \frac{1}{4}; \\ 0, & P(X=0) = \frac{1}{2}; \\ +2, & P(X=+2) = \frac{1}{4}; \end{cases}$$

Define-se **distribuição de probabilidade** da variável aleatória X o conjunto de pares $[x_i, P(x_i)]$, que, geralmente, pode ser representada por meio de tabelas ou gráficos. Veja o gráfico da Figura 24, que representa a distribuição de probabilidade do exemplo 5.2.1, que também pode ser dada pelos pares:

$$\{[-2,\ P(X=-2)]\,;\ [0,\ P(X=0)]\,;\ [1, P(X=1)]\}\,.$$

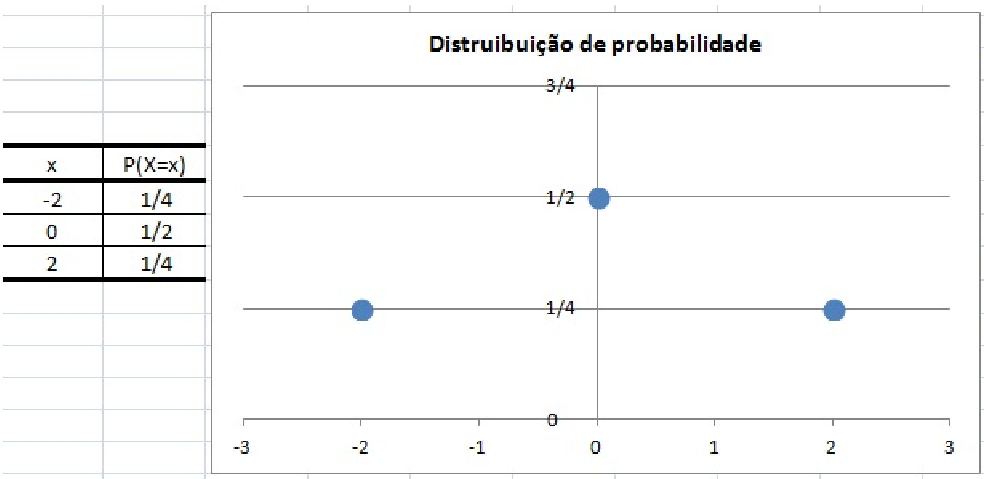

Figura 24: Distribuição de probabilidade para o exemplo 5.2.1, tomando $p = q = \frac{1}{2}$.

Definição formal de uma distribuição de probabilidade discreta:

Definição 5.2.1. *Seja X uma variável aleatória que pode assumir os valores* x_i *para para* $i = \{1,\ 2,\ \ldots\}$. *A cada valor* x_i *correspondem pontos do espaço amostral. Associa-se a* *cada valor x_i a probabilidade p_i de ocorrência de tais pontos no espaço amostral. Assim tem-se:*

$$\sum p_i = 1.$$

Os valores x_i e suas correspondentes probabilidades p_i definem uma distribuição de probabilidade.

Exemplo 5.2.2. *Para o exemplo 4.1.5, "Um dado é lançado tal que a chance de ocorrer a face de número j, com $j \in \{1,\ 2,\ 3,\ 4,\ 5,\ 6\}$, é proporcional a j", a função de probabilidade pode ser representada como mostra Figura 25, em que foi feita a representação da distribuição de probabilidade, em tabela e em gráfico.*

$$P(x_i) = \begin{cases} \frac{x_i}{21}, & para \quad x_i = \{1,\ 2,\ 3,\ 4,\ 5,\ 6\}. \\ \\ 0, & para\ outros\ valores\ de\ x_i. \end{cases}$$

x	P(X=x)
1	1/21
2	2/21
3	1/7
4	4/21
5	5/21
6	2/7

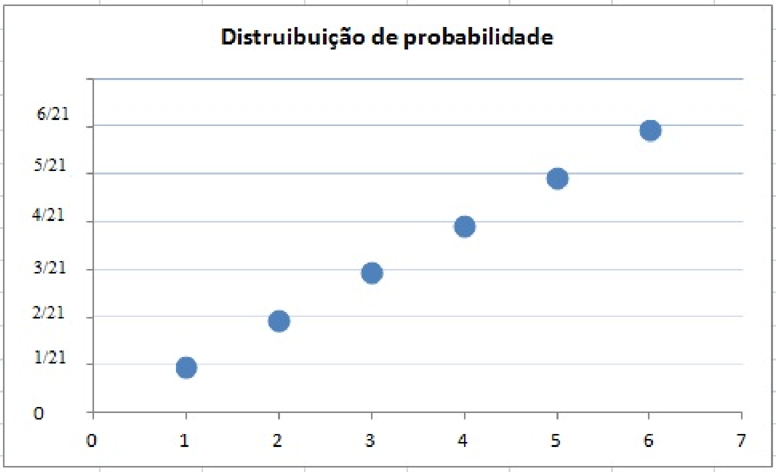

Figura 25: Distribuição de probabilidade: exemplo 4.1.5.

As tabelas e os gráficos são utilizados para mostrar como a probabilidade total atribuída a um espaço amostral é distribuída pelos diversos restultados daquele espaço.

5.3 Função de Distribuição Acumulada.

É sempre interessante saber a probabilidade de uma variável aleatória X tomar determinado valor. Mas, às vezes, necessita-se saber a probabilidade de X ser no máximo igual a um valor particular x' (DOWNING; CLARK, 2002). Por exemplo, o jogo 21 que consiste em retirar cartas de um baralho até que se obtenha soma dos pontos no máximo 21. Vence aquele que atingir o valor mais próximo ou igual a 21. Assim, se um jogador já tirou um 7 e um 8, interessa-lhe a probabilidade de que o valor da próxima carta seja no máximo igual a 6.

De modo geral, seja $x_1,\ x_2,\ x_3,\ \ldots,\ x'$ todos os valores possíveis de uma variável X no máximo igual a x', então:

$$P(X \leq x') = f(x_1) + f(x_2) + f(x_3) + \ \ldots\ + f(x').$$

Esta função, que dá a probabilidade de X ser no máximo igual a um valor particular, é denominada **função de distribuição acumulada da variável aleatória X** e é representada por um F maiúsculo, sendo:

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

Exemplo 5.3.1. *Seja a variável X definida como o número de caras obtidas em quatro jogadas de um moeda equilibrada. O espaço amostral* Ω *é dado por:*

$$\Omega = \left\{ \begin{array}{ccccc} & & KKCC & & \\ & & KCKC & & \\ & CKKK & KCCK & CCCK & \\ & KCKK & CKCK & CCKC & \\ & KKCK & CKKC & CKCC & \\ KKKK & KKKC & CCKK & KCCC & CCCC \end{array} \right\}.$$

Assim X pode assumir os valores $\{0, 1, 2, 3, 4\}$.

As probabilidades para estes valores são:
Probabilidade de ocorrer 0 caras:

$$P(0) = \frac{1}{16},$$

Probabilidade de ocorrer 1 caras:

$$P(1) = \frac{4}{16},$$

Probabilidade de ocorrer 2 caras:

$$P(2) = \frac{6}{16},$$

Probabilidade de ocorrer 3 caras:

$$P(3) = \frac{4}{16} \quad e$$

Probabilidade de ocorrer 4 caras:

$$P(4) = \frac{1}{16}.$$

A função de probabilidade para um número de caras X que aparece em 4 jogadas de uma moeda equilibrada é dada pelos pares:

$$f(0) = \frac{1}{16},$$

$$f(1) = \frac{4}{16},$$

$$f(2) = \frac{6}{16},$$

$$f(3) = \frac{4}{16} \; e$$

$$f(4) = \frac{1}{16}.$$

Veja na Figura 26 a distribuição de probabilidade representada graficamente.

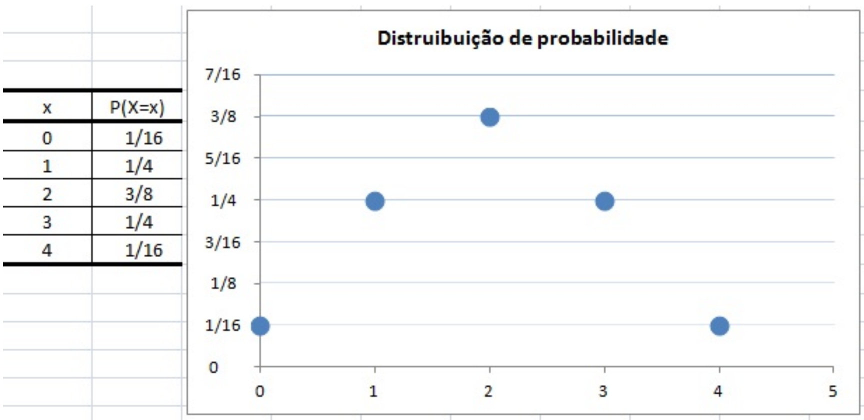

Figura 26: Distribuição de probabilidade da variável X definida como o número máximo de caras obtidas em quatro jogadas de um moeda equilibrada.

A função de distribuição acumulada para o número de caras X, que apresenta em 4 *jogadas de uma moeda equilibrada pode ser dada por:*

$$F(x) = \begin{cases} 0, se\ x < 0 \\ \frac{1}{16},\ se\ 0 \leq x < 1 \\ \frac{5}{16},\ se\ 1 \leq x < 2 \\ \frac{11}{16},\ se\ 2 \leq x < 3 \\ \frac{15}{16},\ se\ 3 \leq x < 4 \\ 1,\ se\ 4 \leq x \end{cases}$$

Veja na Figura 27 a função distribuição acumulada de probabilidades representada graficamente.

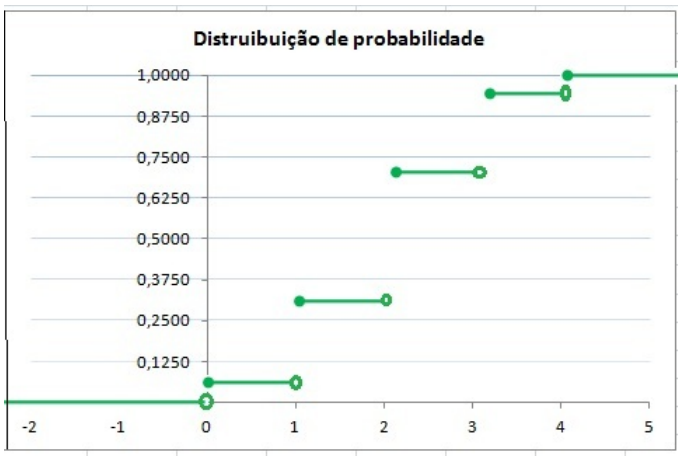

Figura 27: A função de distribuição acumulada F(x) dá as probabilidades de a variável aleatória ser no máximo igual a um valor particular: $F(x') = P(X \leq x')$.

5.4 Exemplo de Variável Aleatória: Provas de Bernoulli.

Considere um experimento aleatório que pode ser repetido em condições semelhantes um número finito de vezes, e que neste experimento há uma probabilidade p do experimento resultar em sucesso e de $1-p$ de resultar em fracasso, em qualquer prova. Experimento como esse, que pode apresentar somente dois resultados possíveis, "sucesso"ou "fracasso", é denominado **Prova de Bernoulli** (DOWNING; CLARK, 2006, p.87).

Exemplo 5.4.1. *Uma população tem aqueles que tem uma certa característica de interesse (diabético, daltônico, etc.) e aqueles que não a tem. Selecione ao acaso e com reposição, indivíduos, um de cada vez, até encontrar um portador da característica de interesse. Em cada seleção considere* S = *"selecionar um portador da característica"e* F *"selecionar um não portador da característica"'. Considere o experimento "selecio-*

nar n indivíduos até ocorrer o 1º sucesso". Este problema é conhecido como "Ensaio de Bernoulli"ou "Prova de Bernoulli"'. Tem-se:

$$P(S) = p \ \ e \ \ P(F) = q, \quad onde \ p + q = 1$$

e

$$\Omega = \{S, FS, FFS, FFFS, FFFFS, \ldots\}$$

Então:
$P(S) = p$
$P(FS) = q.p$
$P(FFS) = q.q.p = q^2.p$
$P(FFFS) = q.q.q.p = q^3.p$
$\vdots$
$P(FFF \ \ldots \ S) = q.q.q. \ \ldots \ .p = q^k.p \ (k \ fracassos)$

A variável aleatória X, que assume apenas os valores 1 (sucesso) e 0 (fracasso), com função de probabilidade representada por $(x, p(x))$ tal que:

$$p(0) = P(X = 0) = 1 - p$$

e

$$p(1) = P(X = 1) = p$$

é denominada "variável aleatória de Bernoulli".

O valor de p é uma probabilidade de valor desconhecido, que a partir de uma verificação da população calcula-se ou estima-se este valor (DANTAS, 2004).

5.5 Características de uma variável aleatória

Conhecer a função de probabilidade ou a função de distribuição acumulada de uma v.a. é muito importante. Mas para que se tenha uma completa informação sobre esta variável aleatória o ideal é que se procure calcular o **valor médio** que esta variável assumiria após repetir o experimento um número muito grande de vezes que é denominado **esperança matemática** (ou valor esperado), e também **a variância** e o **desvio padrão** correspondente.

5.5.1 Esperança (ou média) de uma variável aleatória discreta

A esperança de uma variável aleatória é o valor médio que ocorreria se observasse a variável aleatória um número grande de vezes.

Suponha $\Omega = \{w_1,\ w_1,\ \ldots,\ w_N\}$ finito. Seja $X : \Omega \longrightarrow \mathbb{R}$ uma variável aleatória discreta e que X assuma os valores $x_1,\ x_2,\ \ldots\ x_n$, onde $n \leq N$. A esperança (ou valor esperado) de X é dado:

$$E(X) = \sum_{j=1}^{N} X(w_j).P(w_j).$$

Segundo Downing e Clark (2002, p. 86), a esperança pode ser entendida como sendo a média ponderada de $X(w_1),\ X(w_2),\ \ldots, X(w_n)$ com pesos $P(\{w_1\}),\ P(\{w_2\}),\ \ldots,\ P(\{w_n\})$, ou seja,

$$E(X) = \sum_{j=1}^{n} x_i.f(x_i) = \sum_{j=1}^{n} x_i.P(X = x_i).$$

Obs No caso em que a variável aleatória discreta assume um número infinito, porém enumerável, de valores, $\{w_1,\ \ldots,\ w_n, \ldots\}$, com probabilidades $\{p_1,\ \ldots,\ p_n, \ldots\}$, tal que $p_i > 0$ e $\sum_{i=1}^{\infty} p_i = 1$ a definição de esperança é definida como sendo uma "soma infinita convergente" (BUSSAB, 2006). Veja os exemplos 4.3.7 e 4.3.8.

Exemplo 5.5.1. *Considere a variável aleatória X do exemplo 5.3.1., (número de caras obtidas em 4 jogadas de um moeda equilibrada). A esperança da variável X é*

$$E(X) = \left(0 \times \frac{1}{16}\right) + \left(1 \times \frac{4}{16}\right) + \left(2 \times \frac{6}{16}\right) + \left(3 \times \frac{4}{16}\right) + \left(4 \times \frac{1}{16}\right)$$

$$\Leftrightarrow\ E(X) = \frac{32}{16} = 2$$

Exemplo 5.5.2. *Considere a variável aleatória Y definida como sendo o número que aparece na face superior no lançamento de um dado equilibrado. A esperança da variável*

Y é

$$E(Y) = \left(0 \times \frac{1}{6}\right) + \left(1 \times \frac{1}{6}\right) + \left(2 \times \frac{1}{6}\right) + \left(3 \times \frac{1}{6}\right) + \left(4 \times \frac{1}{6}\right) + \left(5 \times \frac{1}{6}\right) + \left(6 \times \frac{1}{6}\right)$$

$$\Leftrightarrow \; E(Y) = \frac{21}{6} = 3,5$$

5.5.2 Variância e desvio padrão de uma variável aleatória discreta

Seja X uma variável aleatória discreta que assume os valores $x_1, \; x_2, \; \ldots \; x_n$. Para o cálculo da variância deve-se tomar o quadrado das distâncias de cada valor da variável e em seguida calcular a esperança desses valores. Isto significa calcular o desvio médio quadrático de X em torno da média, onde o desvio é dado pela diferença entre cada valor da variável e a média (DOWNING; CLARK, 2006). Assim denomina-se **variância** de uma variável aleatória discreta X, que é representada por $Var(X)$ ou σ^2, o valor

$$Var(X) = \sum_{j=1}^{n} \left[x_i - E(X)\right]^2 .P(X = x_i).$$

Então tem-se:

$$Var(X) = E\left[(X - E(X))^2\right].$$

Mas

$$\begin{aligned} E\left[(X - E(X))^2\right] &= E\left[X^2 - 2.X.E(X) + E(X)^2\right] \\ &= E(X^2) - E\left[2.X.E(X)\right] + E\left[E(X)^2\right] \\ &= E(X^2) - 2.E(X).E(X) + E(X)^2 \end{aligned}$$

Daí

$$Var(X) = E(X^2) - \left[E(X)^2\right].$$

Esta fórmula torna os cálculos mais fáceis.

O **desvio padrão** de X, representado por DP(X) ou σ, é definido como a raiz quadrada positiva da variância.

Exemplo 5.5.3. *A esperança (ou média) da variável aleatória de Bernoulli é o próprio p, já que*

$$E(X) = \sum_{j=1}^{n} x_i.f(x_i) = \sum_{j=1}^{n} x_i.P(X = x_i) = 0 \times (1 - p) + 1 \times p = p.$$

A variância é $p(1-p)$. *Como* $E(X^2) = 0^2 \times (1-p) + 1^2 \times p = p$, *tem-se, a partir da fórmula* $Var(X) = E(X^2) - [E(X)^2]$, *que:*

$$Var(X) = p - p^2 = p(1-p).$$

5.6 Algumas distribuições de variáveis aleatórias discretas.

Nesta unidade será feita uma discussão de algumas distribuições de probabilidades discretas com alguns exemplos de aplicação. Tais distribuições partem da pressuposição de certas hipóteses bem definidas. Como diversar situações reais muitas vezes se aproximam destas hipóteses, esses modelos são úteis no estudo de tais situações, daí a importância deste estudo.

5.6.1 Distribuição Uniforme

É a mais simples de todas as distribuições discretas de probabilidade. É aquela na qual a v.a. assume todos os seus valores com a mesma probabilidade. Veja na Figura 28 a representação gráfica para um caso geral, em $P(x_i) = \frac{1}{k}$ $\forall i \in \mathbb{N}$.

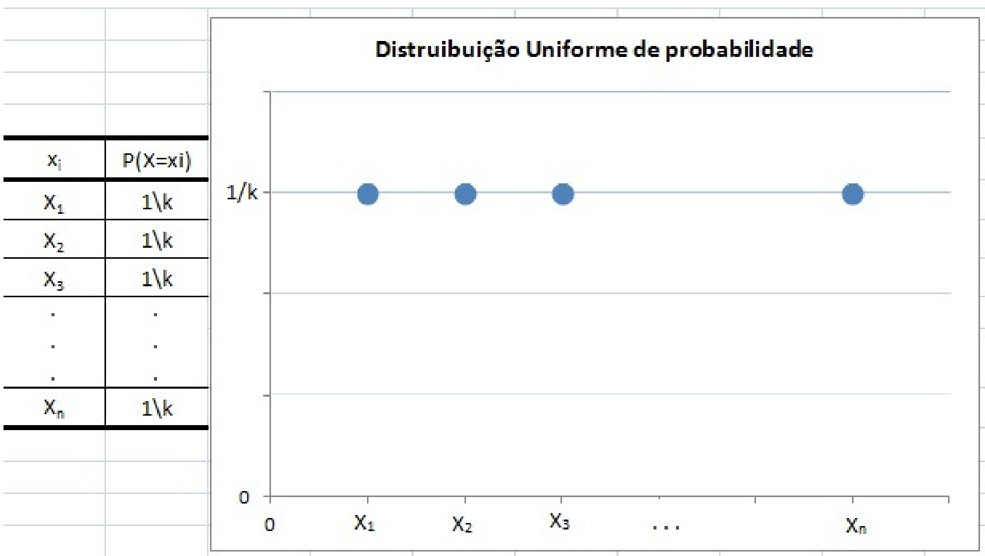

Figura 28: Disbribuição Uniforme: a variável aleatória assume os seus valores com a mesma probabilidade.

Definição: A variável aleatória discreta X, qua assume os valores x_1, x_2, ... x_k tem **distribuição uniforme** se, e somente se,

$$P(x,\ k)=\frac{1}{k}=P(X=x_i),\ \forall i\in\{1,2,3,\dots\ k\}$$

Verifica-se que a **esperança** e a **variância** de uma distribuição uniforme são dadas, respectivamente, por (BUSSAB; MORETTIN, 2006):

$$E(X)=\frac{1}{k}.\sum_{i}^{k}x_i$$

$$Var(X)=\frac{1}{k}.\left\{\sum x_i^2-\frac{\left(\sum x_i^2\right)}{k}\right\}.$$

Exemplo 5.6.1. *Para o lançamento de um dado equilibrado tem-se:*

$$E(X)=\frac{1}{6}\{1+2+3+4+5+6\}=3,5$$

e

$$Var(X)=\frac{1}{6}\left\{\left(1^2+2^2+3^3+4^2+5^2+6^2\right)-\frac{21^2}{6}\right\}=2,92$$

5.6.2 Distribuição Binomial (Uma extensão da distribuição de Bernoulli).

Suponha um experimento cujos resultados podem ser classificados em "sucesso"ou "fracasso". Se fizer $X=1$ quando obtiver um sucesso e $X=0$ quando obtiver um fracasso, então a função de probabilidade de X será dada por:

$$P(X=1)=p \quad e$$

$$P(X=0)=1-p.$$

Nesse caso se diz que a variável aleatória tem distribuição de Bernoulli (BUSSAB; MORETTIN, 2006, p.142).

A distribuição Binomial advém da distribuição de Bernoulli quando repete-se um ensaio (ou prova) de Bernoulli n vezes.
Assim tem-se:

- São realizados n ensaios independentes e do mesmo tipo;
- Cada ensaio admite dois resultados: "sucesso"ou "fracasso";
- A probabilidade de sucesso em cada ensaio é p (valor constante). Daí, a probabilidade de fracasso será $q=1-p$;

Assim, ao associar uma v.a. X="número de sucessos"nesses n ensaios, X poderá assumir os valores $\{0, 1, 2, 3, \; \ldots n\}$.

Exemplo 5.6.2. *A probabilidade de um atirador acertar um alvo é de 75%. Tome para "sucesso"acertar o alvo e "fracasso"errar. Assim:*

$$P(sucesso) \; = \; P(S) = 0,75$$

e

$$P\,(fracasso) = \; P(S) \; = \; 0,25.$$

Para oito tentativas o espaço amostral possui 2^8 *elementos. Suponha a variável aleatória X "o número de vezes em que o atirador acerta o alvo". X poderá assumir os valores* $\{0, \; 1, \; 2, \; 3, \; 4, \; 5, \; 6, \; 7, \; 8\}$.

- *0S8F, onde o atirador erra todas as tentativas. Assim só há uma única situação que pode ser representada por FFFFFFFF. Como* $P(F) = 0,25$ *tem-se:*

  $$P(0S8F) = 1 \times 0,25^8 \times 0,75^0.$$

- *1S7F, onde há 7 situações distintas, pois o S pode figurar em 7 posições, sendo: SFFFFFFF, FSFFFFFF, FFSFFFFF, ... FFFFFFFS.*
  Daí tem-se:

  $$\frac{8!}{1! \times 7!} \times 0,25^7 \times 0,75^1.$$

- *2S6F, onde há 15 situações distintas, já que há um permutação com 8 elementos sendo 2S e 6F, ou seja: SSFFFFFF, FSSFFFFF, FFSSFFFF, ... FFFFFFSS.Daí tem-se:* $\frac{8!}{2! \times 6!} \times 0,25^6 \times 0,75^2$.

  Análogo tem-se:

- *3S5F,* $\frac{8!}{3! \times 5!} \times 0,25^5 \times 0,75^3$.
- *4S4F,* $\frac{8!}{4! \times 4!} \times 0,25^4 \times 0,75^4$.

- *5S3F*, $\frac{8!}{5!\times 3!}\times 0,25^3\times 0,75^5$.
- *6S2F*, $\frac{8!}{6!\times 2!}\times 0,25^2\times 0,75^6$.
- *7S1F*, $\frac{8!}{7!\times 1!}\times 0,25^1\times 0,75^7$.
- *8S0F*, $\frac{8!}{8!\times 0!}\times 0,25^0\times 0,75^8$.

Fazendo os cálculos utizando uma planilha eletrônica obtém-se a distribuição de probalidade da variável aleatória discreta X, veja a Figura 29, dada por:

Variável X (Número de Sucessos)	P(X=xi)
0	0,000015
1	0,000366
2	0,003845
3	0,023071
4	0,086517
5	0,207642
6	0,311462
7	0,266968
8	0,100113

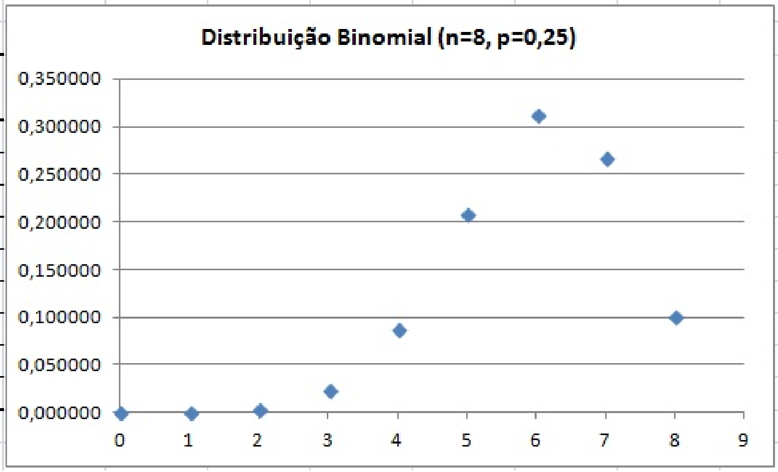

Figura 29: Disbribuição de Binomial: n=8 e P(S)=0,75.

Assim, pode-se concluir que, a distribuição de probabilidades de uma variável aleatória que tem **distribuição binomial** é dada por (BUSSAB; MORETTIN, 2006):

$$P\left(x,\ n,\ p\right)=\begin{cases} C_n^x p^x q^{n-x}, & se\ \ x=\{0,\ 1,\ 2,\ 3,\ \ldots,\ n.\} \\ \\ 0, & se\ \ caso\ contrário. \end{cases}$$

Veja outros exemplos de aplicação da distribuição binomial.

Exemplo 5.6.3. *Um casal planeja ter n filhos. Considere que se o nascituro for do sexo masculino, a variável assume valor 1 (sucesso) e caso contrário assume valor 0 (fracasso).*

Esta é uma distribuição binomial que resulta da aplicação de n provas (ou ensaios) de Bernoulli (a soma de n variáveis binárias). Para estudar a distribuição do número de *crianças do sexo masculino em n nascimentos, é necessário fazer* $n = \{1, 2, 3, \ldots,\ n\}$. *Seja*

$$P(1) = p = \frac{1}{2}$$

a probabilidade de um nascimento ser uma criança do sexo masculino e

$$P(0) = q = 1 - p$$

a probabilidade de esta criança ser do sexo feminino. Logo, como $p + q = 1$ *tem-se* $q = \frac{1}{2}$.

Para n=3, isto é, no caso de três nascimentos, o número de meninos assume os valores $\{0,\ 1,\ 2,\ 3.\}$. *Veja na Tabela 14 a distribuição de probabilidades para este caso em que n=3.*

Números de nascituros do sexo masculino	Probabilidade	
0	$\{qqq\} = q^3$	$= \frac{1}{8} = 0,125$
1	\$ $\{qqp,\ qpq,\ pqq\} = 3q^2p$	$= \frac{3}{8} = 0,375$
2	$\{qpp,\ pqp,\ ppq\} = 3qp^2$	$= \frac{3}{8} = 0,375$
3	$\{ppp\} = p^3$	$= \frac{1}{8} = 0,125$
Total	$q^3 + 3p^2q + 3pq^2 + p^3$	$= 1$

Tabela 14: A distribuição binomial para $n = 3$ e $p = \frac{1}{2}$.

A distribuição binomial fica definida quando são dados dois parâmetros: o número n de variáveis aleatórias binárias observadas (por exemplo, 3 nascimentos) e a probabilidade p de ocorrer valor 1 em uma única observação (no exemplo anterior, a probabilidade de nascer uma criança do sexo masculino). Assim, dados n e p, a probabilidade de a variável aleatória discreta assumir valor x é obtido pela expressão

$$P(x, n, p) = C_n^x p^x q^{n-x}.$$

O nome Binomial é devido a expressão $C_n^x p^x q^{n-x}$ ser um termo do binômio $(p + q)^n$.

Exemplo 5.6.4. *Suponha, no exemplo anterior, n=4. Sendo* $p = \frac{1}{2}$ *a probabilidade de nascer um menino, a probabilidade de ocorrerem 2 meninos em 4 nascimentos é:*

$$P\left(2,\ 4,\ \frac{1}{2}\right) = C_4^2 \left(\frac{1}{2}\right)^2 \cdot \left(\frac{1}{2}\right)^{4-2} = 6.\frac{1}{4}.\frac{1}{4} = 0,375 \quad ou \quad 37,5\%$$

Exemplo 5.6.5. *Uma empresa de seguros faz uma pesquisa e verifica que, numa grande cidade, a probabilidade de que um carro de certo modelo seja roubado, no período de um ano, é 5%. Considerando uma amostra aleatória de 10 destes carros pode-se calcular a* *probabilidade de que exatamente um carro seja roubado no período de um ano.*

Assim:

$$P(X) = C_{10}^{1}\left(\frac{1}{20}\right)^{1}.\left(\frac{19}{20}\right)^{9} = 10.0,05.0,63 \cong 0,3151 = 31,51\%.$$

Admita-se neste caso independência entre os eventos associados aos roubos de cada carro.

Exemplo 5.6.6. *Se 5% dos parafusos produzidos por uma máquina são defeituosos, é possível determinar a probabilidade de, entre n parafusos escolhidos ao acaso, no máximo k deles ($k \leq n$) sejam defeitusos.*

Sendo a Distribuição Binomial uma soma de n variáveis do tipo "Bernoulli", a **esperança** e a **variância** são dadas, respectivamente, por:

$$E(X) = np$$

e

$$Var(X) = np(1-p).$$

Exemplo 5.6.7. *Para o exemplo 5.6.4 tem-se:*

$$E(X) = 10.\frac{1}{20} = 0,5$$

e

$$Var(X) = 10.\frac{1}{20}.\frac{19}{20} = 0,475$$

5.6.3 Distribuição Multinomial

Essa distribuição é uma extensão da distribuição Binomial. O experimento binomial se transforma em um experimento multinomial se em cada tentativa (prova ou ensaio) tem-se mais de dois possíveis resultados de interesse. Como exemplo considere o caso em que se tem 3 possíveis resultados: classificação de um produto em defeituoso, aceitável e recuperável; ou de acidentes de trânsitos em leves, graves ou fatais.

Generalizando, se numa dada tentativa podem ocorrer k resultados, $A_1,\ A_2,\ A_3,\ \ldots,\ A_k$, com probabilidades $p_1,\ p_2,\ p_3,\ \ldots,\ p_k$, então a distribuição de probabilidades das variáveis aleatórias discretas $X_1,\ X_2,\ X_3,\ \ldots,\ X_k$, número de ocorrências para $A_1,\ A_2,\ A_3,\ \ldots,\ A_k$, em n tentativas independentes é (FONSECA; MARTINS, 2006):

$$P\left(x_1,\ x_2,\ x_3,\ldots\ x_k;\ p_1,\ p_2,\ p_3,\ldots\ p_k;\ n\right) = \frac{n!}{x_1!x_2!x_3!\ \ldots\ x_k!}p_1^{x_1}p_2^{x_2}p_3^{x_3}\ \ldots\ p_k^{x_k}$$

Para k=2 tem-se a distribuição binomial.

A **esperança** e a **variância** de uma distribuição multinomial são dadas, respectivamente, por (DANTAS, 2004):

$$E(X_i) = np_i$$

e

$$Var(X_i) = np_i(1 - p_i).$$

Exemplo 5.6.8. *Em uma fábrica, uma peça é considerada "boa"(tipo 1) se uma de suas dimensões L está entre l_1 e l_2; "recuperável"(tipo 2) se $L > l_1$ e "perdida"(tipo 3) se $L < l_1$. Numa sequência de produção de 12 peças resultou em 5 peças do tipo 1, 4 do tipo 2 e 3 do tipo 3. Retira-se, aleatoriamente, uma peça dentre as 12 e mede-se sua dimensão, devolvendo a peça para o conjunto (com reposição). São realizadas 6 ensaios dessa natureza.*
Tem-se:

- $p_1 = \frac{5}{12}$ *é a probabilidade de que a peça retirada seja do tipo 1;*
- $p_2 = \frac{4}{12}$ *é a probabilidade de que a peça retirada seja do tipo 2;*
- $p_3 = \frac{3}{12}$ *é a probabilidade de que a peça retirada seja do tipo 3.*

A probabilidade de entre as seis peças obsevadas obter 3 peças do tipo 1, 2 do tipo 2 e 1 do tipo 3 é:

$$P\left(x_1,\ x_2,\ x_3\right) = \frac{n!}{x_1!x_2!x_3!}p_1^{x_1}p_2^{x_2}p_3^{x_3}$$

$$P\left(x_1,\ x_2,\ x_3\right) = \frac{6!}{3!2!1!} \times \left(\frac{5}{12}\right)^3 \times \left(\frac{4}{12}\right)^2 \times \left(\frac{3}{12}\right)^1$$

$$P\left(x_1,\ x_2,\ x_3\right) \cong 0,1206\ ou\ 12,06\%.$$

5.6.4 Distribuição Geométrica

Numa sequência de provas de Bernoulli independentes, o número de "falhas"antes do acontecimento do primeiro "sucesso"tem uma distribuição denominada **geométrica** com parâmetro p, sendo p a probabilidade de ocorrer "sucesso"(DANTAS, 2006, p.145).

Seja X uma variável aleatória discreta (número de "falhas"antes de um "sucesso"). Esta distribuição é dada por:

$$P\left(x,\ p\right)=\begin{cases} p.(1-p)^x, & se\ \ x=\{0,\ 1,\ 2,\ 3,\ \ \dots,\ \ n,\dots\} \\ \\ 0, & se\ \ caso\ contrário. \end{cases}$$

Exemplo 5.6.9. *Se a probabilidade de certo experimento dar reação "positiva"for igual a* $0,4$*, qual será a probabilidade de que menos de 5 reações"negativas"ocorram antes da primeira "positiva"?*

Neste caso a variável aleatória X poderá assumir os valores $\{0,\ 1,\ 2,\ 3,\ 4\}$. *Esta distribuição pode ser representada pela Tabela 15.*

x	P(X=x)	
0	q	$=0,4$
1	p.q	$=0,4.0,6^1=0,24$
2	$p.q^2$	$=0,4.0,6^2=0,144$
3	$p.q^3$	$=0,4.0,6^3=0,864$
4	$p.q^4$	$=0,4.0,6^4=0,05184$
Total	$q+p.q+p.q^2+p.q^3+p.q^4$	$=0,92224$

Tabela 15: A distribuição geométrica para $x=4$ *e* $p=0,4$.

Portanto a probabilidade de que menos de 5 reações "negativas"ocorram antes da primeira "positiva"é de 92,22% aproximadamente.

As probabilidades dessa distribuição sempre diminuem numa progressão geométrica à medida que o valor da variável aleatória discreta X aumenta, veja Tabela 15 . Daí o nome **distribuição geométrica**.

Deve-se tomar o cuidado na interpretação de experimentos que existe uma longa sequência de "falhas"numa série de provas de Bernoulli independentes com $p=0,5$. Contrário ao pensamento intuitivo de alguns jogadores, isto não tem efeito algum nos resultados futuros. Por exemplo, ao lançar uma moeda equilibra 9 vezes se os resultados foram 9 caras consecutivas, a probabilidade de se obter coroa na décima jogada continua sendo 0,5. Por outro lado, a probabilidade de obter, em 10 lançamentos de uma moeda, 9 caras nos primeiros lançamentos e coroa no último lançamento é dada por $0,5^9.0,5^1 \cong 0,00097$.

A **experança** e a **variância** de uma distribuição geométrica, veja (DANTAS, 2004, p. 146), são dadas, respectivamente, por:

$$E(X) = \frac{1-p}{p}$$

e

$$Var(X) = \frac{1-p}{p^2}.$$

Exemplo 5.6.10. *Um banco de sangue necessita de sangue do tipo O-Rh negativo. Seja p=0,1 a proporção de indivíduos na população com esta característica. As pessoas serão escolhidas ao acaso para serem examinadas se tem esse tipo de sangue. O número de pessoas examinadas antes de se encontrar a primeira pessoa com este tipo de sangue tem distribuição geométrica com parâmetro $p = 0,1$. A* **experança** *e a* **variância** *desta distribuição são, respectivamente:*

$$E(X) = \frac{1-0,1}{0,1} = 9$$

e

$$Var(X) = \frac{1-0,1}{0,1^2} = 90.$$

5.6.5 Distribuição Hipergeométrica

Seja um conjunto com N pessoas (M brasileiros e N-M ingleses). Extraímos uma amostra aleatória com n pessoas, **sem reposição**; a ordem de extração é irrelevante. Qual a probabilidade de examente x pessoas serem brasileiros? A resolução deste problema ajudará a entender a definição de uma **distribuição hipergeométrica** (DOWNING; CLARK, 2005, P.64). Note que:

- Há $\begin{pmatrix} N \\ n \end{pmatrix}$ maneiras de extrair a amostra;
- Há $\begin{pmatrix} k \\ x \end{pmatrix}$ maneiras de escolher os x brasileiros na amostra;

- Há $\begin{pmatrix} N-k \\ n-x \end{pmatrix}$ maneiras de escolher os $n-x$ ingleses.

Portanto há $\begin{pmatrix} k \\ x \end{pmatrix} \times \begin{pmatrix} N-k \\ n-x \end{pmatrix}$ possibilidades da escolha da amostra com k brasileiros e $n-x$ ingleses, e a probabilidade de isso ocorrer é:

$$\frac{\begin{pmatrix} M \\ x \end{pmatrix} \times \begin{pmatrix} N-k \\ n-x \end{pmatrix}}{\begin{pmatrix} N \\ n \end{pmatrix}}.$$

A **distribuição hipergeométrica** se aplica a experimentos como este, em que interessa-se calcular a probabilidade de selecionar x sucessos de k elementos rotulados como sucessos e $n-x$ falhas de $N-k$ rolulados como falhas ou defeitos quando uma amostra aleatória de tamanho n é retirada de uma população de N elementos (DOWNING; CLARK, 2002).

Exemplo 5.6.11. *Considere um grupo de 20 pessoas, sendo 12 brasileiras e 8 inglesas. Escolhe-se aleatoriamente 8 destas pessoas. Seja X a v.a. dada pelo "número de pessoas brasileiras escolhidas dentre as 8 escolhidas". A variável X poderá assumir os valores* $\{0, 1, 2, 3, \ \ldots \ 8\}$*. Assim tem-se uma variável aleatória com distribuição hipergeométrica com polução* $N = 20$ *de onde se retira uma amostra de tamanho* $n = 8$ *e nesta polução há k=12 possibilidades de sucesso. Então, de*

$$P(X = x_i) = \frac{\begin{pmatrix} k \\ x \end{pmatrix} \times \begin{pmatrix} N-k \\ n-x \end{pmatrix}}{\begin{pmatrix} N \\ n \end{pmatrix}}$$

tem-se:

- $P(X = 0) = \frac{\begin{pmatrix} 12 \\ 0 \end{pmatrix} \times \begin{pmatrix} 8 \\ 8 \end{pmatrix}}{\begin{pmatrix} 20 \\ 8 \end{pmatrix}} = \frac{1}{125970}$

- $P(X=1) = \dfrac{\binom{12}{1} \times \binom{8}{7}}{\binom{20}{8}} = \frac{96}{125970}$

- $P(X=2) = \dfrac{\binom{12}{2} \times \binom{8}{6}}{\binom{20}{8}} = \frac{1848}{125970}$

- $P(X=3) = \dfrac{\binom{12}{3} \times \binom{8}{5}}{\binom{20}{8}} = \frac{12320}{125970}$

- $P(X=4) = \dfrac{\binom{12}{4} \times \binom{8}{4}}{\binom{20}{8}} = \frac{34650}{125970}$

- $P(X=5) = \dfrac{\binom{12}{5} \times \binom{8}{3}}{\binom{20}{8}} = \frac{44352}{125970}$

- $P(X=6) = \dfrac{\binom{12}{6} \times \binom{8}{2}}{\binom{20}{8}} = \frac{25872}{125970}$

- $P(X=7) = \dfrac{\binom{12}{7} \times \binom{8}{1}}{\binom{20}{8}} = \frac{6336}{125970}$

- $P(X=8) = \dfrac{\binom{12}{8} \times \binom{8}{0}}{\binom{20}{8}} = \frac{495}{125970}$

Assim, esta distribuição pode ser representada graficamente como mostra a Figura 30.

N=20, n=8 e k=12.	
x	P(X=xi)
0	0,00001
1	0,00076
2	0,01467
3	0,09780
4	0,27507
5	0,35208
6	0,20538
7	0,05030
8	0,00393

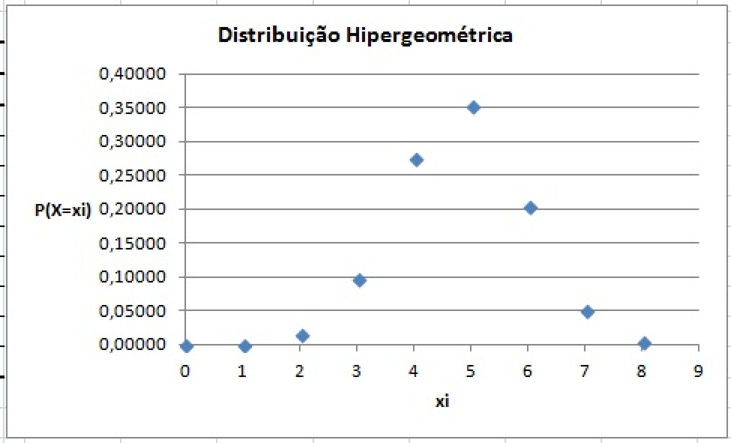

Figura 30: Disbribuição de Hipergeométrica de uma polução $N = 20$, amostra $n = 8$ e número possíveis de sucessos na população $k = 12$.

Exemplo 5.6.12. *Uma empresa comprou várias caixas contendo cada caixa 15 lâmpadas. A mesma decidiu fazer uma inspeção por amostragem* **sem reposição** *analisando 5 lâmpadas de cada caixa e aceitando a caixa se encontrar duas ou menos defeituosas. Seja 20% a probabilidade do produto ser defeituoso. Tem-se aqui uma variável aleatória com distribuição hipergeométrica com*

$$N = 15, \quad n = 5 \quad e \quad K = 20\% \quad de \ 15 = 3.$$

Assim a probabilidade de se aceitar uma caixa é dada por:

$$P(X = 0) + P(X = 1) + P(X = 2) =$$

$$= \frac{\begin{pmatrix} 3 \\ 0 \end{pmatrix} \times \begin{pmatrix} 12 \\ 5 \end{pmatrix}}{\begin{pmatrix} 15 \\ 3 \end{pmatrix}} + \frac{\begin{pmatrix} 3 \\ 1 \end{pmatrix} \times \begin{pmatrix} 12 \\ 4 \end{pmatrix}}{\begin{pmatrix} 15 \\ 3 \end{pmatrix}} + \frac{\begin{pmatrix} 3 \\ 2 \end{pmatrix} \times \begin{pmatrix} 12 \\ 3 \end{pmatrix}}{\begin{pmatrix} 15 \\ 3 \end{pmatrix}}$$

$$= 0,2637 + 0,4951 + 0,2198 \cong 0,9780 \ ou \ 97,80\%.$$

A **esperança** e a **variância** de uma distribuição hipergeométrica são dadas, respectivamente, por (BUSSAB; MORETTIN, 2006):

$$E(X) = \frac{nk}{N}$$

e

$$Var(X) = \frac{nk}{N} . \left(1 - \frac{k}{N}\right) . \left(\frac{N-n}{N-1}\right).$$

Exemplo 5.6.13. *Para o exemplo anterior, sendo* $N = 15,\ n = 5$ *e* $K = 3$, *tem-se:*

$$E(X) = \frac{5.3}{15} = 1 \ e$$

$$Var(X) = \frac{5.3}{15} . \left(1 - \frac{3}{15}\right) . \left(\frac{15-5}{15-1}\right) = \frac{4}{7}.$$

5.6.6 Distribuição de Poisson

Geralmente se conhece o número de sucessos, porém o número de fracassos ou número total de provas seria difícil de ser determinado, ou muitas vezes, não faria nehum sentido prático. Por exemplo pode-se, num determinado intervalo de tempo, anotar o número de veículos que passaram em um trecho de uma rodovia, porém, o número de carros que deixaram de passar não poderá ser determinado. Verifica-se então que há uma variável t e quando $t \to \infty$ a probabilidade de ocorrer um sucesso tende a aumentar.

Para Fonseca e Martins (2006, p.67), a distribuição de Poisson é uma forma limite de distribuição binomial, quanto N tende ao infinito e p tende a zero. Assim para se encontrar a expressão que dá $P(X = k, t)$, ou seja, a probabilidade de $X = k$ sucessos no intervalo de tempo t, pode-se calcular o limite de uma distribuição Binomial com parâmetros n e $\lambda.\frac{t}{n}$, onde λ é o coeficiente de proporcionalidade em que dado intevalo de tempo $\Delta t = \frac{t}{n}$ tem-se $P(X = 1, \Delta t) = \lambda.\frac{t}{n}$.

λ também pode ser interpretado como sendo a média da distribuição no intervalo de tempo (média esta sempre positiva) ou também como sendo a medida considerada.

Para encontrar a expressão que dá a probabilidade de k sucessos num intervalo de tempo t, deve-se admitir que a probabilidade de um sucesso num intervalo de tempo Δt seja proporcional à amplitude do intevalo e as ocorrências de sucessos em intervalos são independentes. Assim, a função de probabilidade da distribuição de Poisson (FONSECA; MARTINS, 2006), que dará a probabilidade de ocorrer examente k sucessos em um intevalo de tempo t, é dada por:

$$P(X = k) = \lim_{n \to \infty} C_n^x \left(\lambda.\frac{t}{n}\right) \left(1 - \lambda.\frac{t}{n}\right)^{n-x},$$

que resulta em:

$$P(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}, \quad (com\ k\ =\ 0,\ 1,\ 2,\ 3,\ \ldots).$$

Note que a distribuição de Poisson fica completamente caracterizada por um único parâmetro (BUSSAB; MORETTIN, 2006), a média do processo, pois nesta distribuição a média é igual a variância, ou seja,

$$E(X) = Var(X) = \lambda.$$

Assim, sabendo-se que uma v.a. tem resultados distribuídos segundo uma distribuição de Poisson e conhecendo o número médio de ocorrências por unidade de medida, podemos determinar a probabilidade de qualquer dos resultados possíveis nos intervalos para os quais deseja-se.

Exemplo 5.6.14. *Numa indústria há em média 3 acidentes por mês. A probabilidade de ocorrer um acidente no próximo mês é dada por:*

$$P(X = 2) = \frac{e^{-3}.3^2}{2!}$$

$$P(X = 2) = 0,049787.4,5 = 0,2204 \quad ou \quad P(X = 2) = 22,04\%.$$

Exemplo 5.6.15. *Uma central de telefones recebe em média 5 chamadas por minuto. Supondo que a distribuição de Poisson seja adequada nessa situação, a probabilidade de que esta central não receba chamadas durante um intervalo de um minuto é dada por:*

$$P(X = 0) = \frac{e^{-5}.5^0}{0!} = 0,0067 \quad ou \quad P(X = 0) = 0,67\%.$$

Veja que no exemplo 4.3.8, em que se deseja calcular a probabilidade do número de chamadas n chegar à uma central de telefones, tomou-se $\lambda = 1$.

Pode-se, a partir exemplo 5.6.15, obter a Tabela 16 fazendo os cálculos a partir de uma planilha eletônica, que mostra as probabilidades de obter exatamente K chamadas no intervalo de tempo de 1 hora.

Veja na Figura 31 a representação gráfica dessa distribuição de probabilidade.

Para Dantas (2004, p. 148-149), a Distribuição de Poisson é uma aproximação para a Distribuição Binomial quando o número de ensaios de Bernoulli tende ao infinito, porém de modo que $n.p = \lambda$ permanece constante. Neste caso, considera-se para cada inteiro n

k	P(X=k)=(Probabilidade de se obter exatamente k chamadas)
0	0,006738
1	0,033690
2	0,084224
3	0,140374
4	0,175467
5	0,175467
6	0,146223
7	0,104445
8	0,065278
9	0,36226
10	0,018133
11	0,008242
12	0,003434

Tabela 16: A distribuição de Poisson para $\lambda = 5$.

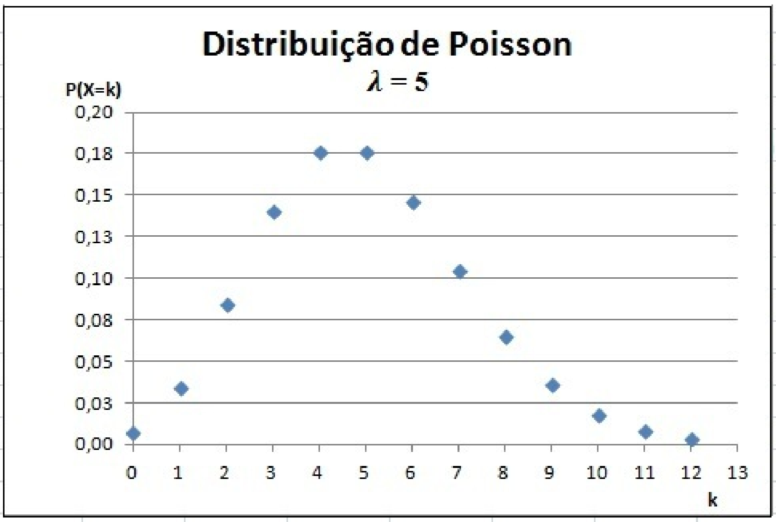

Figura 31: Disbribuição de Poisson para $\lambda = 5$.

um conjunto de n ensaios de Bernoulli em que a probabilidade de sucesso em cada um dos n ensaios é p. Assim para $P(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}$, $(com\ k\ =\ 0,\ 1,\ 2,\ 3,\ \ldots)$. tem-se na realidade não uma distribuição de Poisson, mas sim uma família de distribuições de Poisson dependendo do parâmetro λ.

5.6.7 Distribuição Binomial Negativa ou Distribuição de Pascal

Se ensaios de Bernoulli independentes com probabilidade p de ocorrer "sucesso"são repetidos X vezes até que um número fixo x de "sucessos"tenha ocorrido, X terá distribuição Binomial Negativa. Essa distribuição pode ser considerada como a soma de variáveis aleatórias $X_1,\ X_2,\ \ldots\ X_n$, cada uma com a mesma Distribuição Geométrica.

Considere um experimento em que as propriedades são as mesmas que aquelas enumeradas para o experimento Binomial, com excessão as tentativas serão repetidas até que número fixo de sucessos venha a ocorrer. Portanto em vez de calcular a probabilidade de x sucessos em n tentativas, onde n é fixo, interessa-se, para a Distribuição Binomial Negativa, calcular a probabilidade de que o k-ésimo sucesso ocorra na x-ésima tentativa. O número de tentativas X para produzir k sucessos em um experimento binomial negativo é chamado variável aleatória discreta binomial negativa. Sua função de probabilidade é dada por:

$$P\left(x;\ k,\ p\right) = \begin{cases} C_{x-1}^{k-1}\ p^x.(1-p)^{x-k}, & se\ \ x = \{k,\ k+1,\ k+2,\ k+3,\ \ldots\}; \\ \qquad\qquad para\ 0\ \leq p \leq 1. & \\ & \\ 0, \quad se\ \ caso\ contrário. & \end{cases}$$

Exemplo 5.6.16. *No arremesso de uma moeda equilibrada, qual seria a probabilidade de ocorrer a segunda "cara"no quarto arremesso?*
Pode-se verificar neste exemplo que:

$$p = \frac{1}{2},\ x = 4,\ e\ k = 2.$$

Então:

$$P\left(2;\ 2,\ \frac{1}{2}\right) = C_{2-1}^{4-1}\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^{4-2}$$
$$P\left(2;\ 2,\ \frac{1}{2}\right) = C_1^3\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^2$$
$$P\left(2;\ 2,\ \frac{1}{2}\right) = \frac{3}{16}$$
$$P\left(2;\ 2,\ \frac{1}{2}\right) = 0,1875\ ou\ 18,75\%$$

Essa probabilidade fica mais fácil de calcular se for possível obter o espaço amostral como segue.

$$\Omega = \left\{ \begin{array}{ccccc} & & KKCC & & \\ & & KCKC & & \\ & CKKK & KCCK & CCCK & \\ & KCKK & CKCK & CCKC & \\ & KKCK & CKKC & CKCC & \\ KKKK & KKKC & CCKK & KCCC & CCCC \end{array} \right\}.$$

Note que os sucessos são KKCC, KCKC e CKKC em um total de 16 possibilidades.

A distribuição binomial negativa ou distribuição de Pascal é uma distribuição de probabilidade discreta. Note que esta distribuição indica o número de tentativas necessárias para obter k sucessos de igual probabilidade p ao fim de n experiências, sendo a última tentativa um sucesso. Caso queira calcular o número esperado de tentativas para se obter k sucessos, dado o valor de p, é necessário que se calcule a média dos valores da variável aleatória binomial negativa.

Se X é uma variável aleatória discreta binomial negativa então a **esperança** e a **variância** são dadas, respectivamente, por:

$$E(X) = \frac{k}{p}$$

e

$$Var(X) = \frac{x.(1-p)}{p^2}.$$

Exemplo 5.6.17. *A probabilidade de que um experimento seja bem sucedido é* 80%. *Se o experimento for repetido até que quatro resultados bem sucedidos tenham ocorridos, qual será o número esperado de repetições necessárias?*

O que se quer calcular é a esperanda da variável aleatória discreta dada. Como

$$E(X) = \frac{k}{p},$$

então:

$$E(X) = \frac{4}{0,8} = 5.$$

Uma comparação importante entre as disbribuições Binomial e a Binomial Negativa é que em amabos os casos estamos tratando com provas de Bernoulli. A distribuição Binomial surge quando lida-se com números fixados n dessas provas e preocupa-se com o número de sucessos que venha a ocorrer. A distribuição Binomial Negativa é econtrada quando o número de sucessos a ser obtido é pré-fixado e daí registra-se o número de provas de Bernoulli necessárias. Portanto em uma distribuição Binomial o número de amostragens é pré-determinado, número de sucessos é aleatório e em uma distribuição Binomial Negativa o número de sucessos é pré-determinado e o número de amostragens é aleatório.

6 CONSIDERAÇÕES FINAIS

O fato de a incerteza estar sempre presente em muitas situações do cotidiano, o ser humano, em suas múltiplas relações e interações, faz com que crianças e adultos tenham um conhecimento intuitivo sobre temas originados dessa incerteza, tais como: o acaso, a aleatoriedade e a probabilidade. Esse conhecimento está relacionado com sorte, destino, crença, e outros, adquiridos a partir de situações do cotidiano, como a participação de um sorteio, adivinhações ou previsões, sem, no entanto, adotar nenhuma metodologia científica, originando assim, uma **ideia intuitiva** de probabilidade.

O cálculo de probabilidade está ligado a experimentos, chamados experimentos aleatórios, que não se sabe ao certo o resultado, mas é possível fazer uma estimativa das chances de sucesso ou fracasso deste evento.

Durante a história da humanidade surgiram três significados diferentes para probabilidade e a partir desses é que se originaram três definições para calcular probabilidade: a definição clássica, a definição frequentista e a definição subjetiva. A **definição clássica** de probabilidade é a que está presente nos livros do ensino fundamental e médio. Esta definição relaciona a probabilidade de um sucesso ocorrer em um número finito de modalidades, o que faz com que a probabilidade seja definida como uma fração cujo numerador é o número de casos favoráveis ao sucesso e o denominador da fração é o número de casos possíveis. A partir desta definição é possível enunciar um grande número de propriedades que são muito úteis na resolução de vários problemas em que se deseja saber as chances de um evento ocorrer. Mas devido esta definição tomar como base o número de elementos de dois conjuntos, é necessário, para se conhecer a probabilidade de um evento, que se conte o número de elementos destes conjuntos. Daí o uso de técnicas de contagem no cálculo de probabilidade se tornar a ferramenta fundamental. No entanto, o número de elementos destes conjuntos pode ser impossível de calcular e essa fração que é atribuída à probabilidade pode não existir. Assim, há situações em que não se aplica a definição clássica.

Um experimento é repetido, em condições semelhantes, um número n finito de vezes e, a partir de uma análise em seus resultados pode-se chegar a conclusão que um evento tem uma probabilidade de ocorrência que tende a um determinado valor. A partir da necessidade de calcular esta probabilidade é que surgiu a **definição frequentista** de probabilidade. Esta definição toma como base a frequência relativa do evento. Então define-se a probabilidade de um evento como sendo a fração $\frac{n_A}{n}$ em que n_A é a frequência com que o evento A surge nas n repetições do experimento, quando o número n tende a um número muito grande. Um problema prático é que, com o enfoque frequencial, nunca se obtém o valor exato da probabilidade, sendo somente esta razão uma estimativa da probabilidade real. Por outro lado, há experimentos que são impossíveis de serem realizados exatamente nas mesmas condições. Também é difícil saber com certeza qual o número de vezes que se deve repetir o experimento para se obter uma boa estimativa da probabilidade. Outra restrição da definição frequentista, e talvez a que mais a compromete, é que certos experimentos, por exemplo, no campo da economia, da história, no esporte, apesar de serem aleatórios, não se repetem em condições semelhantes e, de acordo com o significado frequentista, não seria possível aplicar a teoria das probabilidades para seu estudo.

Embora, o significado frequencial amplia o campo de aplicação desta teoria, ainda há situações em que ela não se aplica. A partir da definição de probabilidade condicional e a regra de Bayes, que permite transformar as probabilidades "*a priori*"(probabilidades que se tem antes de realizar o experimento) em probabilidades "*a posteriori*"(que incorporam a informação dos dados analisados) é que surge a **definição subjetiva** de probabilidade. Esta definição se baseia em medir o grau de crença pessoal, baseados no conhecimento e na experiência da pessoa. Neste caso, a probabilidade de sucesso de um dado experimento está condicionada a certo conjunto de conhecimentos e pode ser, portanto, diferentes para diferentes pessoas. Assim, o enfoque subjetivo não necessita de contar os elementos dos conjuntos e nem de repetir o experimento em condições "quase idênticas"para dar sentido a probabilidade, e sim, aplica-se o conhecimento e o grau de crença pessoal, como é necessário em um diagnóstico médico (avalição de um paciente), na economia (oscilação dos preços das ações na bolsa), etc. Uma dificuldade inicial deste enfoque subjetivo foi de criar regras matemáticas que atribuíssem números para as probabilidades que expressam graus de crenças pessoais. Para que a teoria construída sobre esta ou outras questões subjetivas (opiniões pessoais) tenha consistência, seja coerente, algumas regras gerais e de comportamentos racionais foram estabelecidas. Estas regras são baseadas em alguns axiomas conhecidos como axiomas de Kolmogorov.

A partir das ideias de Kolmogorov, a probabilidade passou a ser simplesmente modelos matemáticos que podem ser usados para descrever e interpretar fenômenos da realidade e fenômenos aleatórios. Estes modelos têm mostrado sua importância e utilidade em quase todos os campos da atividade humana, como em pesquisas científicas, na economia, na gestão de empresas, em departamentos de recursos humanos, na pscilologia, etc.

Ideia intuitiva de probabilidade	Sorteios e advinhações.	Opinião, crença, imprevisibilidade.	Sorte, destino.
Siginificado de Probabilidade	Campos de Problemas	Definições e Propriedades	Alguns conceitos relacionados
Clássica	Cálculo de Esperança ou riscos em jogos de azar.	Razão entre casos favoráveis e possíveis. Equiprobabilidade de eventos simples.	Esperança, equiprovável, independência.
Frequentista	Estimação de parâmetros em problemas.	Limite das frequências relativas. Caráter objetivo baseado na evidência empírica.	Frequência retativa. Espaço amostral. Variável aleatória. Distribuição de probabilidades.
Subjetiva	Melhor conhecimento sobre sucessos e incertezas atribuídos ao experimento.	Caráter subjetivo. Voloriza a experiência e conhecimento pessoal.	Probabilidade condicional. Probabilidades a priori e a posteriori.
Axiomática	Quantificar as incertezas de resultados em experimentos aleatórios abstratos.	Função mensurável.	Espaço amostral. Espaço de probabilidadeses.

Tabela 17: Elementos que caracterizam as diferentes concepções de probabilidade.

A Tabela 17 resume o que foi apresentado anteriormente. Os diferentes siginificados de probabilidade que deram origem as definições que ainda persistem, são utilizadas na prática e são ensinadas nos cursos fundamentais e de graduação, geralmente ligados à disciplina Estatística. As diferenças não estão somente nas definições, mas também em campos de aplicação e métodos de cálculos. A definição subjetiva, em muitas situações, se confude com a axiomática. Os axiomas foram criados com o objetivo de criar regras de comportamentos para todas as definições de probabilidade. Qualquer pessoa, seja um cidadão comum, um educador ou pesquisador, poderá elaborar um modelo probabilístico, desde que este modelo obedeça os axiomas de Kolmogorov.

A teoria das probabilidades se tornou tão importante que na maioria dos cursos de graduação este conteúdo está inserido em suas ementas, geralmente ligados à disciplina Estatística. A proposta do PROFMAT é que as dissertações dos alunos concluintes do curso de mestrado em matemática fossem elaboradas com vista na criação de materiais que possam auxiliar o professor na aplicação e prática docente, e que possam ser utilizado em sala de aula em algum momento. Como professor das séries finais do ensino médio e de cursos superiores noturnos de carga horária reduzida como Administração, Ciências Contábeis, Psicologia, Sistema de Informações, Tecnologia em Cafeicultura, Tecnologia em Agronegócio, no Centro Universitário do Cerrado - UNICERP - Patrocínio - MG, surgiu a necessidade de escrever sobre a teoria das probabilidades, e tentar escrever um texto que possa ser usado pelos professores desta disciplina em suas aulas. Geralmente, os cursos citados, possuem carga horária de 60 horas aulas semestrais e que além da Estatística descritiva nos programas consta regressão, correlação e probabilidade, o que torna difícil cumprir esse programa.

Outro fato importante e que justifica este estudo é que os alunos chegam no curso superior com uma bagagem muito pequena de conhecimentos e portanto, o tratamento que foi dado neste texto (citar as três definições de probabilidade com exemplos simples e com resolução, enunciar a definição de probabilidade condicional e suas propriedades a partir de um exemplo, a preocupação de cada propriedade apresentada ter um exemplo de aplicação com a resolução e finalmente citar alguns exemplos de distribuição de probabilidade de variável aleatória discreta) foi com o objetivo de criar um material que possa ser utilizado nas aulas daqueles cursos de graduação, paralelo com um livro didático, de modo a facilitar o entendimento desta disciplina que em muitas situções amedronta os alunos pelo seu grau de dificuldade.

Referências

1 BATANERO, C. **Significados de la Probabilidade en la Educación Secundária. Revista Latinoamerica de Investigation en Matematica Educativa**. México. Comité Latinoamericano de Matemática Educativa. V. 8, n. 3 pp. 247-263, nov. 2005.

2 BUSSAB, W. O.; MORETTIN, P. A.; **Estatística Básica**. 5. ed. São Paulo: Saraiva, 2006.

3 COSTA, S. F. **Introdução ilustrada à Estatística**. 4. ed. São Paulo: HARBRA LTDA, 2005.

4 DANTAS, C. A. B. **Probabilidade: um curso introdutório**. 2. ed. São Paulo: EDUSP, 2004.

5 DOWNING, D.; CLARK, J. **Estatística aplicada**. 2. ed. São Paulo: Saraiva, 2002.

6 GONGALVES, M. C. **Concepções de Professores e Ensino de Probabilidade na Escola Básica**. São Paulo: MESTRADO PUC, 2004.

7 FONSECA, J. S.; MARTINS, G. A. **Curso de Estatística**. 6. Ed. São Paulo: Atlas, 2006.

8 JAMES, B. R. **Probabilidade: um curso em nível intermediário**. Rio de Janeiro: IMPA, 2004.

9 MAGALHÃES, M. N. **Probabilidade e Variáveis Aleatórias**. 2. ed. São Paulo: EDUSP, 2006.

10 MORGADO, A.C. et al. **Análise Combinatória e Probabilidade**. Rio de Janeiro: IMPA, 1991.

11 MURTEIRA, B.J.F. **Probabilidade e Estatística**. 2. Edição. Portugal: McGraw-Hill, 1990.

12 SÁNCHEZ, E.; VALDEZ, J.C. **La cuantificación del azar: una articulación de las definiciones subjetiva, frecunecial y clásica de probabilidad**, Cinvestav-IPN. México.

7 ANEXO I - OPERAÇÕES ENTRE EVENTOS

7.1 Definições

Considere o espaço amostral $\Omega = \{1, 0\}$. Os conjuntos $\{\ \}$, $\{0\}$, $\{1\}$ e $\{1, 0\}$ são todos os subconjuntos (ou eventos) de Ω. O conjunto destes subconjuntos é denominado conjuntos das partes de Ω e representado por $\mathcal{P}(\Omega)$. O evento $\{\ \}$ que também pode ser representado por ϕ é o evento impossível e $\{1, 0\}$, que é o próprio espaço amostral, é o evento certo.

Assim o **Conjunto das Partes do Espaço Amostral $\boldsymbol{\Omega}$**, é definido como sendo:

$$\mathcal{P} = \{A; A \subset \Omega\}.$$

Exemplo 7.1.1. *Considere $\Omega = \{1, 2, 3\}$ o conjunto das partes de Ω é a coleção de todos os seus subconjuntos. Então*

$$\mathcal{P} = \{\emptyset, \Omega, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}\}.$$

Se o conjunto A for finito com n elementos, então o conjunto das partes tem 2^n elementos, pois pode-se entender como a formação de subconjuntos com zero elementos, subconjuntos com 1 elementos, ..., subconjuntos com n elementos. Daí:

$$C_n^0 + C_n^1 + \ldots + C_n^n = 2^n$$

.

Diz-se que o evento A implica o evento B, e representa-se por $A \subset B$, se para todo $w \in A$ tem-se $w \in B$. Isto equivale a dizer que a ocorrência de A garante inevitavelmente a ocorrência de B.

Dois eventos são iguais se $A \subset B$, e $B \subset A$.

A **união** de dois eventos A e B, representada por $A \bigcup B$, é o evento que ocorre se pelo menos um deles ocorre. Então $A\bigcup B = \{w \in \Omega; w \in A \ ou \ w \in B.\}$. Veja na Figura 32 que da mesma forma que se utiliza diagramas para representar conjuntos pode-se usar também diagramas para representar eventos.

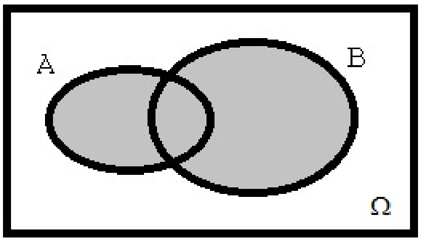

Figura 32: Diagrama representando a união dos conjuntos A e B. Fonte: próprio autor.

A **interseção** de dois eventos A e B, representada por $A\bigcup B = \{w \in \Omega; w \in A \ e \ w \in B.\}$, é o evento que ocorre se ambos ocorrem. A Figura 33 representa o evento $A\bigcap B$.

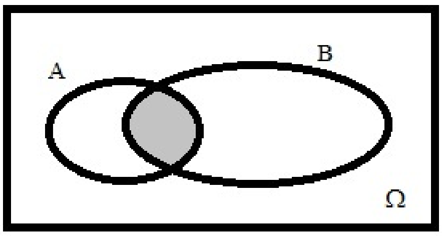

Figura 33: Diagrama representando a interseção dos conjuntos A e B. Fonte: próprio autor.

Se A e B são tais que a ocorrência de A implica na não ocorrência de B, diz-se que A e B são **disjuntos** e a união destes eventos é denominada uma **união disjunta**, veja Figura 34, ou ainda, se os eventos A e B não podem ocorrer simultaneamente estes eventos são ditos **mutuamente exclusivos**. Isto equivale a dizer que $A \cap B = \phi$.

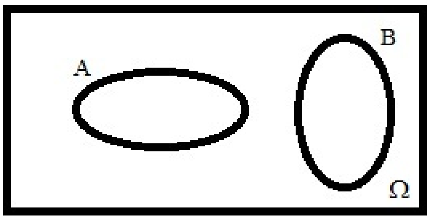

Figura 34: Conjuntos A e B disjuntos ou eventos mutuamente exclusivos. Fonte: próprio autor.

De forma geral, se $A_i \cap A_j = \phi$, com $i \neq j$, dizemos que os eventos são mutuamente exclusivos.

A **diferença** entre dois eventos A e B, representada por $A - B$ é o evento em que A ocorre e B não ocorre (veja a Figura 35). Daí $A - B = \{w \in \Omega\ ;\ w \in A\ e\ w \notin B\}$.

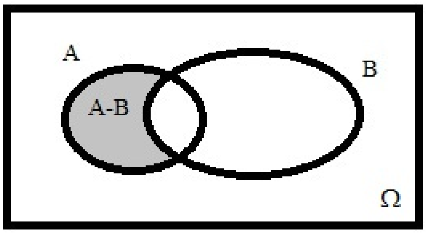

Figura 35: Diferença entre os eventos A e B. Fonte: próprio autor.

O evento **complementar** do evento A, representado por A^c, veja Figura 36, é o evento que ocorre quando A não ocorre, assim tem-se $A^c = \{w \in \Omega\ ;\ w \in A\}$ ou $A^c = \Omega - A$.

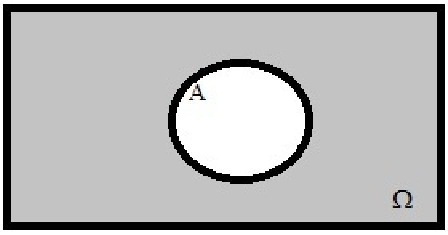

Figura 36: O complementar do evento A, dado por A^c. Fonte: próprio autor.

Suponha os eventos $A_1,\ A_2,\ A_3\ \ldots\ , A_n$. O evento $\bigcup_{i=1}^{n} A_i$ é o evento quando pelo menos um dos eventos A_i (para $i = 1,\ 2,\ 3,\ \ldots$) ocorre.

O evento $\bigcap_{i=1}^{n} A_i$ é o evento quando todos os eventos A_i (para $i = 1,\ 2,\ 3,\ \ldots$) ocorre.

Dados três eventos A, B e C tem-se

$$A \cup B \cup C = \{w \in \Omega\ ;\ w \in A\ ou\ w \in B\ ou\ w \in C\},$$

ou seja, $A \cup B \cup C$ ocorre se A ou B ou C ocorrer. E,

$$A \cap B \cap C = \{w \in \Omega\ ;\ w \in A\ e\ w \in B\ e\ w \in C\},$$

ou seja, $A \cap B \cap C$ ocorre se ocorrer A e B e C simultaneamente.

7.2 Princípios

1. **Princípio aditivo I** Seja $|A|$ o número de pontos amostrais do evento A. Se A e B são eventos finitos tais que $A \cap B = \phi$, tem-se que o número de elementos ou de pontos amostrais do evento $A \cup B$, representado por $|A \cup B|$, é dado por:

$$|A \cup B| = |A| + |B|$$

2. **Princípio aditivo II** Se $A_1,\ A_2,\ A_3,\ \dots\ , A_n$ forem eventos 2 a 2 mutuamente exclusivos, ou seja, $A_i \cap A_j = \phi$, com $i \neq j$, então:

$$|A_1 \cup A_2 \cup A_3 \cup \ldots \cup A_n| = |A_1| + |A_2| + |A_3| + \ \dots\ + |A_n|$$

3. **Princípio da inclusão e exclusão para 2 conjuntos (eventos)** Se A e B são conjuntos (eventos) finitos tem-se que o número de elementos ou de pontos amostrais do evento $A \cup B$, representado por $|A \cup B|$, é dado por:

$$|A \cup B| = |A| + |B| - |A \cap B|$$

Demonstração:
Seja os eventos $(A-B),\ (A\cap B)$ e $(B-A)$. Note que esses conjuntos são mutuamente exclusivos. Então:

I) $(A \cup B) = (A - B) \cup (A \cap B) \cup (B - A)$;

II) $A = (A - B) \cup (A \cap B)$;

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

Pelo princípio aditivo tem-se:

- em $I)\ |A \cup B| = |A - B| + |A \cap B| + |B - A|$;
- em $II)\ |A| = |A - B| + |A - B|;\ e$
- em $II)\ |B| = |B - A| + |A \cap B|$.

Logo,

$$|A \cup B| = |A - B| + |A \cap B| + |B - A| \Leftrightarrow$$

$$|A \cup B| = |A - B| + |A \cap B| + |B - A| + |A \cap B| - |A \cap B| \Leftrightarrow$$

$$|A \cup B| = |A| + |B| - |A \cap B|$$

Análogo, pode-se obter para os eventos A, B e C, que:

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

4. **Princípio da Inclusão e Exclusão para n conjuntos.** Seja Ω o espaço amostral de um experimento, e, $A_1,\ A_2,\ A_3,\ \ldots\ , A_n$ eventos (subconjuntos) de Ω. Denomina-se:

- $S_0 = |\Omega|$
- $S_1 = |A_1| + |A_2| + |A_3| +\ \ldots\ + |A_n| = \sum_{i=1}^{n} |A_i|$
- $S_2 = |A_1 \cap A_2| + |A_1 \cap A_3| +\ \ldots\ + |A_1 \cap A_n| +\ \ldots\ + |A_{n-1} \cap A_n| =$ $= \sum_{1 \leq i < j \leq n}^{n} |A_i \cap A_j|.$
- $S_3 = |A_1 \cap A_2 \cap A_3| + \ldots + |A_1 \cap A_2 \cap A_n| +\ \ldots\ |A_(n-2) \cap A_{n-1} \cap A_n| =$ $= \sum_{1 \leq i < j < k \leq n}^{n} |A_i \cap A_j \cap A_k|.$

$$\vdots$$

- $Sn = |A_1 \cap A_2 \cap\ \ldots\ \cap A_n| = |A_1 \cap A_2 \cap\ \ldots\ \cap A_n|.$

Daí, demonstra-se que:

i) O número de elementos de Ω que pertencem a exatamente p $(p \leq n)$ dos conjuntos $A_1,\ A_2,\ A_3,\ \ldots\ , A_n$ é:

$$a_p = \sum_{k=o}^{n-p} (-1)^k C_{p+k}^{k} S_{p+k}.$$

ii) O número de elementos de Ω que pertencem a pelo menos p $(p \leq n)$ dos conjuntos $A_1,\ A_2,\ A_3,\ \ldots\ , A_n$ é:

$$b_p = \sum_{k=o}^{n-p} (-1)^k C_{p+k-1}^{k} S_{p+k}.$$

iii) O número de elementos do conjunto Ω é:

$$|\Omega| = |A_1 \cup A_2 \cup\ \ldots\ \cup A_n| = S_1 - S_2 + S_3 -\ \ldots + (-1)^{n-1} S_n$$

7.3 Propriedades

Sejam A, B e C eventos do espaço amostral Ω.

a) $A \cup \phi = A \; e \; A \cap \phi = A$

b) $A \subset B \Leftrightarrow A \cup B = B$

c) $A \subset B \Leftrightarrow A \cap B = A$

d) $(A \cup B) \cup C = A \cup (B \cup C)$

e) $(A \cap B) \cap C = A \cap (B \cap C)$

f) $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$

g) $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$

h) $A - B = A \cap B^c$

i) $A \cup A^c = \Omega, \; A \cap A^c = \phi, \; \phi^c = \Omega, \; e \; \Omega^c = \phi$

j) $(A^c)^c = \Omega$

k) $A \subset B \Leftrightarrow B^c \subset A^c$

l) $(A \cup B)^c = A^c \cap B^c$ (fórmula de Morgan).

m) $(A \cap B)^c = A^c \cup B^c$ (fórmula de Morgan).

Aqui será feito algumas demonstrações.

f) Para demonstrar essa igualdade precisa-se demonstrar que todo elemento pertencente ao lado esquerdo pertence ao lado direito e vice-versa. Logo, se $w \in (A \cup B) \cap C$ então $w \in (A \cup B)$ e $w \in C$. Daí decorre que $w \in A$ ou $w \in B$ e $w \in C$, e portanto, $w \in A$ e $w \in C$ ou $w \in B$ e $w \in C$, que implica em $(A \cap C) \cup (B \cap C)$. Fazendo a leitura de trás para frente verifica-se que são verdadeiras, daí conclui-se que a igualdade é verdadeira.

k) $(\Rightarrow)$ Hipótese: $A \subset B$. Assim tome $w \in B^c$. Isto implica $w \notin B$, o que implica (por hipótese) $w \notin A$. Logo $w \in A^c$, então $B^c \subset A^c$.
$(\Leftarrow)$ Hipótese: $B^c \subset A^c$. Usando o que se provou em $(\Rightarrow)$ temos que: $(A^c)^c \subset (B^c)^c \Rightarrow A \subset B$.

l) Temos que $w \in (A \cup B)^c \Leftrightarrow w \notin (A \cup B) \Leftrightarrow w \notin A$ e $w \notin B \Leftrightarrow w \in A^c$ e $w \in B^c \Leftrightarrow w \in (A^c \cap B^c)$. Logo, a igualdade é verdadeira.

m) Seja $w \in (A \cap B)^c$. Então $w \notin (A \cap B)$, o que implica $w \notin A$ e $w \notin B$, e que por sua vez implica em $w \in A^c$ ou $w \in B^c$, ou seja, $w \in (A^c \cup B^c)$. Partindo de $w \in (A^c \cup B^c)$. Fazendo as implicações inversas conclui-se que a igualdade é verdadeira.

8 ANEXO II - ANÁLISE COMBINATÓRIA

8.1 Princípio Multiplicativo

Se um evento A pode ocorrer de m maneiras distintas e se, para cada uma dessas maneiras outro evento B pode ocorrer de n maneiras distintas, então o número de maneiras de ocorrer o evento A seguido do evento B é $m \times n$.

Exemplo 8.1.1. *Uma empresa de ônibus transporta passageiros que são coletados em 10 estações. Quantos tipos distintos de passagens existem em circulação, sabendo que cada bilhete contém impressos apenas a estação de partida e a estação de chegada? (Suponha que os ônibus sejam todos iguais).*
Deve-se primeiramente saber que para escolher a estação de origem há 10 opções e, para escolher a estação de destino há 9 opções. Pelo princípio multiplicativo tem-se $10 \times 9 = 90$.

8.2 Permutação - P_n

Considere o conjunto A $= \{a_1,\ a_2,\ a_3,\ \dots\ a_n\}$. Chama-se permutações dos n elementos de A, qualquer sequência formada pelos n elementos de A.

Exemplo 8.2.1. *Seja o conjunto $A = \{a,\ b,\ c,\ d\}$. Todas as permutações possíveis de A são:*

abcd	bcda	cabd	dabc
abdc	bcad	cadb	dacb
adcb	bacd	cbda	dcba
adbc	badc	cbad	dcab
acdb	bdac	cdab	dbca
acbd	bdca	cdba	dbac

Tabela 18: Permutações dos elementos do conjunto $\{a,\ b,\ c,\ d\}$.

O cálculo do número de permutações de um conjunto com n elementos é dado por $P_n = n!$. A Tabela 19 indica alguns valores de $n!$.

n	0	1	2	3	4	5	6	7	8	9	10
n!	1	1	2	6	24	120	720	5040	40320	362880	3628800

Tabela 19: Número de permutações dos elementos de conjunto com n elementos: n!.

Exemplo 8.2.2. *Um locutor de rádio tem disponível somente 10 músicas para fazer seu programa diário. Em quanto tempo no mínimo ele poderá tocar as mesmas 10 músicas sem repetir uma sequência? As 10 músicas podem ser ordenadas em 10! =3628800 sequências distintas. Logo o locutor terá 3628800 dias, o que corresponde quase 100 séculos.*

8.3 Permutações com elementos repetidos.

Se se tem um conjunto com n elementos, dos quais n_1 são iguais a a_1, n_2 são iguais a a_2, n_3 são iguais a a_3, $\ldots$, n_r são iguais a a_r, o número de permutações possíveis é dado por:

$$P_n^{n_1,\ n_2,\ n_3,\ \ldots,\ n_r} = \frac{n!}{n_1! \times n_2! \times n_3! \times\ \ldots\ \times n_r!}$$

Exemplo 8.3.1. *Considere o conjunto formado pelas letras da palavra ARARA. Quantas permutações distintas são possíveis. Como a palavra possui 5 letras tem-se* $n = 5$ *(total de elementos do conjunto). Mas há 2 letras A repetidas* $n_1 = 2$ *e 3 letras R repetidas* $n_2 = 3$. *Portanto:*

$$P_n^{n_1,\ n_2} = \frac{n!}{n_1! \times n_2!} = \frac{5!}{2! \times 3!} = 10$$

Exemplo 8.3.2. *O número de permutações dos elementos do conjunto formado pelas letras da palavra ESTATÍSTICA é:*

$$P_{11}^{2,\ 3,\ 2,\ 2} = \frac{11!}{2! \times 3! \times 2! \times 2!} = 831600.$$

Exemplo 8.3.3. *Uma pessoa quer comprar 6 empada numa lanchonete. Há de camarão, frango, legumes e palmito. Sabendo que podem ser compradas de 0 a 6 empadas de cada tipo, de quantas maneiras diferentes esta compra pode ser feita?*
*Considerando que cada uma das 6 empadas seja representada por * e que | seja uma maneira de escolher a quantidade de cada tipo de empada a ser compradas, então*

$$* \ * \ * \,|\ \ |\, * \ * \,|\, *$$

representa a opção de comprar 3 empadas de camarão, 0 de frango, 2 de legumes e 1 de palmito. Outra opção seria

$$|\, * \ * \ * \ * \,|\ \ |\, * \ *$$

que representa a opção de comprar 0 empadas de camarão, 4 de frango, 0 de legumes e 2 de palmito. Assim, calculando o número de permutações de um conjunto de 9 elementos onde há 3 elementos repetidos do tipo $|$ *e 6 do tipo* $*$ *estaremos calculando o número de maneiras diferentes de fazer a compra dos 6 salgados. Como* $P_9^{3,\ 6} = \frac{9!}{3! \times 6!} = 84$, *então há 84 maneiras diferentes de comprar os seis salgados numa lanchonete que há 4 opções de escolha.*

8.4 Arranjos - $A_n^p = A_{n,\ p}$

Considere o conjunto A $= \{a_1,\ a_2,\ a_3,\ \ldots,\ a_n\}$. Chama-se arranjo simples dos n elementos, tomados p a p, qualquer sequência (obedecendo a ordem) formada por p elementos distintos de A. Pelo princípio multiplicativo, o cálculo do número de arranjos é dado por $A_{n,p} = \{(n).(n-1).(n-2).\ldots.(n-p+1)\}$, com $n \geq p$, ou também tem-se:

$$A_n^p = A_{n,\ p} = \frac{n!}{(n-p)!}.$$

Exemplo 8.4.1. *Uma empresa colocará 3 produtos dos cinco* $\{A,\ B,\ C,\ D,\ E\}$ *em promoção. Quantas possíveis sequências distintas destes produtos podem ser formadas?*
1º modo - Pelo princípio multiplicativo tem-se: $5 \times 4 \times 3 = 60$.
2º modo - Utilizando a fórmula tem-se: $A_5^3 = \frac{5!}{(5-3)!} = 60$.

8.5 Combinações Simples - $C_n^p = C_{n,\ p}$

Considere o conjunto A $= \{a_1,\ a_2,\ a_3,\ \ldots,\ a_n\}$. Chama-se combinações simples dos n elementos, tomados p a p, qualquer subconjunto (não importando a ordem) de A com p elementos. O cálculo do número de combinações se faz obsevando que qualquer permutação de uma determinada sequência ordenada dá origem a uma única combinação. Assim:

$$C_n^p = C_{n,\ p} = \frac{A_n^p}{p!},$$

ou

$$C_n^p = \frac{n!}{p! \times (n-p)!}.$$

Exemplo 8.5.1. *De um grupo de 5 funcionários* $\{A,\ B,\ C,\ D,\ E\}$ *de uma empresa 3 serão escolhidos para fazerem parte da diretoria. De quantas maneiras diferentes pode ser* *escolhido este grupo?*

Como os funcionários ABC sejam os mesmos que CBA, CAB, ACB, BAC e BCA então tem-se:

$$\textit{1º modo: } \frac{(5\times4\times3)}{3!} = 10.$$
$$\textit{2º modo: } C_5^3 = \frac{5!}{3!\times(5-3)!} = 10.$$

Portanto os grupos (combinações) são:

ABC, ABD, ABC, ACD, ACE, BCD, BCE, BCE, BCE e CDE.

Note que para cada uma dessas combinações é possível formar 6 novos grupos (arranjos), daí, $10 \times 6 = 60$, *o que foi obtido no exemplo anterior.*