Universidade Federal de Goiás
Instituto de Matemática e Estatística
Programa de Mestrado Profissional em Matemática em Rede Nacional

Modelos Binomiais: caracterização e aplicações

Dorgival Fidellis de Souza

Goiânia

2014

sistema de bibliotecas ufg

UFG

TERMO DE CIÊNCIA E DE AUTORIZAÇÃO PARA DISPONIBILIZAR ELETRONICAMENTE OS TRABALHOS DE CONCLUSÃO DE CURSO NA BIBLIOTECA DIGITAL DA UFG

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Goiás (UFG) a disponibilizar, gratuitamente, por meio da Biblioteca Digital de Teses e Dissertações (BDTD/UFG), sem ressarcimento dos direitos autorais, de acordo com a Lei nº 9610/98, o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou *download*, a título de divulgação da produção científica brasileira, a partir desta data.

1. Identificação do material bibliográfico: **Trabalho de Conclusão de Curso de Mestrado Profissional**

2. Identificação do Trabalho

Autor (a):	Dorgival Fidellis de Souza					
E-mail:	dogyn@hotmail.com					
Seu e-mail pode ser disponibilizado na página? [x]Sim [ ] Não						
Vínculo empregatício do autor		Estatutário				
Agência de fomento:		Prefeitura de Senador Canedo			Sigla:	
País:	Brasil	UF:	GO	CNPJ:	25.107.525/0001-51	
Título:	Modelos Binomiais: caracterização e aplicações					
Palavras-chave:	Definição axiomática da probabilidade, Modelo Binomial, Matemática do Ensino Médio.					
Título em outra língua:	Binomial Models: characterization and applications					
Palavras-chave em outra língua:	probability axiomatic definition, binomial model, brazilian high school.					
Área de concentração:	Matemática do Ensino Básico					
Data defesa: (07/03/2014)						
Programa de Pós-Graduação:	ProfMat – Mestrado Profissional em Matemática em rede nacional					
Orientador (a):	Prof. Dr. Valdivino Vargas Júnior					
E-mail:	vvjunior@gmail.com					
Co-orientador(a):*						
E-mail:						

*Necessita do CPF quando não constar no SisPG

3. Informações de acesso ao documento:

Concorda com a liberação total do documento [x] SIM [ ] NÃO[1]

Havendo concordância com a disponibilização eletrônica, torna-se imprescindível o envio do(s) arquivo(s) em formato digital PDF ou DOC do trabalho de conclusão de curso.

O sistema da Biblioteca Digital de Teses e Dissertações garante aos autores, que os arquivos contendo eletronicamente as teses, dissertações ou trabalhos de conclusão de curso, antes de sua disponibilização, receberão procedimentos de segurança, criptografia (para não permitir cópia e extração de conteúdo, permitindo apenas impressão fraca) usando o padrão do Acrobat.

Dorgival Fidellis de Souza

Assinatura do (a) autor (a)

Data: 17 / 07 / 2014

[1] Neste caso o documento será embargado por até um ano a partir da data de defesa. A extensão deste prazo suscita justificativa junto à coordenação do curso. Os dados do documento não serão disponibilizados durante o período de embargo.

Dorgival Fidellis de Souza

Modelos Binomiais: caracterização e aplicações

Trabalho de Conclusão de Curso apresentado ao Instituto de Matemática e Estatística da Universidade Federal de Goiás, como parte dos requisitos para obtenção do grau de Mestre em Matemática.
Área de Concentração: Matemática do Ensino Básico
Orientador: Prof. Dr. Valdivino Vargas Júnior

Goiânia
2014

Dados Internacionais de Catalogação na Publicação na (CIP)
GPT/BC/UFG

S729m

Souza, Dorgival Fidellis de.
Modelos Binomiais [manuscrito] : caracterização e aplicações/ Dorgival Fidellis de Souza. - 2014.
xv, 86 f. : il., figs, tabs.

Orientador: Prof. Dr. Valdivino Vargas Júnior.
Dissertação (Mestrado) – Universidade Federal de Goiás, Instituto de Matemática e Estatística, 2014.
Bibliografia.
Inclui lista de figuras, abreviaturas, siglas e tabelas.
Apêndices.
1. Probabilidade (matemática) - Aplicacões 2. Matemática – Ensino fundamental 3. Matemática – Estudo e ensino I. Título.

CDU: 51:37

Dorgival Fidellis de Souza

Modelos Binomiais: Caracterização e Aplicações

Trabalho de Conclusão de Curso defendido no Programa de Mestrado Profissional em Matemática em Rede Nacional – PROFMAT/UFG, do Instituto de Matemática e Estatística da Universidade Federal de Goiás, como requisito parcial para obtenção do título de Mestre em Matemática, área de concentração Matemática do Ensino Básico, aprovado no dia 07 de março de 2014, pela Banca Examinadora constituída pelos professores:

Valdivino Vargas Júnior

Prof. Dr. Valdivino Vargas Júnior
Instituto de Matemática e Estatística-UFG
Presidente da Banca

Rui Seimetz

Prof. Dr. Rui Seimetz
Departamento de Matemática-UnB

Prof. Dr. Luis Rodrigo Fernandes Baumann
Instituto de Matemática e Estatística-UFG

Dorgival Fidellis de Souza graduou-se em Matemática na modalidade Licenciatura pela Universidade Federal de Goiás em 2011.

Este trabalho dedico a Jesus Cristo, o Mestre dos mestres que com sua didática nos leva à esperança, ao amor e à fé.

Agradecimentos

São tantas pessoas que passando pelo meu caminho me ajudaram, algumas até sem saber o que estavam fazendo.

Agradeço aos meus pais, José Dorgival de Souza e Lúcia Dias da Cunha.

Aos meus tios, Valdecy, Hilário, Sebastião e José.

À minha tia e também mãe Maria Cândida.

Aos amigos mestrandos da UFG bem como aos professores do IME. Especialmente, ao Professores Doutores Mário José de Souza, Rogério de Queiroz Chaves, Paulo Henrique de Azevedo Rodrigues, José Yunier Bello Cruz e às mestres Lucimeire Carvalho e Aline do IFG.

À meu orientador Professor Valdivino Vargas Júnior.

À fraternidade GOD FIRE.

Ao Ministério Luz para Os Povos.

À minha madrinha Izabel.

À minha avó Florinda Fernandes Dias.

À CAPES pelo suporte financeiro.

À Universidade Federal de Goiás.

À Republica Federativa do Brasil.

Ao único Deus verdadeiro, o Grande EU SOU.

Resumo

Neste trabalho, fizemos um ensaio teórico sobre os fundamentos da definição axiomática da probabilidade proposto por A. N. Kolmogorov e suas implicações ao modelo probabilístico de um experimento binomial, explorando conceitos e elementos próprios à Matemática discreta que se relacionam aos princípios aditivo e multiplicativo, combinações simples, triângulo de Pascal e binômio de Newton. Também, estudamos a probabilidade aplicada ao caso discreto passando pela definição de experimentos, eventos e variáveis aleatórias a fim de fundamentar nossa proposta de aplicação de um problema sobre transmissão de informações entre centrais em série ao Ensino Médio.

Palavras-chave: Definição axiomática da probabilidade, Modelo Binomial, Matemática do Ensino Médio.

Abstract

In this work, we made a theoric essay about the foundations of probability axiomatic definition, proposed by A. N. Kolmogorov and its implications in probabilistic models of the binomial experiment, exploring concepts and elements innate from discrete mathematics that is relate to the additive and multiplicative principles, simple combinations, Pascal ´s triangle and Newton´s binomy. Also, we study the probability applied to the discrete case throughout experiments definition, events and random variables, so that we could substantiate our problem application proposal, about information transmitting between central series to brazilian high school.

Keywords: probability axiomatic definition, binomial model, brazilian high school.

Sumário

1 Introdução

De acordo com Gnedenko (2008, p.11) a teoria das probabilidades tem origem no século XVII quando alguns matemáticos começaram a canalizar seus esforços para entender os eventos do acaso. Assim, através de trabalhos e correspondências, matemáticos tais como Pascal, Fermat, Huyghens e James Bernoulli, motivados por problemas ligados à jogos de azar e apostas abriram o caminho para formalização do conhecimento e definição daquilo que hoje conhecemos como *eventos aleatórios.* Contudo, apenas mais tarde, em meados do século XX, essa teoria começou a ganhar um corpo mais robusto e complexo motivada por problemas próprios ao estado democrático e a negócios de seguros que levaram alguns matemáticos, cujos trabalhos em probabi-lidade são considerados, nos dias de hoje, fundamentais a essa teoria, a buscarem uma crescente formalidade da matéria ligada à probabilidade.

Nesse processo a escola Russa de estudos probabilísticos ofereceu contribuições chaves para o caminho da formalização através de cientistas bem conhecidos como Tchebychev, Markov e Liapunov que introduziram o conceito de variável aleatória e se aproveitaram dele para apresentar soluções aos problemas próprios da época. Essas contribuições aliadas ao desenvolvimento dos métodos analíticos na teoria das probabilidades propostos por DeMoivre, Laplace, Gauss, Poisson e Lobachevsky possibilitaram à Emile Borel, Alexander Khintchin e Andrei Kolmogorov entre outros, em meados de 1920, aproximarem a teoria das probabilidades à análise funcional e teoria da medida, o que proveu soluções definitivas para problemas clássidos formulados por Tchebychev. A partir de então a definição axiomática de probabilidade conquistou, definitivamente, seu espaço e credibilidade na comunidade científica e acabou por mostrar que os métodos clássicos, criados pelos matemáticos do século XVII, baseados, exclusivamente, em aritmética elementar e processos combinatórios realmente eram abordados de forma correta. Mas, além disso, a definição axiomática possibilitou o estalecimento de meios teóricos consistentes e apropriados aos problemas modernos ligados à indústria de produção em massa, assuntos militares, física quântica, química molecular, dinâmica populacional entre outros assuntos.

Diante disso, temos por objetivo o estudo de *eventos aleatórios,* que são objeto da Probabilidade, propondo este ensaio teório para abordar os fundamentos dessa teoria a fim de encontrar pontos de contato com a matemática própria ao Ensino Médio do modo que está estruturado no Brasil [7]. Assim, falaremos, no primeiro capítulo, sobre os princípios de contagem, depois, no segundo capítulo, abordaremos a definição axiomática da probabilidade com ênfase no caso discreto para então, no terceiro capítulo, discorrermos sobre variáveis binomias. Ao final, traremos, no último capítulo, uma aplicação de modelos binomiais à problemas sobre transmissão de informações entre centrais em série bem como uma adaptação desta para o Ensino Médio.

2 Binômio de Newton

De acordo com Morgado (1991, p.2-3)[5] "O desenvolvimento do binômio $(1+x)^n$ está entre os primeiros problemas [...] ligados à Análise Combinatória. O caso $n = 2$ já pode ser encontrado nos *Elementos* de Euclides, em torno de 300 a.C."No entanto, foi apenas em 1713 que foi desenvolvido formalmente e publicado em *Ars Conjectondi* de Jaime Bernoulli a identidade objeto desta secção:

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

Os problemas que envolvem contagem tem na Análise Combinatória seu lugar de abordagem. O que almejamos fazer aqui é um ensaio teórico que nos conduza naturalmente ao conhecido *binômio de Newton.* Começaremos pelos princípios que fundamentam essa teoria conforme (Morgado, 1991)[5] os enuncia.

Princípio aditivo das partes disjuntas:

Definição 2.1. *Considere* n *conjuntos* $A_1, A_2, ..., A_n$, *dois a dois disjuntos, com* $a_1, a_2, ..., a_n$ *elementos respectivamente. Então,* $A_1 \cup A_2 \cup ... \cup A_n$ *possui* $a_1 + a_2 + ... + a_n$ *elementos. Isto é,*

$$\left| \bigcup_{i=1}^{n} A_i \right| = \sum_{i=1}^{n} a_i.$$

Princípio multiplicativo:

Definição 2.2. *Suponha que há* m_1 *maneiras de tomar uma decisão* d_1 *sendo que tomada essa decisão, temos* m_2 *caminhos para uma segunda decisão* d_2. *A seguir, tomadas as decisões* d_1 *e* d_2 *temos* m_3 *caminhos para uma terceira decisão* d_3, *todas independentes da primeira e segunda escolha. E assim sucessivamente até chegar a uma decisão* d_n *onde há* m_n *caminhos para escolher o modo a tomar a decisão, independente das* $n-1$ *primeiras escolhas. Então, ao tomar as* n *decisões juntas temos* $m_1 m_2 ... m_n$ *possibilidades.*

Agora vamos à seguinte questão: possuindo n objetos distintos desejamos retirar k objetos entre eles. Quantos modos existem de fazê-lo?

O primeiro passo para a resposta é escolher um primeiro objeto, depois um segundo, terceiro até o k-ésimo objeto que formará, ao final, um grupo de k objetos.

Para a primeira escolha dispomos de n objetos. Na escolha do segundo, podemos optar por $n-1$ objetos, pois um já foi escolhido. Para o terceiro são $n-2$ escolhas e, repetindo-se o processo até a k-ésima escolha vamos dispor de $n-(k-1)=n-k+1$ objetos. Usando o princípio multiplicativo podemos ordenar esses k objetos de $n\left(n-1\right)\left(n-2\right)...\left(n-k+1\right)$ modos. No entanto, dessa forma vários grupos que possuem os mesmos objetos foram contados repetidas vezes. Na verdade, um mesmo grupo foi contado $k!$ vezes. Observe, como referência, o grupo formado pelos elementos $\{a_{1,}a_{2,}a_{3,...,}a_k\}$. Quando escolhemos o grupo $\{a_{2,}a_{1,}a_{3,...,}a_k\}$ escolhemos um grupo com os mesmo objetos do anterior. Ou seja, um mesmo grupo foi contado novamente, pois neste caso a ordem dos elementos não é importante nessa contagem. Então para retirar as repetições na contagem devemos responder quantas vezes o grupo $\{a_{1,}a_{2,}a_{3,...,}a_k\}$ foi contado durante a contagem inicial cujo resultado foi $n\left(n-1\right)\left(n-2\right)...\left(n-k+1\right).$

A pergunta é simples de ser respondida, basta pensarmos nela da seguinte forma: De quantas maneiras podemos reordenar um grupo com k objetos?

Ora, escolhendo um objeto de cada vez para definir sua ordem, fazendo um processo similiar à contagem inicial e aplicando, novamente, o princípio multiplicativo obteremos $k\left(k-1\right)\left(k-2\right)...\left(k-\left(k-2\right)\right)\left(k-\left(k-1\right)\right)=k!$ Aplicando raciocínio análogo para os demais grupos percebemos que todos os grupos obtidos foram contados $k!$ vezes, o que nos permite concluir que os modos pelos quais podemos retirar k objetos entre n objetos distintos totalizam $\frac{n(n-1)(n-2)...(n-k+1)}{k!}$.

Definição 2.3. *Considere $n,k\in\mathbb{N}$ tais que $k\leq n$. Definimos, combinação simples $C_{n,k}$ como $C_{n,k}=\frac{n!}{k!(n-k)!}$.*

Observe que $C_{n,k}=\binom{n}{k}$. Então, a partir dessa definição usamos a notação $\binom{n}{k}$ para nos referir à quantidade de modos existentes de escolher k objetos numa coleção de n. Isto porque

$$\begin{aligned}\frac{n\left(n-1\right)\left(n-2\right)...\left(n-k+1\right)}{k!} &= \frac{n\left(n-1\right)\left(n-2\right)...\left(n-k+1\right)\left(n-k\right)!}{k!\left(n-k\right)!}\\ &= \frac{n!}{k!\left(n-k\right)!}\\ &= \binom{n}{k}.\end{aligned}$$

2.1 Triângulo de Pascal

O triângulo de Pascal reune todos os elementos do tipo $\binom{n}{k}$. Chamamos a seguinte estrutura de triângulo de Pascal:

$$
\begin{array}{llllll}
\binom{0}{0} & & & & & \\
\binom{1}{0} & \binom{1}{1} & & & & \\
\binom{2}{0} & \binom{2}{1} & \binom{2}{2} & & & \\
\binom{3}{0} & \binom{3}{1} & \binom{3}{2} & \binom{3}{3} & & \\
\binom{4}{0} & \binom{4}{1} & \binom{4}{2} & \binom{4}{3} & \binom{4}{4} & \\
. & \cdots & & & . & \\
. & & \cdots & & & . \\
. & & & \cdots & & . \\
\binom{n}{0} & \binom{n}{1} & \binom{n}{2} & \cdots & \binom{n}{k-1} & \binom{n}{k}
\end{array}
$$

Que numericamente é:

1						
1	1					
1	2	1				
1	3	3	1			
1	4	6	4	1		
1	5	10	10	5	1	
1	6	15	20	15	6	1
...			...			...

A primeira característica curiosa desse triângulo de números é a igualdade entre o elemento que está na k-ésima posição da direita para esquerda numa dada linha e o k-ésimo elemento da esquerda para direita da mesma linha. A seguir a afirmação é demonstrada.

Proposição 2.1. *Considere $n, k \in \mathbb{N}$ tais que $k \leq n$. Então,* $\binom{n}{k} = \binom{n}{n-k}$.

Demonstração: Da definição 2.3 temos,

$$\begin{aligned}
\binom{n}{k} &= \frac{n!}{k!\,(n-k)!} \\
&= \frac{n!}{(n-k)!k!} \\
&= \frac{n!}{(n-k)!\,(n-(n-k))!} \\
&= \binom{n}{n-k}.
\end{aligned}$$

□

A chamada relação de Stifel também auxilia em muito na construção de triângulos de Pascal com um número grande de linhas, em resolução de problemas e em demonstrações por indução que envolvam combinações simples. Pois, através deste resultado se estabeleceu uma relação dos elementos da n-ésima linha aos da linha anterior.

Proposição 2.2. (*Relação de Stifel*) *Considere* $n, k \in \mathbb{N}$ *tais que* $k \leq n$. *Então,* $\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}$.

Demonstração: Note que,

$$\begin{aligned}
\binom{n}{k} + \binom{n}{k+1} &= \frac{n!}{k!\,(n-k)!} + \frac{n!}{(k+1)!\,(n-k-1)!} \\
&= \frac{n!}{k!\,(n-k-1)!}\left(\frac{1}{(n-k)} + \frac{1}{(k+1)}\right) \\
&= \frac{n!}{k!\,(n-k-1)!}\left(\frac{k+1+n-k}{(n-k)\,(k+1)}\right) \\
&= \frac{n!}{k!\,(n-k-1)!}\left(\frac{1+n}{(n-k)\,(k+1)}\right) \\
&= \frac{(n+1)!}{(k+1)!\,(n-k)!} \\
&= \frac{(n+1)!}{(k+1)!\,(n+1-(k+1))!} \\
\binom{n}{k} + \binom{n}{k+1} &= \binom{n+1}{k+1}.
\end{aligned}$$

□

O próximo resultado nos oferece uma forma muito econômica para o cálculo da soma entre elementos de uma mesma linha. Veja:

Proposição 2.3. *Sendo $n, k \in \mathbb{N}$ tais que $k \leq n$. Então,* $\displaystyle\sum_{i=0}^{n} \binom{n}{i} = 2^n$.

Demonstração: Seja L_n a soma de todos os termos da n-ésima linha do triângulo de Pascal. Então,

$$\begin{aligned}
L_n &= \sum_{i=0}^{n} \binom{n}{i} \\
&= \binom{n}{0} + \binom{n}{n} + \sum_{i=1}^{n-1} \left[\binom{n-1}{i-1} + \binom{n-1}{i} \right]; \text{ (Relação de Stifel)} \\
&= \binom{n-1}{0} + \binom{n-1}{n-1} + \sum_{i=1}^{n-1} \binom{n-1}{i-1} + \sum_{i=1}^{n-1} \binom{n-1}{i} \\
&= \left[\binom{n-1}{n-1} + \sum_{i=1}^{n-1} \binom{n-1}{i-1} \right] + \left[\binom{n-1}{0} + \sum_{i=1}^{n-1} \binom{n-1}{i} \right]
\end{aligned}$$

$$\begin{aligned}
&= \sum_{i=1}^{n} \binom{n-1}{i-1} + \sum_{i=0}^{n-1} \binom{n-1}{i} \\
&= 2 \sum_{i=0}^{n-1} \binom{n-1}{i}; \text{ (Ajustando os índices)} \\
L_n &= 2L_{n-1}.
\end{aligned}$$

Isto nos diz que a soma das linhas do Triângulos de Pascal formam uma sequência de razão constante $q = \frac{L_n}{L_{n-1}} = 2$. Sendo assim, como o primeiro termo dessa progressão geométrica é $L_0 = 1$ concluímos que $L_n = L_0 q^n = 2^n$.

□

A seguir, vamos demonstrar um resultado relativo ao Triângulo de Pascal que também é conhecido como Teorema das Colunas, pois através dele podemos, sempre que precisar, somar os elementos de uma determinada coluna até certa linha para obter o elemento da próxima linha que está na coluna adjacente à direita da coluna inicial.

Proposição 2.4. (*Teorema das Colunas*)*Seja c correspondente à c-ésima coluna do triângulo de Pascal e n o número de elementos a serem somados. Então* $\displaystyle\sum_{i=1}^{n} \binom{c-2+i}{c-1} = \binom{c-1+n}{c}$.

Demonstração: Usaremos o método de indução para demonstrar. Para $n = 1$ temos $\displaystyle\sum_{i=1}^{1}\binom{c-2+i}{c-1} = \binom{c-1}{c-1} = 1 = \binom{c-1+1}{c}$; constata-se a igualdade.

Suponha que seja válida para algum $n-1 \geq 2$. Isto é,

$$\sum_{i=1}^{n-1}\binom{c-2+i}{c-1} = \binom{c-2+n}{c}.$$

Então,

$$\begin{aligned}
\sum_{i=1}^{n}\binom{c-2+i}{c-1} &= \sum_{i=1}^{n-1}\binom{c-2+i}{c-1} + \binom{c-2+n}{c-1}; \\
&= \binom{c-2+n}{c} + \binom{c-2+n}{c-1}; \text{ (Hipótese de indução)} \\
\sum_{i=1}^{n}\binom{c-2+i}{c-1} &= \binom{c-1+n}{c}; \text{ (Relação de Stifel)}
\end{aligned}$$

□

Proposição 2.5. *(Teorema das Diagonais) Seja* $D_n = \displaystyle\sum_{i=0}^{n}\binom{d+i}{i}$ *com* $d, n \in \mathbb{N}$. *Então* $D_n = \binom{d+n+1}{n}$.

Demonstração: Utilizando o método indutivo, temos para $n = 0$ que $D_0 = \displaystyle\sum_{i=0}^{0}\binom{d+i}{i} = \binom{d}{0} = 1 = \binom{d+1}{0}$.

Suponha que haja algum $n \in \mathbb{N}$ tal que,

$$\sum_{i=0}^{n}\binom{d+i}{i} = \binom{d+n+1}{n}$$

Assim,

$$\begin{aligned}
D_{n+1} &= \sum_{i=0}^{n+1}\binom{d+i}{i} \\
&= D_n + \binom{d+n+1}{n+1} \\
&= \binom{d+n+1}{n} + \binom{d+n+1}{n+1}; \text{ (Hipótese de indução)} \\
D_{n+1} &= \binom{d+(n+1)+1}{n+1}; \text{ (Relação de Stifel)}
\end{aligned}$$

$\square$

O resultado central desta seção será anunciado a seguir e será evocado sempre que necessário nas provas posteriores.

2.2 Teorema Binomial

Proposição 2.6. *(*Teorema Binomial*) Se $x, y \in \mathbb{R}$ e $n \in \mathbb{N}$, então $(y+x)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$.*

Demonstração: Usaremos indução. Note que,

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

De fato, a identidade é válida para $n = 0$.

Suponha,

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

Para $n+1$ temos,

$$\begin{aligned}
\sum_{k=0}^{n+1} \binom{n+1}{k} x^k y^{n+1-k} &= y^{n+1} + \sum_{k=1}^{n} \binom{n+1}{k} x^k y^{n+1-k} + x^{n+1}; \text{ (Relação de Stifel)} \\
&= y^{n+1} + \sum_{k=1}^{n} \left[\binom{n}{k-1} + \binom{n}{k} \right] x^k y^{n+1-k} + x^{n+1} \\
&= y^{n+1} + \sum_{k=1}^{n} \binom{n}{k-1} x^k y^{n+1-k} + \sum_{k=1}^{n} \binom{n}{k} x^k y^{n+1-k} + x^{n+1}; \; j = k-1 \\
&= y^{n+1} + \sum_{j=0}^{n-1} \binom{n}{j} x^{j+1} y^{n-j} + \sum_{k=1}^{n} \binom{n}{k} x^k y^{n+1-k} + x^{n+1} \\
&= y^{n+1} + x \sum_{j=0}^{n-1} \binom{n}{j} x^j y^{n-j} + y \sum_{k=1}^{n} \binom{n}{k} x^k y^{n-k} + x^{n+1} \\
&= y \left[y^n + \sum_{k=1}^{n} \binom{n}{k} x^k y^{n-k} \right] + x \left[\sum_{j=0}^{n-1} \binom{n}{j} x^j y^{n-j} + x^n \right] \qquad (1)
\end{aligned}$$

$$
\begin{aligned}
\sum_{k=0}^{n+1} \binom{n+1}{k} x^k y^{n+1-k} &= y \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} + x \sum_{j=0}^{n} \binom{n}{j} x^j y^{n-j}; \text{ (Hipótese de indução)} \\
&= y\,(y+x)^n + x\,(y+x)^n \\
\sum_{k=0}^{n+1} \binom{n+1}{k} x^k y^{n+1-k} &= (y+x)^{n+1}\,.
\end{aligned}
$$

$\square$

Este resultado, útil aos mais diversos ramos da Matemática, será utilizado neste trabalho nos estudos relativos à teoria da probabilidade e suas aplicações.

3 Probabilidade: o caso discreto

3.1 Experimentos Aleatórios

A Matemática sempre foi a ciência da certeza e exatidão, daquilo que é sempre válido para toda uma classe de elementos, em teoremas e proposições que evocando um certo conjunto de hipóteses garantem uma tese proposta pelo enunciado e isso, tem influenciado a filosofia das demais ciências. No entanto, há longo tempo a sociedade já se deparava com algo um tanto quanto curioso e inusitado para a ciência das certezas. Os jogos de azar com cartas e dados foram um dos primeiros fenômenos permeados por incertezas registrados por estudiosos da Matemática. E esta é a diferença fundamental entre a Probabilidade e as demais áreas da Matemática: a incerteza. Muitos dos fenômenos da natureza e da sociedade têm como característica principal a incerteza e a Ciência, sobretudo a Matemática tem tido crescente interesse no estudo e modelagem desses fenômenos. Diante disso, de acordo com Gnedenko (2008, p.20)[6], os eventos[1] podem ser classificados, quanto à sua ocorrência, como:

i) eventos certos, aqueles cuja ocorrência é inevitável sempre que está estabelecido um mesmo conjunto de condições φ.

ii) eventos impossíveis, eventos que nunca ocorrem quando se estabelece o conjunto de condições φ.

iii) eventos aleatórios, que podem ou não ocorrer mediante a realização de todas as condições que compõem φ, isto é, eventos cuja ocorrência é incerta.

Este ensaio teórico, tem por objetivo o estudo destes últimos, que são objeto da Probabilidade. Assim, podemos dizer que um experimento, realizado sob as mesmas condições, é aleatório quando seus resuldados determinam um conjunto com pelo menos dois elementos. Formalmente,

Definição 3.1. Experimentos Aleatórios *são experimentos que ao serem repetidos, obedecendo um mesmo conjunto de condições* φ, *podem não produzir os mesmos resultados.*

As comunidades científicas de várias epócas e campos do conhecimento cujos trabalhos se concentravam em experimentos aleatórios percorriam diversos caminhos para a definição da probabilidade. De modo que, por muito tempo não houve consenso sobre

[1] acontecimentos observados na realização de um experimento científico desenvolvimento de acordo com um conjunto de condições pré-definidas

a fundamentação da teoria das probabilidades. Gnedenko (2008, p.24)[6] divide, de forma geral, as diversas definições de probabilidade defendidas nos seguintes grupos:

A) Definições de probabilidade concebidas como uma medida quantitativa do "grau de certeza"do observador;

B) Definições que reduzem a conceito de probabilidade à noção mais primitiva de "verossimilhança"(a chamada definição clássica);

C) Definições que tomam como ponto de partida a "frequência relativa"de ocorrência do evento num grande número de testes (definição "estatística")

Este autor critica severamente as definições do grupo A, pois eventos aleatórios, probabilisticamente falando, não tem sua chance de ocorrência baseada no olhar subjetivo de quem observa. Se assim fosse, o mundo externo e as leis naturais que regem fenômenos aleatórios (como as relacionadas ao clima) não seriam, de fato, leis, porquanto estariam condicionadas a um observador. No entanto, o que se percebe nesses fenômenos é que eles ocorrem, independentemente, da interpretação subjetiva mensurada, quatitativamente, como "grau de certeza"do observador. Além disso,

> "se assumimos que a avaliação da probabilidade depende só do estado do observador, então todas as concluções baseadas em julgamentos probabilísticos [...] são deprovidas do significado objetivo que elas têm, independemente do observador"(Gnedenko, 2008, p.25) [6]

Este primeiro grupo, como podemos perceber, não expressam em seus conceitos a objetividade inerente à Matemática em todos os seus sólidos ramos de desenvolvimento teórico-axiomáticos. Ao contrário, se concentra em estados psíquicos de dúvida situados entre a aceitação e a negação absoluta emitida por alguém que se utiliza de uma gama de aspectos subjetivos para emitir tal opinião. Deste modo, Gnedenko (2008, p.27) [6] conclui dizendo, "por mais interessante que possa parecer de um ponto de vista psicológico, [isso] nos levaria muito além do campo de nosso problema básico de explicar o significado das leis probabilísticas que têm valor científico objetivo".

As definições pertencentes ao grupo B residem em fundamentos e concepções anteriores à construção axiomática da teoria da probabilidade e têm poder de alcance muito limitado na modelagem de fenômenos aleatórios, uma vez que partem do pressuposto de que todos os eventos de um conjunto de eventos aleatórios inerentes a um dado experimento possuem probabilidades iguais de ocorrência, o que não se constata em vários experimentos aleatórios. Assim, apesar de se desviar da subjetividade do grupo A, sendo até certo ponto frutífera, a definição clássica de probabilidade não consegue suprir todas as demandas modernas de modelagem de fenômenos aleatórios. Para isso, então vários matemáticos contribuíram para a criação do modelo axiomático de probalidade.

O problema da construção axiomática da probabilidade foi inicialmente colocada e resolvida em 1917, por S. N. Bernstein (renomado matemático), seu trabalho se baseou na comparação qualitativa de eventos aleatórios de acordo com sua maior ou menor probabilidade. Contudo, existe outro enfoque proposto por A. N. Kolmogorov. Este relaciona intimamente a teoria da probabilidade com a teoria dos conjuntos e utiliza os aspectos teóricos modernos da medida na teoria das funções de uma variável real. (Gnedenko, 2008, p.67-68)[6]

Nosso trabalho seguirá o caminho proposto por Kolmogorov.

3.2 Modelo Probabilístico

De maneira geral, o tratamento axiomático das funções da teoria da probabilidade advêm de um arcabouço de conhecimentos desenvolvidos a partir de propriedades fundamentais constatadas pela experimentação e da filosofia matemática intrínseca aos grupos B e C. Assim, Gnedenko (2008, p.67)[6] explica que "a definição axiomática de probabilidade inclui como um caso especial ambas as definições: clássica [grupo B] e estatística [grupo C] e ultrapassa as limitições de cada uma delas."Além disso, os modelos probabilísticos, que definiremos logo a diante, permitem o estudo de experimentos aleatórios que se vinculam ao caso contínuo e discreto. No entanto, ainda que o modelo probabilístico se aplique a esses dois casos, a partir de certo ponto, o leitor perceberá que abandonamos o primeiro para nos focar no segundo: o caso discreto. Vamos às definições.

Definição 3.2. *Seja Ω o conjunto dos possíveis resultados de um experimento aleatório. Uma classe de eventos aleatórios $\mathcal{F}$ é um conjunto tal que*

i) $\Omega \in \mathcal{F}$.

ii) Seja $A \subset \Omega$. Se $A \in \mathcal{F}$, então $A^c \in \mathcal{F}$.

iii) Se $A_1, A_2, ..., A_n, .. \in \mathcal{F}$, então $\displaystyle\bigcup_{k=1}^{\infty} A_k \in \mathcal{F}$.

Definição 3.3. *Seja o espaço amostral Ω de um experimento aleatório e $\mathcal{F}$ uma classe de eventos aleatórios. Se $A \in \mathcal{F}$ então, A é um* evento aleatório.

Desta definição, naturalmente, concluímos que o próprio Ω e o conjunto vazio $\varnothing$ são eventos vínculados a um experimento aleatório e, ainda, podemos dizer que Ω é o evento certo ao passo que $\varnothing$ é o evento impossível. A partir daqui chamaremos o conjunto Ω de espaço amostral.

Definição 3.4. *A função $P : \mathcal{F} \longrightarrow [0,1]$ é chamada* função de probabilidade *e é regida pelos seguintes axiomas:*

AXIOMA 1: Se $A \in \mathcal{F}$, então $0 \leq P(A) \leq 1$.
AXIOMA 2: $P(\Omega) = 1$.
AXIOMA 3: Se $A_1, A_2, ..., A_n, .. \in \mathcal{F}$ e $A_i \cap A_j = \varnothing$ sempre que $i \neq j$, então

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

Definição 3.5. *Dado Ω definimos* modelo probabilístico *como um espaço de três coordenadas $(\Omega, \mathcal{F}, P)$ onde $\mathcal{F}$ é uma classe de eventos aleatórios de Ω e P é uma função de probabilidade.*

A seguir, daremos um exemplo de uma clase de eventos $\mathcal{F}$ muito adequado ao caso discreto que é o conjunto das partes de Ω.

Proposição 3.1. *Se $\mid \Omega \mid = n$ finito e $\mathcal{F}$ o conjunto das partes de Ω, então $\mid \mathcal{F} \mid = 2^n$.*

Demonstração: Primeiramente vamos mostrar que o conjunto das partes de Ω é uma classe de eventos. Seja A um conjunto formado por elementos de Ω. Então, $\Omega - A = A^c \subset \Omega$ o que significa que A^c é formado por elementos de Ω e, portanto, compõe o conjunto da partes. Considere os conjuntos $A_1, A_2, ..., A_n, .. \subset \Omega$, então $A_i \in \mathcal{F}$ conjunto das partes $\forall \ i \in \mathbb{N}$ e, além disso, $\bigcup_{k=1}^{\infty} A_k \subset \Omega$, por isso pertence a $\mathcal{F}$. Logo, $\mathcal{F}$ é uma classes de eventos de Ω. Agora basta contar os subconjuntos de Ω. Notamos que existe $\binom{n}{0} = 1$ subconjunto de Ω vazio, $\binom{n}{1} = n$ subconjuntos de Ω unitários, $\binom{n}{2}$ subconjuntos de Ω com dois elementos e assim por diante. Genericamente, os subconjuntos de Ω com k elementos totalizam $\binom{n}{k}$. Como $\mathcal{F}$ é o conjunto das partes de Ω segue que,

$$
\begin{aligned}
| \ \mathcal{F} \ | &= \binom{n}{0} + \binom{n}{1} + ... + \binom{n}{n} \\
| \ \mathcal{F} \ | &= \sum_{k=0}^{n} \binom{n}{k}; \text{ (Teorema Binomial, p.21)} \\
| \ \mathcal{F} \ | &= (1+1)^n \\
| \ \mathcal{F} \ | &= 2^n.
\end{aligned}
$$

□

A Definição 3.3 é fundamental para o estudo da probabilidade, proposto por Kolmogorov, pois ela faz a ponte que nos permite utilizar várias implicações da Teoria de Conjuntos em nossos estudos. Assim, nosso próximo passo será trazer alguns conhecimentos dessa Teoria e aplicá-los à eventos aleatórios.

Lema 3.1. *Considere os eventos aleatórios* $A, B, C \in \mathcal{F}$. *Então,*

a) $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$.
b) $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$.
c) $(A \cap B)^c = A^c \cup B^c$.
d) $(A \cup B)^c = A^c \cap B^c$.

Demonstração:

a) De fato, pois $w \in (A \cap C) \cup (B \cap C) \iff w \in (A \cap C)$ ou $w \in (B \cap C) \iff w \in C$ e $w \in A \cup B \iff w \in (A \cup B) \cap C$.

□

b) Agora suponha $w \in (A \cup C) \cap (B \cup C) \iff w \in (A \cup C)$ e $w \in (B \cup C) \iff w \in C$ ou $w \in A \cap B \iff w \in (A \cap B) \cup C$.

□

c) De fato, pois $w \in A^c \cup B^c \iff w \notin A$ ou $w \notin B \iff w \notin A \cap B \iff w \in (A \cap B)^c$.

□

d) De fato, pois $w \in A^c \cap B^c \iff w \notin A$ e $w \notin B \iff w \notin A \cup B \iff w \in (A \cup B)^c$.

□

Proposição 3.2. *Sejam* $A, B \in \mathcal{F}$, *então* $A \cup B \in \mathcal{F}$ *e* $A \cap B \in \mathcal{F}$.

Demonstração: Tome $A = A_1$, $B = A_2$ e $A_n = \varnothing$ para $n \geq 3$. Então, $A_1, A_2, ..., A_n \in \mathcal{F}$. Logo, $\bigcup_{k=1}^{\infty} A_k \in \mathcal{F}$. Mas $\bigcup_{k=1}^{\infty} A_k = A \cup B$. Portanto, $A \cup B \in \mathcal{F}$.

Note também que $A \cap B = (A^c \cup B^c)^c$. Como $A \in \mathcal{F}$ e $B \in \mathcal{F}$ segue que $A^c \in \mathcal{F}$ e $B^c \in \mathcal{F}$. Assim, $A^c \cup B^c \in \mathcal{F}$ e com isso, pela definição 3.2, se completa o resultado $(A^c \cup B^c)^c = A \cap B \in \mathcal{F}$.

$\square$

Os itens c) e d) do lema 3.1 são bastante úteis e vamos generalizá-los para um número qualquer de eventos.

Lema 3.2. *(Leis de DeMorgan) Sejam os eventos $A_1, A_2, ..., A_n, ... \in \mathcal{F}$. Então, são válidas:*

$$\begin{aligned} \text{i)} \left(\bigcup_{k=1}^{\infty} A_k\right)^c &= \bigcap_{k=1}^{\infty} (A_k)^c. \\ \text{ii)} \left(\bigcap_{k=1}^{\infty} A_k\right)^c &= \bigcup_{k=1}^{\infty} (A_k)^c. \end{aligned} \tag{2}$$

Demontração:

Comecemos por i). Para w arbitário temos,

$$\begin{aligned} w \quad \in \quad \bigcap_{k=1}^{\infty} (A_k)^c &\Longleftrightarrow w \notin A_k \qquad \forall \qquad k \in \mathbb{N} \\ \Longleftrightarrow \quad w \notin \bigcup_{k=1}^{\infty} A_k &\Longleftrightarrow w \in \left(\bigcup_{k=1}^{\infty} A_k\right)^c. \end{aligned}$$

Para a prova de ii) considere $w \in [(A_k)^c]^c \Longleftrightarrow w \notin (A_k)^c \Longleftrightarrow w \in A_k$, assim, $[(A_k)^c]^c = A_k$. Agora, por i) temos

$$\begin{aligned} \left[\bigcup_{k=1}^{\infty} (A_k)^c\right]^c &= \bigcap_{k=1}^{\infty} [(A_k)^c]^c \\ \left\{\left[\bigcup_{k=1}^{\infty} (A_k)^c\right]^c\right\}^c &= \left(\bigcap_{k=1}^{\infty} A_k\right)^c \\ \bigcup_{k=1}^{\infty} (A_k)^c &= \left(\bigcap_{k=1}^{\infty} A_k\right)^c. \end{aligned}$$

$\square$

Todos esses resultados, sobre eventos aleatórios, são de suma importância para as operações entre eventos aleatórios, o cálculo de probabilidades e fundamentar propriedades próprias a esses elementos. Há de se mencionar que as propriedades são as relações mais fundamentais em uma teoria, uma vez que a partir delas se estabelece as "regras do jogo"num campo teórico. Além disso, é a partir das propriedades que podemos concluir os resultados mais avançados e abstratos. Então, vejamos as propriedades de uma probabilidade P.

3.3 Propriedades da probabilidade

Proposição 3.3. *Seja a tripla $(\Omega, \mathcal{F}, P)$ vinculada a um experimento aleatório qualquer. Tome $A, B \in \mathcal{F}$ e $\{A_k\}_{k\in\mathbb{N}}$ em $\mathcal{F}$. Então são válidas as propriedades:*

Propriedade 1: $P(\varnothing) = 0$.

Propriedade 2: $P\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} P(A_k)$ sempre que $A_i \cap A_j = \varnothing$ com $i \neq j$.

Propriedade 3: $P(A^c) = 1 - P(A)$.

Propriedade 4: Se $A \subset B$, então $P(A) \leq P(B)$.

Propriedade 5: $P\left(\bigcup_{k=1}^{n} A_k\right) \leq \sum_{k=1}^{n} P(A_k)$.

Propriedade 6: $P\left(\bigcup_{k=1}^{n} A_k\right) = 1 - P\left[\bigcap_{k=1}^{n} (A_k)^c\right]$.

Propriedade 7: $P\left(\bigcap_{k=1}^{n} A_k\right) \geq 1 - \sum_{k=1}^{n} P[(A_k)^c]$.

Propriedade 8: Se $A \subset B$, então $P(B \cap A^c) = P(B) - P(A)$.

Essas propriedades são implicações imediatas da definição axiomática da probabilidade e se aplicam a qualquer tripla $(\Omega, \mathcal{F}, P)$ tanto a casos contínuos como discretos. Vamos às **demonstrações**:

Propriedade 1: Devemos recorrer ao Axioma 3 e considerar a igualdade entre as probabilidades

$$\begin{aligned} P(\Omega) &= P(\Omega \cup \varnothing \cup \varnothing \cup ...) \\ &= P(\Omega) + \sum_{k=1}^{\infty} P(\varnothing);\ \text{(Axioma 3).} \end{aligned}$$

Disto, $\sum_{k=1}^{\infty} P(\varnothing) = P(\Omega) - P(\Omega) = 0$

Mas estamos falando de uma soma de termos não-negativos, pois $0 \leq P(\varnothing) \leq 1$ (Axioma 1). Portanto, cada parcela deve ser zero, isto é, $P(\varnothing) = 0$.

$\square$

Propriedade 2: Note que,

$$\bigcup_{k=1}^{n} A_k = \left(\bigcup_{k=1}^{n} A_k\right) \cup \left(\bigcup_{k=n+1}^{\infty} \varnothing\right)$$

$$\Longrightarrow P\left(\bigcup_{k=1}^{n} A_k\right) = P\left[\left(\bigcup_{k=1}^{n} A_k\right) \cup \left(\bigcup_{k=n+1}^{\infty} \varnothing\right)\right].$$

Como todos os conjuntos são disjuntos, pelo Axioma 3 temos,

$$\begin{aligned} P\left(\bigcup_{k=1}^{n} A_k\right) &= \sum_{k=1}^{n} P(A_k) + \sum_{k=n+1}^{\infty} P(\varnothing) \\ P\left(\bigcup_{k=1}^{n} A_k\right) &= \sum_{k=1}^{n} P(A_k). \end{aligned}$$

$\square$

Propriedade 3: É claro que

$$\begin{aligned} \Omega &= A \cup A^c \\ P(\Omega) &= P(A \cup A^c);\ \text{(propriedade 2)} \\ P(\Omega) &= P(A) + P(A^c); \end{aligned}$$

Para finalizar a prova basta aplica o Axioma 2, $P(A) = 1 - P(A^c)$.

$\square$

Propriedade 4: Como $A \subset B$ vale a seguinte identidade.

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

Como a união é disjunta segue que,

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

Por outro lado, $0 \leq P\left(A^c \cap B\right) \leq 1$. Ao somar $P\left(A\right)$ concluímos a demonstração $P\left(A\right) \leq P\left(A\right) + P\left(A^c \cap B\right) = P\left(B\right)$.

$\square$

Propriedade 5: Considere a identidade

$$A_1 \cup A_2 = A_1 \cup \left[A_2 \cap A_1^c\right].$$

Como $A_1 \cup \left[A_2 \cap A_1^c\right]$ é uma união disjunta ao aplicar a função de probabilidade temos,

$$P\left(A_1 \cup A_2\right) = P\left(A_1\right) + P\left(A_2 \cap A_1^c\right).$$

Mas, $A_2 \cap A_1^c \subset A_2$, então $P\left(A_2 \cap A_1^c\right) \leq P\left(A_2\right)$. O que implica em $P\left(A_1 \cup A_2\right) \leq P\left(A_1\right) + P\left(A_2\right)$. A tese constata-se para $n = 2$. Suponha, por indução, que $P\left(\bigcup_{k=1}^{n} A_k\right) \leq \sum_{k=1}^{n} P\left(A_k\right)$. Então,

$$\begin{aligned}
P\left(\bigcup_{k=1}^{n+1} A_k\right) &= P\left(\bigcup_{k=1}^{n} A_k \cup A_{n+1}\right) \\
&\leq P\left(\bigcup_{k=1}^{n} A_k\right) + P\left(A_{n+1}\right); \text{ (Caso } n = 2\text{)} \\
&\leq \sum_{k=1}^{n} P\left(A_k\right) + P\left(A_{n+1}\right); \text{ (Hipótese de indução)} \\
&= \sum_{k=1}^{n+1} P\left(A_k\right).
\end{aligned}$$

$\square$

Propriedades 6:

Das leis de DeMorgan (p. 28) temos, $\bigcup_{k=1}^{n} A_k = \left[\bigcap_{k=1}^{n} \left(A_k\right)^c\right]^c$

$$\begin{aligned}
P\left(\bigcup_{k=1}^{n} A_k\right) &= P\left(\left[\bigcap_{k=1}^{n} \left(A_k\right)^c\right]^c\right); \text{ (Propriedade 3)} \\
&= 1 - P\left[\bigcap_{k=1}^{n} \left(A_k\right)^c\right].
\end{aligned}$$

$\square$

Propriedade 7:

Aplicando DeMorgan segue a identidade, $\bigcap_{k=1}^{n} A_k = \left[\bigcup_{k=1}^{n} (A_k)^c\right]^c$

$$\begin{aligned}
P\left(\bigcap_{k=1}^{n} A_k\right) &= P\left(\left[\bigcup_{k=1}^{n} (A_k)^c\right]^c\right); \text{ (Propriedade 3)} \\
&= 1 - P\left[\bigcup_{k=1}^{n} (A_k)^c\right]; \text{ (Propriedade 5)} \\
P\left(\bigcap_{k=1}^{n} A_k\right) &\geq 1 - \sum_{k=1}^{n} P\left[(A_k)^c\right].
\end{aligned}$$

$\square$

Propriedade 8: Fazendo $B = A \cup (A^c \cap B) \Longrightarrow P(B) = P(A) + P(A^c \cap B) \Longrightarrow P(A^c \cap B) = P(B) - P(A)$.

$\square$

Apesar de haver outras propriedades importantes, decorrentes da definição axiomática da probabilidade, as omitiremos aqui por não haver necessidade de utilização de tais propriedades no decorrer do trabalho. Por isso, iremos a diante nos focando na caracterização de variáveis aleatórias discretas. Para o leitor interessado em mais propriedades consultar [2], [3] e [6] na seção de referências.

3.4 Probabilidade Condicional e Independência

Outro conceito muito importante no estudo de eventos aleatórios é a probabilidade condicional. Podemos entendê-la como uma redução do espaço amostral e, ainda, usá-la para calcular probabilidades através da partição de um espaço amostral quando conveniente.

Definição 3.6. *Seja $(\Omega, \mathcal{F}, P)$ um espaço de probabilidade. Se $A, B \in \mathcal{F}$ e $P(B) > 0$, a probabilidade condicional de A dado B, representada por $P(A \mid B)$, é*

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

Passaremos agora ao conceito probabilístico de independência cuja utilização na teoria da probabilidade é fundamental para operações com probabilidades, entre eventos e variáveis aleatórias.

Definição 3.7. *Os eventos aleatórios* $A_1, A_2, ..., A_n$ *são* independentes *se a ocorrência de qualquer subconjunto de* $\{A_1, A_2, ..., A_{k-2}, A_{k-1}, A_{k+1}, A_{k+2}, ..., A_n\}$ *não altera a probabilidade* $P(A_k)$. *Isto é,*

$$P(A_k) = P(A_k \mid S_{n,k})$$

onde $S_{n,k} \subset \{A_1, A_2, ..., A_{k-2}, A_{k-1}, A_{k+1}, A_{k+2}, ..., A_n\}$

Veja o exemplo abaixo.

Exemplo 3.1. *Se* A_1, A_2 *e* A_3 *são eventos aleatórios independentes, então são válidas as igualdades:*

a) $P(A_1 \mid A_2) = P(A_1)$
b) $P(A_1 \mid A_3) = P(A_1)$
c) $P(A_3 \mid A_1) = P(A_3)$
d) $P(A_3 \mid A_2) = P(A_3)$
e) $P(A_2 \mid A_1) = P(A_2)$
f) $P(A_2 \mid A_3) = P(A_2)$
g) $P(A_1 \mid A_2 \cap A_3) = P(A_1)$
h) $P(A_2 \mid A_1 \cap A_3) = P(A_2)$
i) $P(A_3 \mid A_2 \cap A_1) = P(A_3)$

O próximo resultado é conhecido como Teorema da Multiplicação.

Proposição 3.4. *Num espaço de probabilidade* $(\Omega, \mathcal{F}, P)$ *em que* $A_i \in \mathcal{F} \quad \forall \quad i = 1, 2, ..., n$ *temos:*

$$P\left(\bigcap_{k=1}^{n} Ai\right) = P(A_1)\, P(A_2 \mid A_1)\, P(A_3 \mid A_1 \cap A_2) \ldots P(A_n \mid A_1 \cap A_2 \cap \ldots \cap A_{n-1}).$$

Demonstração: Faremos por indução. Da Definição 3.6 segue que $P(A_1 \cap A_2) = P(A_1)\, P(A_1 \mid A_2)$, isto é, para $n = 2$ constata-se. Suponha, então, a identidade vale para algum $n \in \mathbb{N}$.

$$\begin{aligned} P\left(\bigcap_{k=1}^{n+1} Ai\right) &= P\left[\left(\bigcap_{k=1}^{n} Ai\right) \cap A_{n+1}\right]; \ (n = 2) \\ &= P\left(\bigcap_{k=1}^{n} Ai\right) P\left(A_{n+1} \mid \bigcap_{k=1}^{n} Ai\right); \ (\text{hipótese de indução}) \\ &= P(A_1)\, P(A_2 \mid A_1)\, P(A_3 \mid A_1 \cap A_2) \ldots P(A_{n+1} \mid A_1 \cap A_2 \cap \ldots \cap A_n). \end{aligned}$$

□

Corolário 3.1. *Se os eventos aleatórios* $A_1, A_2, ..., A_n$ *são* independentes*, então*

$$P\left(\bigcap_{k=1}^{m} A_{i_k}\right) = \prod_{k=1}^{m} P\left(A_{i_k}\right)$$

com $1 < j \leq n$ *e* $\forall \quad \{i_1, i_2, ..., i_k\} \subset \{1, 2, ..., n\}$ *tal que* $1 \leq i_1 < i_2 < ... < i_k \leq n$ *e* $k = 1, 2, ..., n.$

Demonstração: Segue imediatamente da definição 3.7 e da proposição 3.4, pois quando $A_1, A_2, ..., A_n$ são *independentes* a probabilidade $P\left(\bigcap_{k=1}^{n} Ai\right)$ é, exatamente, o produto das probabilidades individuais $P\left(A_i\right)$. Portanto, através deste corolário podemos, quando constatada a independência entre eventos, calcular a probabilidade de ocorrência simultânea de dois, três, quatro até n eventos aleatórios pela multiplicação das probabilidades de cada evento.

Exemplo 3.2. *Sendo* A_1, A_2 *e* A_3 *eventos independentes as probabilidade das ocorrências conjuntas são*

a) $P\left(A_1 \cap A_2\right) = P\left(A_1\right) P\left(A_2\right)$
b) $P\left(A_1 \cap A_3\right) = P\left(A_1\right) P\left(A_3\right)$
c) $P\left(A_2 \cap A_3\right) = P\left(A_2\right) P\left(A_3\right)$
d) $P\left(A_1 \cap A_2 \cap A_3\right) = P\left(A_1\right) P\left(A_2\right) P\left(A_3\right).$

3.5 Variáveis Aleatórias Discretas

As variáveis aleatórias, que definiremos formalmente a seguir, estão sempre presentes num experimento, e assumem valores numéricos de acordo que os critérios estabelecidos ao objetivo do experimento de modo a caracterizar o fenômeno estudado. Como estamos lidando com experimentos aleatórios essas variáveis são capazes de assumir vários valores sob a influência do acaso. De modo que, é impossível predizer com absoluta certeza qual valor a variável assumirá quando o experimento é realizado. Entretanto, vimos que todo experimento aleatório define um modelo probabilístico, a tripla $(\Omega, \mathcal{F}, P)$, e a partir dela podemos concluir como se comporta a variável aleatória. Se esta pode assumir valores de maneira contínua, isto é, na vizinhança de qualquer ponto de um intervalo, então trata-se de uma variável aleatória contínua, pois a quantidade de seus possíveis valores é infinita e não enumerável. Caso a variável possa assumir, apenas, uma quantidade de valores finita ou infinita enumerável, então estamos no caso discreto. Usaremos como base para esta definição Ross (2010, p.157)[2].

Definição 3.8. *Considere um experimento aleatório definido pelo modelo probabilístico* $(\Omega, \mathcal{F}, P)$. *Uma variável aleatória é uma função* $X : \Omega \rightarrow \mathbb{R}$ *tal que:*

$$\{w \in \Omega : X(w) \leq x\} \in \mathcal{F} \text{ para todo } x \in \mathbb{R}.$$

As variáveis aleatórias discretas são aquelas que assumem valores em um conjunto finito ou infinito enumerável.

A partir da Definição 3.8 podemos dizer que uma Variável Aleatória Discreta - VAD, determina um conjunto finito ou infinito enumerável de valores reais para os quais está associada uma probabilidade. Além disso, observe que através dela podemos estabelecer uma relação conveniente entre os elementos de Ω e um subconjunto de $\mathbb{R}$. Este conceito foi criado para viabilizar a análise numérica de fenônemos aleatórios, uma vez que dispondo apenas do conjunto $\mathcal{F}$ não podemos fazê-la.

Como o leitor pode notar, a VAD não tem sua razão de existência em si mesma. Mas é imprescindével para se estabelecer uma relação entre uma probabilidade e um número real. O que nos leva, naturalmente, ao conceito de Função de Probabilidade - FP e Função de Distribuição Acumulada -FDA:

Definição 3.9. *Seja uma Variável Aleatória Discreta X. Definimos sua Função de Probabilidade como a função* $p : \mathbb{R} \longrightarrow [0, 1]$ *obedecendo a seguinte lei de formação*

$$p(x) = P(X = x)$$

e Função de Distribuição Acumulada $F : \mathbb{R} \longrightarrow [0, 1]$ *sendo*

$$F(x) = \sum_{k \leq x} p(k). \tag{3}$$

Através da FP podemos analisar qual o comportamento da probabilidade para uma dada variável aleatória, ao passo que FDA nos fornece a soma das probabilidades para valores menores ou iguais a um determinado número real. Essas duas funções são essenciais para a caracterização de uma variável aleatória de modo que todo o estudo do comportamento das variáveis discretas se baseia nessas funções.

Para finalizar a caracterização das variáveis discretas traremos aqui duas características numéricas fundamentais que auxiliam a caracterizar uma variável. Estamos falando do valor esperado e da variância. De acordo com Ross (2010, p.160;168-169)[2] segue as definições.

Definição 3.10. *Seja X uma variável aleatória discreta com função de probabilidade $p(x)$. Então o valor esperado de X representado por $E[X]$ é definido como*

$$E[X] = \sum_{k|p(k)>0} kp(k). \tag{4}$$

Sua variância, representada por $Var[X]$, é dada por

$$Var[X] = E\left[(X-\mu)^2\right] \text{ onde } \mu = E[X].$$

Para facilitar os cálculos da variância daremos uma fórmula alternativa que decorre de sua definição.

Proposição 3.5. *Seja X uma variável aleatória discreta com função de probabilidade $p(x)$ e uma função $h : \mathbb{R} \to \mathbb{R}$. Então, temos:*
i) $E[h(X)] = \sum_{k|p(k)>0} h(k)p(k)$.
ii) $Var[X] = E[X^2] - (E[X])^2$.

Demonstração: Omitiremos a prova do item i). Da definição de variância temos que,

$$\begin{aligned}
Var[X] &= E\left[(X-\mu)^2\right] \text{ onde } \mu = E[X] \\
&= \sum_{x_i}(x_i-\mu)^2 P(X=x_i) \\
&= \sum_{x_i}(x_i^2 - 2\mu x_i + \mu^2) P(X=x_i) \\
&= \sum_{x_i} x_i^2 P(X=x_i) - 2\mu \sum_{x_i} x_i P(X=x_i) + \mu^2 \sum_{x_i} P(X=x_i) \\
&= E\left[X^2\right] - 2\mu E[X] + \mu^2 \\
&= E\left[X^2\right] - 2(E[X])^2 + (E[X])^2.
\end{aligned}$$

Assim,

$$Var[X] = E\left[X^2\right] - (E[X])^2. \tag{5}$$

□

Podemos entender a $Var[X]$ como o valor esperado do quadrado da diferença entre a média da variável e seus possíveis valores ponderados por suas probabilidades. Isto nos informa sobre a grau de concentração dos valores da variável em torno da esperança, sendo esse grau de concentração inversamente proporcional à variância.A seguir apresentamos três propriedades importantes do Valor Esperado.

Definição 3.11. *Considere um vetor* $V_n = (X_1, X_2, ..., X_n)$ *formado pelas variáveis aleatórias* $X_1, X_2, ..., X_n$*. A este vetor* V_n *denominamos* vetor aleatório.

Em particular, se todas as variáveis $X_1, X_2, ..., X_n$ são discretas, então v_n é um *vetor aleatório discreto.*

Definição 3.12. *Seja um vetor aleatório discreto* $V_n = (X_1, X_2, ..., X_n)$*.Se todos os eventos associados a qualquer subconjunto das variáveis* $X_1, X_2, ..., X_n$ *não altera a probabilidade de eventos associados a variáveis fora deste subconjunto então* $X_1, X_2, ..., X_n$ *são independentes.*

Proposição 3.6. *Seja* $V_n = (X_1, X_2, ..., X_n)$ *um vetor aleatório discreto e uma função* $h : \mathbb{R}^n \to \mathbb{R}$*, então*

$$E[h(V_n)] = \sum_{x_1|p(x_1)>0} \sum_{x_2|p(x_2)>0} ... \sum_{x_n|p(x_n)>0} h(x_1, x_2, ..., x_n) p(x_1, x_2, ..., x_n)$$

Demonstração: Omitiremos sua demonstração.

Proposição 3.7. *O valor esperado do produto entre um variável aleatória discreta* X *e um número real* a *é igual ao valor esperado do produto* aX *. Ou seja,* $E(aX) = aE(X)$*.*

Demonstração: Da definição de $E(X)$, temos $aE(X) = a \sum x_i P(X = x_i)$ para $i = 0, 1, ..., n$.

De $\{X = x_i\} \iff \{aX = ax_i\}$ segue que $P(aX = ax_i) = P(X = x_i)$.

Assim, $aE(X) = a \sum x_i P(X = x_i) = \sum ax_i P(aX = ax_i) = E(aX)$.

□

Proposição 3.8. *Sejam* $X_1, X_2, ..., X_n$ *variáveis aleatórias discretas e* $a_1, a_2, ..., a_n \in \mathbb{R}_*$*. O valor esperado de* $a_1X_1 + a_2X_2 + ... + a_nX_n$ *é igual à soma dos valores esperados de cada uma delas multiplicado pelas respectivas constantes* $a_1, a_2, ..., a_n$*. Isto é,*

$$E\left[\sum_{i=1}^{n} a_i X_i\right] = \sum_{i=1}^{n} a_i E[X_i].$$

Demonstração: Faremos por indução. Considere o caso $n = 2$.

$$E\left[a_1X_1 + a_2X_2\right] = \sum_{x_1|p(x_1)>0}\sum_{x_2|p(x_2)>0}\left(a_1x_1 + a_2x_2\right)P\left(X_1 = x_1, X_2 = x_2\right)$$

$$\begin{aligned}
&= \sum_{x_1|p(x_1)>0}\sum_{x_2|p(x_2)>0} a_1x_1P\left(X_1 = x_1, X_2 = x_2\right) + \sum_{x_1|p(x_1)>0}\sum_{x_2|p(x_2)>0} a_2x_2P\left(X_1 = x_1, X_2 = x_2\right)\\
&= a_1\sum_{x_1|p(x_1)>0} x_1 \sum_{x_2|p(x_2)>0} P\left(X_1 = x_1, X_2 = x_2\right) + a_2\sum_{x_2|p(x_2)>0} x_2 \sum_{x_1|p(x_1)>0} P\left(X_1 = x_1, X_2 = x_2\right)\\
&= a_1\sum_{x_1|p(x_1)>0} x_1P\left(X_1 = x_1\right) + a_2\sum_{x_2|p(x_2)>0} x_2P\left(X_2 = x_2\right)
\end{aligned}$$

Logo,

$$E\left(a_1X_1 + a_2X_2\right) = a_1E\left(X_1\right) + a_2E\left(X_2\right)$$

Suponha, então, que
$E(a_1X_1 + a_2X_2 + ... + a_{n-1}X_{n-1}) = a_1E(X_1) + a_2E(X_2) + ... + a_{n-1}E(X_{n-1})$
Agora considere
$E(a_1X_1 + ... + a_{n-1}X_{n-1} + a_nX_n) = E(a_1X_1 + ... + a_{n-1}X_{n-1}) + a_nE(X_n)$
e assim, pela hipótese de indução, segue o resultado

$$\begin{aligned}
E\left(a_1X_1 + a_2X_2 + ... + a_{n-1}X_{n-1} + a_nX_n\right) &= a_1E(X_1) + a_2E(X_2) + ... + a_nE(X_n)\\
E\left[\sum_{i=1}^{n} a_iX_i\right] &= \sum_{i=1}^{n} a_iE\left[X_i\right].
\end{aligned}$$

□

Proposição 3.9. *As variáveis aleatórias discretas $X_1, X_2, ..., X_n$ são independentes se, e somente se, para qualquer escolha de $x_1, x_2, ..., x_n$ temos que*

$$P\left(X_1 = x_1, X_2 = x_2, ..., X_n = x_n\right) = P\left(X_1 = x_1\right)P\left(X_2 = x_2\right)...P\left(X_n = x_n\right)$$

Demonstração: Omitida.

Proposição 3.10. *Seja as funções $h_i : \mathbb{R} \to \mathbb{R}$ para $i = 1, 2, ..., n$. Se as variáveis aleatórias discretas $X_1, X_2, ..., X_n$ são independentes, então*

$$E\left[h_1\left(X_1\right)h_2\left(X_2\right)...h_n\left(X_n\right)\right] = E\left[h_1\left(X_1\right)\right]E\left[h_2\left(X_2\right)\right]...E\left[h_n\left(X_n\right)\right]$$

Demonstração: Omitida.

Proposição 3.11. *Sejam $X_1, X_2, ..., X_n$ variáveis aleatórias discretas independentes e $a_1, a_2, ..., a_n \in \mathbb{R}_*$. A variância de $a_1X_1+a_2X_2+...+a_nX_n$ é igual à soma das variâncias de cada uma delas multiplicado pelas respectivas constantes $a_1, a_2, ..., a_n$. Isto é,*

$$Var\left[\sum_{i=1}^{n} a_iX_i\right] = \sum_{i=1}^{n} a_i^2 Var\left[X_i\right].$$

Demonstração: Usaremos indução. Da identidade (5) temos,

$$Var\left[a_1X_1 + a_2X_2\right] = E\left[(a_1X_1 + a_2X_2)^2\right] - \left(E\left[a_1X_1 + a_2X_2\right]\right)^2$$

Que é,

$$\begin{aligned}
&= E\left[a_1^2X_1^2 + 2a_1a_2X_1X_2 + a_2^2X_2^2\right] - \left(a_1E\left[X_1\right] + a_2E\left[X_2\right]\right)^2 \\
&= a_1^2E\left[X_1^2\right] + a_2^2E\left[X_2^2\right] + 2a_1a_2E\left[X_1X_2\right] - \left(a_1^2\left(E\left[X_1\right]\right)^2 + a_2^2\left(E\left[X_2\right]\right)^2 + 2a_1a_2E\left[X_1\right]E\left[X_2\right]\right) \\
&= a_1^2E\left[X_1^2\right] - a_1^2\left(E\left[X_1\right]\right)^2 + a_2^2E\left[X_2^2\right] - a_2^2\left(E\left[X_2\right]\right)^2 \\
&= a_1^2\left\{E\left[X_1^2\right] - \left(E\left[X_1\right]\right)^2\right\} + a_2^2\left\{E\left[X_2^2\right] - \left(E\left[X_2\right]\right)^2\right\}
\end{aligned}$$

Assim,

$$Var\left[a_1X_1 + a_2X_2\right] = a_1^2Var\left[X_1\right] + a_2^2Var\left[X_2\right]$$

Suponha por indução que

$$Var\left[\sum_{i=1}^{n} a_iX_i\right] = \sum_{i=1}^{n} a_i^2 Var\left[X_i\right].$$

Logo,

$$\begin{aligned}
Var\left[\sum_{i=1}^{n+1} a_iX_i\right] &= Var\left[\sum_{i=1}^{n} a_iX_i + a_{n+1}X_{n+1}\right] \\
Var\left[\sum_{i=1}^{n+1} a_iX_i\right] &= Var\left[\sum_{i=1}^{n} a_iX_i\right] + Var\left[a_{n+1}X_{n+1}\right] \\
Var\left[\sum_{i=1}^{n+1} a_iX_i\right] &= \sum_{i=1}^{n} a_i^2Var\left[X_i\right] + a_{n+1}^2Var\left[X_{n+1}\right] \\
Var\left[\sum_{i=1}^{n+1} a_iX_i\right] &= \sum_{i=1}^{n+1} a_i^2Var\left[X_i\right]
\end{aligned}$$

□

Exemplo 3.3. *Uma urna possui dez bolas, oito Azuis e duas Brancas. Cinco bolas são retiradas uma a uma ao acaso e com reposição. Seja* $X = X_1 + X_2 + X_3 + X_4 + X_5$ *onde*

$$X_i = \begin{cases} \text{1, se a i-ésima bola retirada é Azul} \\ \text{0, caso contrário.} \end{cases}$$

para $i = 1, 2, 3, 4, 5$*. Então,* $E[X] = 4$ *e* $Var[X] = \frac{4}{5}$*.*

Vamos utilizar a proposição 3.8. Para $i = 1, 2, 3, 4, 5$ é fácil ver que,

$$\begin{aligned} E[X_i] &= 0 \cdot P(X_i = 0) + 1 \cdot P(X_i = 1) \\ E[X_i] &= 0 \cdot \frac{2}{10} + 1 \cdot \frac{8}{10} \\ E[X_i] &= \frac{4}{5} \end{aligned}$$

Agora considere,

$$\begin{aligned} E[X] &= E[X_1 + X_2 + X_3 + X_4 + X_5] \\ E[X] &= E[X_1] + E[X_2] + E[X_3] + E[X_4] + E[X_5] \\ E[X] &= \frac{4}{5} + \frac{4}{5} + \frac{4}{5} + \frac{4}{5} + \frac{4}{5} \\ E[X] &= 4 \end{aligned}$$

Para variância note que,

$$\begin{aligned} E\left[X_i^2\right] &= 0^2 \cdot \frac{2}{10} + 1^2 \cdot \frac{8}{10} \\ E\left[X_i^2\right] &= \frac{4}{5} \end{aligned}$$

Assim,

$$\begin{aligned} Var[X_i] &= E\left[X_i^2\right] - (E[X_i])^2 \\ Var[X_i] &= \frac{4}{5} - \left(\frac{4}{5}\right)^2 \\ Var[X_i] &= \frac{20}{25} - \frac{16}{25} \\ Var[X_i] &= \frac{4}{25} \end{aligned}$$

Aplicando a proposição 3.11 resulta,

$$\begin{aligned} Var\,[X] &= Var\left[\sum_{i=0}^{5} X_i\right] \\ Var\,[X] &= \sum_{i=0}^{5} Var\,[X_i] \\ Var\,[X] &= 5 \cdot \frac{4}{25} \\ Var\,[X] &= \frac{4}{5} \end{aligned}$$

As proposições 3.6 e 3.7 oferecem um novo caminho para o cálculo da esperança e variância como vemos neste exemplo.

Proposição 3.12. *Para X variável aleatória discreta e t um número real não nulo é válida a equação $\sum_{k=0}^{\infty} \frac{X^k t^k}{k!} = e^{Xt}$.*

Demonstração:

Como X assume valores reais o produto $Xt = w \in \mathbb{R}$. Considere, pois, a função $f : \mathbb{R} \longrightarrow \mathbb{R}$ definida como $f(w) = e^w$.

Da série de *Maclaurin* (Leithold, 1994, p. 678)[1] temos:

$$S_n(w) = \sum_{i=0}^{n} \frac{f^{(i)}(0)}{i!} w^i.$$

Usando S_n podemos obter uma aproximação de e^w tão precisa quanto desejarmos com erro dado por $R_n(w) = \frac{f^{(n+1)}(c)}{(n+1)!} w^{n+1}$ onde $c \in (0, w)$

Em se tratando de uma função exponencial temos que $n \in \mathbb{N} \Longrightarrow f^{(n+1)}(w) = f^{(n)}(w) \quad \forall\ w \in \mathbb{R}$. Em particular, $f^{(n+1)}(0) = f^{(n)}(0) = 1$. Logo,

$$S_n(w) = \sum_{i=0}^{n} \frac{w^i}{i!}.$$

Como $R_n(w) \longrightarrow 0 \quad \forall\ w \in \mathbb{R}$, segue que $\lim_{n\to\infty} S_n(w) = e^w$.

Portanto,

$$\sum_{i=0}^{\infty} \frac{X^k t^k}{k!} = e^{Xt}.$$

$\square$

3.5.1 Função Geradora de Momentos - FGM

Essa função é definida, para auxiliar no cálculo de várias características numéricas de uma variável aleatória como, por exemplo, seu valor esperado e sua variância. De acordo com Stirzaker (2003, p. 245)[3],

Definição 3.13. *Considere a variável aleatória discreta X tal que $E\left[X^k\right] < \infty\ \forall\ k \in \mathbb{N}_*$. A* Função Geradora de Momentos *de X é definida como a função $M_X : \mathbb{R} \longrightarrow \mathbb{R}$ dada pela expressão:*

$$M_X(t) = \sum_{k=0}^{\infty} \frac{E\left[X^k\right] t^k}{k!}.$$

Conjugando Proposição 3.7 e 3.8 concluimos que:

$$\begin{aligned}
M_X(t) &= \sum_{k=0}^{\infty} \frac{E\left[X^k\right] t^k}{k!} \\
M_X(t) &= E\left[\sum_{k=0}^{\infty} \frac{X^k t^k}{k!}\right] \\
M_X(t) &= E\left[e^{Xt}\right]. \qquad (6)
\end{aligned}$$

Com o auxílio de $M_x(t)$ vamos calcular $E\,(X)$ e $Var\,(X)$. Derivando a equação 6 em relação a t,

$$\frac{\partial M_X}{\partial t}(t) = E\left[Xe^{Xt}\right].$$

Para $t = 0$ seque que,

$$\frac{\partial M_X}{\partial t}(0) = E\,[X].$$

Derivando pela segunda vez em t,

$$\frac{\partial^2 M_X}{\partial^2 t^2}(t) = E\left[X^2 e^{Xt}\right].$$

Fazendo $t = 0$ temos que,

$$\frac{\partial^2 M_X}{\partial^2 t^2}(0) = E\left[X^2\right].$$

Assim, um caminho alternativo para o cálculo da variância é $\frac{\partial^2 M_X}{\partial^2 t^2}(0) - \left[\frac{\partial M_X}{\partial t}(0)\right]^2$. Além disso, apesar de não ser a tônica deste trabalho, a FGM caracteriza uma variável aleatória, de modo que variáveis com uma mesma FGM são idênticas e variáveis idênticas tem uma mesma FGM. Não iremos demonstrar esse resultado, mas ele é muito importante em estudos teóricos mais avançados sobre probabilidade.

Exemplo 3.4. *Considere a variável X cuja função de probabilidade é dada por $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$. Então $E[X] = np$.*

De fato,

$$\begin{aligned}
M_X(t) &= E\left[e^{Xt}\right] \\
&= \sum_{n=1}^{n} e^{kt} P(X = k) \\
&= \sum_{n=1}^{n} e^{kt} \binom{n}{k} p^k (1-p)^{n-k} \\
&= \sum_{n=1}^{n} \binom{n}{k} \left(pe^t\right)^k (1-p)^{n-k} \text{ ; Teorema Binomial (p.21)} \\
M_X(t) &= \left(pe^t + 1 - p\right)^n
\end{aligned}$$

Derivando $M_X(t)$ em t, temos

$$\frac{\partial M_X}{\partial t}(t) = n\left(pe^t + 1 - p\right)^{n-1} pe^t$$

Como $E[X] = \frac{\partial M_X}{\partial t}(0)$, segue que

$$\begin{aligned}
E[X] &= n\left(pe^0 + 1 - p\right)^{n-1} pe^0 \\
&= n(p + 1 - p)^{n-1} p \\
&= np
\end{aligned}$$

3.6 Variáveis Aleatórias Contínuas

As variáveis aleatórias contínuas são aquelas cujo conjunto de valores possiveis é não-enumerável. Uma variável que mede o tempo de vida útil de uma lâmpada ou que registra o momento de passagem de um veículo num certo ponto da rodovia são bons exemplos disso.

A definição de variável aleatória contínua é dada a seguir:

Definição 3.14. *Uma variável aleatória X é contínua se existe um função f:$\mathbb{R} \to \mathbb{R}$ tal que*

i) $\forall x \in D(f) \Longrightarrow f(x) \geq 0$;

ii) $\int\limits_{-\infty}^{\infty} f(x)\, dx = 1;$

Proposição 3.13. *Seja B é um conjunto de números mensurável, então* $P(x \in B) = \int\limits_{B} f(x)\, dx.$

Demonstração: Omitida.

A função f é chamada de função "densidade"da variável X. Veja alguns exemplos:

Exemplo 3.5. *A função* $f(x) = \begin{cases} \lambda e^{-\lambda x}; \text{ para } x \geq 0. \\ 0; \text{ caso contrário.} \end{cases}$ *é uma função densidade para* $\lambda \in \mathbb{R}_*^+$ *e, por isso, está vinculada a uma variável aleatória contínua.*

De fato, f satisfaz os itens i) e ii) da definição acima. Pois,
$\lambda e^{-\lambda x} > 0$ constata-se (i). Derivando $-e^{-\lambda x}$ em x resulta $f(x) = \lambda e^{-\lambda x}$. Assim,

$$\begin{aligned}
\int\limits_{-\infty}^{\infty} f(x)\, dx &= \lim_{b \to -\infty} \int\limits_{b}^{0} 0dx + \lim_{a \to \infty} \int\limits_{0}^{a} \lambda e^{-\lambda x} dx \\
\int\limits_{-\infty}^{\infty} f(x)\, dx &= \lim_{b \to -\infty} k - k + \lim_{a \to \infty} -e^{-\lambda a} - \left(-e^{-\lambda 0}\right) \\
\int\limits_{-\infty}^{\infty} f(x)\, dx &= 0 - \frac{1}{e^{\infty}} + 1
\end{aligned}$$

Como $\frac{1}{e^{\infty}} = 0$ segue que,

$$\int\limits_{-\infty}^{\infty} f(x)\, dx = 1; \text{ constata-se } (ii).$$

Uma variável aleatória X vinculada a esta função densidade é chamada exponencial com parâmetro λ. Notação: $X \backsim \exp(\lambda)$.

Exemplo 3.6. *Sejam as constantes reais positivas* μ, σ *e a função* $f : \mathbb{R} \to \mathbb{R}$. *Então,*

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

é densidade associada a um tipo de variável aleatória contínua.

A variável aleatória contínua X cuja função densidade é dada por esta função chama-se normal com parâmetros μ e σ (notação: $X \backsim normal\,(\mu, \sigma))$. Para $\mu = 0$ e $\sigma = 1$ temos uma variável normal padrão. Isto é, $X \backsim normal\,(0, 1)$.

Exemplo 3.7. *Seja* $f(x) = \begin{cases} \frac{\lambda e^{-\lambda x}\lambda^{\alpha-1}x^{\alpha-1}}{\Gamma(\alpha)};\ x \geq 1 \\ 0;\ \textit{caso contrário} \end{cases}$ *onde* $\lambda, \alpha \in \mathbb{R}_*^+$ *e* $\Gamma(\alpha) = \int\limits_0^\infty e^{-y}y^{\alpha-1}dy$.

Variáveis aleatórias cuja densidade é f são chamadas gama de parâmetros α e λ. Notação: $X \backsim gama\,(\alpha, \lambda)$.

Em particular para $\alpha \in \mathbb{N}$ a função densidade da variável gama é dada por

$$f(x) = \begin{cases} \frac{\lambda e^{-\lambda x}\lambda^{\alpha-1}x^{\alpha-1}}{(\alpha-1)!};\ x \geq 1 \\ 0;\ \text{caso contrário} \end{cases}$$

3.6.1 Caracterização de Variáveis Aleatórias Contínuas

Definição 3.15. *A função de distribuição acumulada de uma variável aleatória contínua X de densidade f é dada por*

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)\,du$$

Exemplo 3.8. *Se* $X \backsim \exp(\lambda)$ *então* $F(x) = \begin{cases} 1 - e^{-\lambda x};\ x \geq 0 \\ 0;\ \textit{caso contrário} \end{cases}$.

Por definição,

$$\begin{aligned} F(x) &= \int_{-\infty}^{x} f(u)\,du \\ F(x) &= \lim_{b \to -\infty} \int\limits_b^0 0du + \int\limits_0^x \lambda e^{-\lambda u}du \\ F(x) &= \lim_{b \to -\infty} (k - k) + \left[-e^{-\lambda x} - \left(-e^{-\lambda 0}\right)\right] \\ F(x) &= 1 - e^{-\lambda x} \end{aligned}$$

Para $x < 0$ perceba que $F(x) = 0$. Portanto constata-se o enunciado.

Exemplo 3.9. *Para as variáveis contínuas a função de distribuição acumulada é dada pela a integral da função densidade. Para $X \backsim normal\,(0,1)$ tem-se:*

$$P\,(X \leq a) = \int\limits_{-\infty}^{a} \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}} dx$$

Como essa integral não tem solução analítica, foi-se elaborada uma tabela através de técnicas de cálculo numérico. A tabela a seguir fornece os valores destas integrais para alguns valores de a.

Distribuição Normal Padrão

Função tabelada: $\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-x^2/2}\,dx$ para $z \geq 0$.

Parte inteira e primeira decimal de z	Segunda decimal de z										
	0	1	2	3	4	5	6	7	8	9	
0,0	0,5	0,504	0,508	0,512	0,516	0,5199	0,5239	0,5279	0,5319	0,5359	0,0
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753	0,1
0,2	0,5793	0,5832	0,5871	0,591	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141	0,2
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,648	0,6517	0,3
0,4	0,6554	0,6591	0,6628	0,6664	0,67	0,6736	0,6772	0,6808	0,6844	0,6879	0,4
0,5	0,6915	0,695	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,719	0,7224	0,5
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549	0,6
0,7	0,758	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852	0,7
0,8	0,7881	0,791	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133	0,8
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,834	0,8365	0,8389	0,9
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621	1,0
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,877	0,879	0,881	0,883	1,1
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,898	0,8997	0,9015	1,2
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177	1,3
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319	1,4
1,5	0,9332	0,9345	0,9357	0,937	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441	1,5
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545	1,6
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633	1,7
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706	1,8
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,975	0,9756	0,9761	0,9767	1,9
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817	2,0
2,1	0,9821	0,9826	0,983	0,9834	0,9838	0,9842	0,9846	0,985	0,9854	0,9857	2,1
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,989	2,2
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916	2,3
2,4	0,9918	0,992	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936	2,4
2,5	0,9938	0,994	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952	2,5
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,996	0,9961	0,9962	0,9963	0,9964	2,6
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,997	0,9971	0,9972	0,9973	0,9974	2,7
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,998	0,9981	2,8
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986	2,9
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,999	0,999	3,0
3,1	0,999	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993	3,1
3,2	0,9993	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995	3,2
3,3	0,9995	0,9995	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997	3,3
3,4	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998	3,4
3,5	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	3,5
3,6	0,9998	0,9998	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	3,6
3,7	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	3,7
3,8	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	3,8
3,9	1,	1,	1,	1,	1,	1,	1,	1,	1,	1,	3,9

Tabela 3.1.

Definição 3.16. *Seja X variável aleatória contínua, com densidade f, então sua esperança é definida por*

$$E(X) = \int_{-\infty}^{+\infty} xf(x)\,dx$$

Exemplo 3.10. *Se $X \backsim \exp(\lambda)$ então $E(X) = \frac{1}{\lambda}$.*

De fato, basta calcularmos a integral:

$$\begin{aligned}
E(X) &= \int_{-\infty}^{+\infty} xf(x)\,dx \\
E(X) &= \lim_{b \to -\infty} \int_{b}^{0} x0dx + \lim_{a \to \infty} \int_{0}^{a} x\lambda e^{-\lambda x} dx \\
E(X) &= \lim_{a \to \infty} \int_{0}^{a} x\lambda e^{-\lambda x} dx
\end{aligned}$$

Mas,

$$\begin{aligned}
\int x\lambda e^{-\lambda x} dx &= -xe^{-\lambda x} - \int e^{-\lambda x} dx \\
&= -\frac{x}{e^{\lambda x}} - \frac{1}{\lambda e^{\lambda x}} \\
&= \frac{-1}{\lambda e^{\lambda x}} (\lambda x + 1)
\end{aligned}$$

Assim,

$$\begin{aligned}
E(X) &= \lim_{a \to \infty} \frac{-1}{\lambda e^{\lambda a}} (\lambda a + 1) - \left[\frac{-1}{\lambda e^{\lambda 0}} (\lambda 0 + 1) \right] \\
E(X) &= -\lim_{a \to \infty} \frac{a}{e^{\lambda a}} - \lim_{a \to \infty} \frac{1}{\lambda e^{\lambda a}} + \frac{1}{\lambda}
\end{aligned}$$

Como ambos limites tendem a zero segue que,

$$E(X) = \frac{1}{\lambda}$$

Para as variáveis aleatórias contínuas a variância é definida a seguir.

Definição 3.17. *Seja X uma variável aleatória contínua cuja densidade é f e $E(X) = \mu$. Então, sua variância é*

$$Var(X) = \int_{-\infty}^{+\infty} (x - \mu)^2 f(x)\,dx$$

Proposição 3.14. *Seja a função $h : \mathbb{R} \to \mathbb{R}$ e X uma variável aleatória contínua com densidade f. Então,*

$$E[h(X)] = \int_{-\infty}^{+\infty} h(x) f(x) dx$$

Demonstração: Omitida

Exemplo 3.11. *Para $X \backsim \exp(\lambda)$ temos $Var(X) = \frac{1}{\lambda^2}$*

Exemplo 3.12. *Se $X \backsim normal(\mu, \sigma^2)$, então $E(X) = \mu$ e $Var(X) = \sigma^2$.*

Exemplo 3.13. *Quando $X \backsim gama(\alpha, \lambda)$ temos $E(X) = \frac{\alpha}{\lambda}$ e $Var(X) = \frac{\alpha}{\lambda^2}$.*

3.6.2 Alguns resultados do caso contínuo

Nesta seção apresentaremos os mais importantes resultados teóricos, utilizados neste trabalho, inerentes ao caso contínuo da teoria de probabilidade. Omitiremos suas provas por não ser a tônica deste trabalho.

Teorema 3.1. (*Lei Fraca dos Grandes Números*) *Seja $X_1, X_2, ..., X_n$ uma sequência de variáveis aleatórias independentes e identicamente distribuídas, cada uma com média $E(X_i) = \mu$. Então, para qualquer $\epsilon > 0$*

$$\lim_{n\to\infty} P\left(\left|\frac{X_1 + X_2 + ... + X_n}{n} - \mu\right| \geq \epsilon\right) = 0$$

Demonstração: Ver Barry James [8]

Teorema 3.2. (*Lei Forte dos Grandes Números*) *Seja $X_1, X_2, ..., X_n$ uma sequência de variáveis aleatórias independentes e identicamente distribuídas, cada uma com média $E(X_i) = \mu$. Então quando $n \to \infty$ temos*

$$P\left(\frac{X_1 + X_2 + ... + X_n}{n} \to \mu\right) = 1$$

Demonstração: Ver Barry James [8]

Exemplo 3.14. *Suponha uma sequência de tentativas independentes de um experimento aleatório. Seja A um evento de interesse e $P(A)$ a probabilidade de que A ocorra em alguma tentativa particular. Então, $P(A)$ pode ser estimado repetindo-se o experimento.*

Defina $X_i = \begin{cases} 1, \text{ se } A \text{ ocorre na i-ésima tentativa} \\ 0, \text{ caso contrário} \end{cases}$

Pela Lei Forte dos Grandes Números com probabilidade 1,

$$\frac{X_1 + X_2 + ... + X_n}{n} \to E(X_1) = P(A).$$

Como $X_1 + X_2 + ... + X_n$ é o número de vezes que o evento A ocorre nas n primeiras tentativas, podemos inferir que com probabilidade 1 a proporção limite de vezes que o evento A ocorre é exatamente $P(A)$. Intuitivamente, quando n cresce a proporção de vezes que A ocorre se aproxima de $P(A)$. Este fato pode ser usado para estimar a probabilidade de um evento A quando o experimento pode ser replicado muitas vezes.

Teorema 3.3. (*Teorema Central do Limite*)*Sejam variáveis aleatórias* $T_1, T_2, ..., T_n$ *identicamente distribuídas tais que* $E[T_i] = \mu < \infty$ *e* $0 < Var[T_i] = \sigma^2 < \infty$. *Então, a sequência de variáveis aleatórias*

$$Z_n = \frac{T_1 + T_2 + ... + T_n - n\mu}{\sigma\sqrt{n}}$$

converge em distribuição para a normal padrão quando $n \to \infty$. *Isto é, para* $a \in \mathbb{R}$ *temos*

$$P(Z_n \leq a) \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{\frac{-x^2}{2}} dx. \text{ quando } n \to \infty.$$

Demonstração: Ver Barry James [8]

O teorema central do limite pode ser representado pela notação

$$\frac{T_1 + T_2 + ... + T_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{d} N(0,1).$$

Onde d representa convergência em distribuição.

Exemplo 3.15. *Considere o lançamento de um dado honesto 600 vezes. Então, a probabilidade de se obter face cinco* 120 *vezes pode ser aproximada através da distribuição normal padrão.*

Defina $X_i = \begin{cases} 1, \text{ para "face cinco"na i-ésima jogada} \\ 0, \text{ caso contrário} \end{cases}.$

Então, $S_n = X_1 + X_2 + ... + X_n$ é o número de vezes que se obteve face cinco nos n primeiros lançamentos. Claramente, $S_n \backsim binomial\left(n, \frac{1}{6}\right)$. Note também que as variáveis X_i com $i = 1, 2, ..., 600$ são independentes e identicamente distribuídas com média $\mu = E(X_i) = \frac{1}{6}$ e variância $\sigma^2 = Var(X_i) = \frac{5}{36}$.

Pelo Teorema Central do Limite temos,

$$Z_n = \frac{S_n - \frac{n}{6}}{\frac{1}{6}\sqrt{5n}} \xrightarrow{d} Z \backsim N(0,1)$$

Efetuando os cálculos,

$$\begin{aligned} Z_{600} &= \frac{S_{600} - 600 \cdot \frac{1}{6}}{\frac{\sqrt{5}}{6} \cdot \sqrt{600}} \\ Z_{600} &= \frac{S_{600} - 100}{10\sqrt{\frac{5}{6}}} \end{aligned}$$

Logo,

$$\begin{aligned} P(S_{600} \leq 120) &= P\left(\frac{S_{600} - 100}{10\sqrt{\frac{5}{6}}} \leq \frac{120 - 100}{10\sqrt{\frac{5}{6}}}\right) \\ &\implies P(S_{600} \leq 120) = P(Z_{600} \leq 2,2) \\ &\quad\; P(Z_{600} \leq 2,2) \overset{TCL}{\backsimeq} P(Z \leq 2,2) \\ P(S_{600} \leq 120) &= 0,9861 \end{aligned}$$

Para fins práticos e número de repetições significativamente grande $(n = 600)$ o cálculo da probabilidade se torna mais econômico através da normal padrão. Dedicamos a próxima seção inteiramente à variáveis aleatórias binomiais, que são o campo de alcance deste ensaio.

4 Variáveis binomiais: Elementos teóricos

4.1 Ensaio de Bernoulli

Quando temos interesse na ocorrência ou não de um conjunto de eventos de um determinado experimento aleatório, com espaço amostral finito, estudamos seu comportamento associando a este experimento uma variável aleatória Y que é binária, isto é, recebe apenas dois valores: "0" quando o evento não ocorre e "1" quando o evento de interesse ocorre. Um ensaio com estas caraterísticas é chamado *Ensaio de Bernoulli.*

Suponha que neste experimento a probabilidade de ocorrência do evento que se deseja observar é um número real p entre 0 e 1. Então podemos escrever que

$$P(Y=y)=\begin{cases} p, \text{ se } y=1; \\ 1-p, \text{ se } y=0. \end{cases}$$

Considere o seguinte exemplo: Alguém lança dois dados honestos de seis faces e anota a soma das faces sorteadas. Suponha que estejamos interessados na ocorrência da soma igual a onze. Qual é a função de probabilidade da variável de interesse associada a esse experimento aleatório?

O espaço amostral para este experimento é $\Omega=\{(i,j)\,;i,j=1,2,3,4,5,6\}$. A soma dos dois números sorteados nesse experimento resulta um número natural maior que seis e menor que treze. Para a soma ser onze temos dois resultados favoráveis: $(5,6)$ e $(6,5)$. Como os dados são honestos e o resultado de um dado não influencia o resultado do outro, podemos dizer que a probabilidade da soma ser onze é $\frac{1}{18}$. O resultado de interesse que permeia a questão é a ocorrência ou não deste evento (soma igual a onze), então associamos uma variável binária Y ao experimento de modo que $Y=1$ quando o experimento resultar onze e $Y=0$ caso contrário. Logo, a função de probabilidade - FP de Y é dada por,

$$P(Y=y)=\begin{cases} \frac{1}{18}, \text{ se } y=1; \\ \frac{17}{18}, \text{ se } y=0. \end{cases}$$

Note que o modo como definimos uma variável aleatória para um determinado experimento não depende apenas das condições do experimento, mas quais conjuntos de eventos queremos estudar o comportamento. Neste exemplo, o espaço amostral é o par ordenado (i,j) onde $i=1,2,...,6$ e $j=1,2,...,6$ também. E o evento de interesse é $\{(5,6);(6,5)\}$. É óbvio que se o foco de nosso estudo mudasse, por exemplo, para "soma das faces sorteadas igual a sete"então as probabilidades seriam outras. Contudo a variável ainda seria do tipo binária, pois teríamos interesse na ocorrência ou não da soma igual a dez.

4.2 Experimento binomial

A seguir trataremos de características e condições necessárias a um experimento aleatório binomial. Este, como veremos, é uma extensão natural do modelo de *Bernoulli*. Considere uma sequência de n repetições de ensaios de Bernoulli onde há independência entre os ensaios e cada repetição tem mesma probabilidade de sucesso igual $p \in [0,1]$. Então, temos um experimento binomial. Note que do mesmo modo que no ensaio de Bernoulli a noção de sucesso é relativa, pois esta depende do interesse do observador. Vamos ilustrar o experimento binomial a partir da situação onde se lança uma moeda não-honesta a fim de observar o evento "ocorre face coroa virada para cima na i-ésima repetição".

Exemplo 4.1. *Lançar uma moeda* 10 *vezes sabendo-se que a probabilidade de ocorrer "coroa"é* $\frac{12}{25}$.

Evidentemente, ao lançarmos essa moeda 10 vezes nós temos um experimento binomial, pois para cada lançamento da moeda temos apenas duas possibilidades coroa (sucesso) ou cara (fracasso). Além disso, cada repetição do experimento tem probabilidade $\frac{12}{25}$ de sucesso e $\frac{13}{25}$ de fracasso e constata-se independência entre as repetições, pois constatando-se "cara"no k-ésimo lançamento a probabilidade de sair "coroa"em algum j-ésimo lançamento, com $j \neq k$, é $\frac{12}{25}$. Isto é, a mesma probabilidade de ocorrer "coroa"sem nenhuma informação anterior. Matematicamente, escrevemos

$$P\left[S_j \mid (S_k)^c\right] = P\left(S_j\right) \text{ sempre que } j \neq k$$

onde S_j : "sucesso no j-ésimo lançamento da moeda"e $j,k = 1,2,...,10$.

Outra situação que se enquadra nas hipóteses do modelo binomial é o lançamento de um dado. Veja o exemplo abaixo:

Exemplo 4.2. *Suponha que ao lançar um dado honesto de seis faces cem vezes desejamos observar o evento "face 3 virada para cima".*

Sendo o dado honesto temos as seguintes probabilidades associadas aos possíveis resultados de cada repetição.

Sair face	1	2	3	4	5	6
Probabilidade	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Como nosso interesse está na ocorrência do evento aleatório "face 3 virada para cima no i-ésimo lançamento"podemos entendê-lo como sucesso de um ensaio de Bernoulli e o evento "face 3 não virada para cima"como fracasso. Deste modo, a partir da tabela de probabilidades a cima concluímos que nossa probabilidade de sucesso é $\frac{1}{6}$ ao passo que a probabilidade de fracasso é $\frac{5}{6}$. Além disso, como no exemplo da moeda constata-se independência entre todos os lançamentos de modo que o conhecimento da ocorrência de quaisquer eventos não altera a probabilidade de ocorrência de um evento que não pertencente ao conjunto daqueles eventos. Portanto, temos cem repetições independentes de um ensaio de Bernoulli com probabilidade de sucesso igual a $\frac{1}{6}$.

Outro tipo de problema que podemos modelar pelo experimento binomial é a retirada, com reposição, de bolas azuis de uma urna onde existem bolas azuis e brancas. Acompanhe a situação a seguir:

Exemplo 4.3. *Considere uma urna contendo dez bolas idênticas, exceto pela cor. Dentre as dez bolas, sete são azuis e três brancas. Suponha que são feitas vinte retiradas ao acaso em que retira-se uma bola em cada retirada e esta é devolvida à urna logo em seguida e que temos interesse no evento "bola azul foi retirada".*

Das características próprias ao experimento deduz-se facilmente que a probabilidade de se retirar bola azul na i-ésima retirada é $\frac{3}{10}$ e que cada retirada é um ensaio de Bernoulli. Além disso, como as bolas retiradas são devolvidas à urna, a ocorrência de algum evento aleatório anterior ou posterior não altera a probabilidade de sucesso em alguma retirada. Ou seja, as repetições são independentes umas das outras. Portanto, caracteriza um experimento binomial.

Em síntese, para podermos afirmar que um experimento aleatório é binomial devemos mostrar que tal experimento obedece a três hipóteses:

· Há repetição de ensaios de Bernoulli;

· Todas as repetições tem mesma probabilidade de sucesso;

· Em cada repetição a probabilidade de sucesso independe da ocorrência nas demais repetições;

O próximo experimento, também envolve retirar bolas de uma urna e servirá como contra-exemplo pois não contempla todas as hipóteses acima. Veja:

Exemplo 4.4. *(Contra-exemplo) Considere uma urna contendo oito bolas idênticas, exceto pela cor. Dentre as oito bolas, quatro são azuis e quatro brancas. Suponha que são feitas três retiradas uma a uma ao acaso e sem reposição. Suponha interesse no evento "bola branca foi retirada".*

Neste experimento, podemos perceber que temos vários ensaios de Bernoulli, pois em cada retirada temos sucesso (bola branca) e fracasso (bola azul). No entanto, não há independência entre as retiradas, pois a probabilidade de se retirar bola branca na terceira retirada sabendo-se que na primeira e segunda retiradas sairam azuis é maior que se saissem bolas brancas nas duas primeiras retiradas. Isto porque, para o caso de já haverem saído duas bolas azuis teremos três bolas brancas e um azul dentro da urna, o que implica sucesso com probabilidade de $\frac{3}{4}$. No caso de haverem saído bolas brancas teríamos probabilidade igual a $\frac{1}{4}$ de se retirar a terceira bola branca. Portanto, não é um experimento binomial.

Agora vamos estudar o comportamento das variáveis aleatórias binomiais e como se distribui suas probabilidades.

4.3 Variável aleatória binomial

Uma variável aleatória definida a partir de um experimento binomial com n repetições de ensaios de Bernoulli de mesma probabilidade de sucesso p pode receber apenas valores inteiros não-negativos, pois esta variável registra quantos sucessos ocorreram na execução do experimento binomial. Sendo assim, na hipótese de n fracassos a variável recebe *zero*, quando ocorre um sucesso durante o experimento ela recebe *um*, para dois sucessos receberá *dois* etc. De modo que para k sucessos essa variável aleatória receberá k. Assim, percebemos que a variável receberá algum dos valores do conjunto $\{0, 1, 2, ..., n\}$ ao fim desse experimento binomial. À variável aleatória associada a este experimento chamaremos *Variável Binomial.* A seguir temos uma proposição sobre sua distribuição de probabilidade.

Proposição 4.1. *Considere um experimento binomial com n repetiões de ensaios de Bernoulli com probabilidade p de sucesso em cada repetição. Seja X o número de sucessos obtidos na execução do experimento binomial. Então a distribuição de probabilidade* *de X é dada por,*

$$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k},\ \ k=0,1,2,...,n.$$

X é chamada variável aleatória binomial.

Demonstração: Sejam os eventos A_i : "sucesso na i-ésima repetição de um ensaio de Bernoulli com probabilidade p". Então,

$$\begin{aligned}
\{X=k\} &= \bigcup_{1\leq i_1<i_2<...<i_k\leq n}\left[\left(\bigcap_{j=1}^{k}A_{i_j}\right)\bigcap\left(\bigcap_{j=1|j\notin\{i_1,i_2,...,i_k\}}^{n}A_j\right)\right]\Longrightarrow\\
P\left(X=k\right) &= P\left\{\bigcup_{1\leq i_1<i_2<...<i_k\leq n}\left[\left(\bigcap_{j=1}^{k}A_{i_j}\right)\bigcap\left(\bigcap_{j=1|j\notin\{i_1,i_2,...,i_k\}}^{n}A_j\right)\right]\right\}\\
&= \sum_{1\leq i_1<i_2<...<i_k\leq n}P\left[\left(\bigcap_{j=1}^{k}A_{i_j}\right)\bigcap\left(\bigcap_{j=1|j\notin\{i_1,i_2,...,i_k\}}^{n}A_j\right)\right]\\
&= \sum_{1\leq i_1<i_2<...<i_k\leq n}\left[\prod_{j=1}^{k}P\left(A_{i_j}\right)\right]\left[\prod_{j=1|j\notin\{i_1,i_2,...,i_k\}}^{n}P\left(A_j\right)\right]\\
&= \sum_{1\leq i_1<i_2<...<i_k\leq n}\prod_{i=1}^{k}p\prod_{i=1}^{n-k}(1-p)\\
&= \binom{n}{k}p^k\left(1-p\right)^{n-k}.
\end{aligned}$$

Portanto,

$$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k};\ k=0,1,2,...,n. \tag{7}$$

□

A equação 7 é chamada de *função de probabilidade* da variável binomial X com parâmetros n e p, ou melhor, $X\backsim binomial(n;p)$.

A partir dessa função, que associa a cada possível valor de X uma probabilidade do intervalo $(0,1)$, desenvolveremos nosso estudo e chegaremos a importantes conclusões a respeito dos experimentos binomiais. Uma delas segue imediatamente da proposição anterior e será demonstrada conforme (Ross, 2010, p.177) [2].

Proposição 4.2. *Seja $X\backsim binomial(n;p)$ onde $n\in\mathbb{N}_*$ e $p\in\mathbb{R}\cap(0,1)$, então $P(X=k)$ é crescente no intervalo $[0,a]$ e descrescente em $(a,n]$ onde a é o maior inteiro menor ou igual a $(n+1)\,p$.*

Demonstração: Para $k = 1, 2, ..., n$ considere a razão

$$\begin{aligned}
\frac{P(X=k)}{P(X=k-1)} &= \frac{\binom{n}{k} p^k (1-p)^{n-k}}{\binom{n}{k-1} p^{k-1} (1-p)^{n-k+1}} \\
&= \frac{\frac{n!}{(n-k)!k!} p}{\frac{n!}{(n-k+1)!(k-1)!} (1-p)} \\
&= \frac{(n-k+1)!\,(k-1)!p}{(n-k)!k!(1-p)} \\
&= \frac{(n-k+1)\,p}{k(1-p)}.
\end{aligned}$$

Assim, quando $P(X = k)$ é crescente, isto é, $P(X = k) > P(X = k-1)$ então $(n-k+1)\,p > k(1-p) \iff k < (n+1)\,p$. Por outro lado, $P(X = k)$ descrescente $\iff P(X = k) > P(X = k-1) \implies (n-k+1)\,p > k(1-p) \iff k > (n+1)\,p$.

$\square$

A gráfico da figura 4.1 ilustra o resultado da Proposição 4.2. Considere nesse exemplo $X \backsim binomial(7; \frac{2}{5})$:

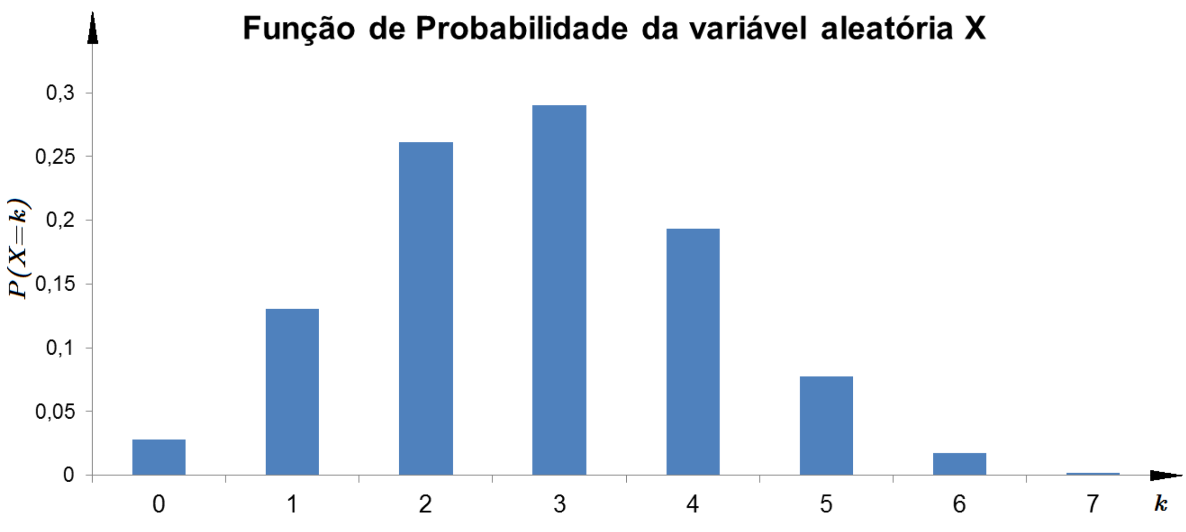

Figura 4.1.

Neste exemplo, note que $(n+1)\,p = (7+1)\,\frac{2}{5} = \frac{16}{5}$. O maior inteiro menor que $\frac{16}{5}$ é 3. Assim, percebemos que, de fato, $P(X = k)$ é crescente em $[0, 3]$ atingindo seu valor máximo em $k = 3$ e depois $P(X = k)$ descresce até o parâmetro $n = 7$.

4.4 Caracterização

Atualmente os esforços de vários estudiosos da área de probabilidade e estatística tem como ponto inicial o comportamento e a caracterização de variáveis aleatórias observadas no estudo da natureza, da sociedade e de estruturas cibernéticas. Assim, nasce a necessidade da criação de modelos probabilísticos para experimentos aleatórios que possuem variáveis aleatórias que não se enquadram a modelos já conhecidos. Nosso interesse neste trabalho não é propor um novo modelo, mas utilizar o conhecimento a respeito de variáveis aleatórias binomiais em situações-problema que podem ser modeladas probabilisticamente por meio de variáveis binomiais. Deste modo, daremos a seguir algumas características intrínsecas a variáveis aleatórias deste tipo. Falaremos sobre Função de Distribuição Acumulada, Esperança, Variância e posteriormente sobre algumas aplicações.

4.4.1 Função de Distribuição Acumulada - FDA

A FDA é uma função que associa um número real positivo a um número do intervalo [0, 1]. Da definição (p.35) $F(x) = P(X \leq x)$ onde X é uma varíavel aleatória. Em se tratando de varíaveis binomiais, percebemos rapidamente que

i) se $x < 0$, então $F(x) = 0$.

ii) se $x \geq n$, entaõ $F(x) = 1$.

Além disso, como as varíaveis binomiais recebem apenas valores no conjunto $\{0, 1, 2, ..., n\}$, o grafíco de sua FDA terá a forma de uma "*escada*" onde cada *"degrau"* começa em um número natural e terminará em seu sucessor. Também o parâmetro n (números de ensaios ou repetições) determinará o número de "*degraus*" de seu gráfico, a saber, n + 2 "*degraus*".

Para exemplificar, imagine uma variável aleatória binomial X que está associada a um experimento de Bernoulli com três repetições em que a probabilidade de sucesso em cada repetição igual a $\frac{9}{20}$. Então, estamos falando sobre um $X \backsim binomial(3; \frac{9}{20})$. Veja o gráfico de FDA de X para este caso na figura 4.2:

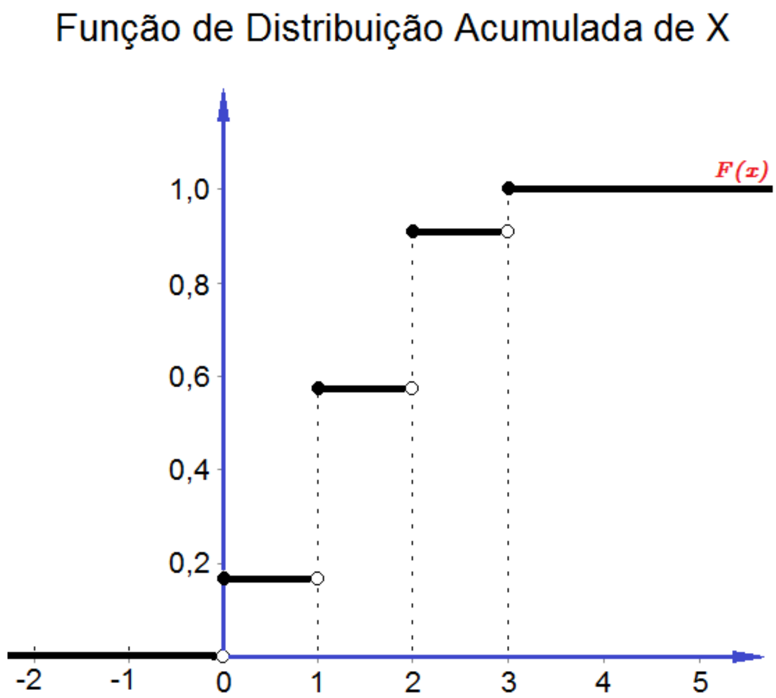

Figura 4.2.

Observe que, neste exemplo, temos $3 + 2 = 5$ "degraus"sendo que o primeiro "degrau"começa em $-\infty$ e termina em zero, o próximo começa em 0 e termina em 1 e assim acontece com os demais "degraus". Note, também, que para valores menores que zero a FDA é sempre nula ao passo que a partir de zero ela segue crescendo de "degrau"em "degrau"até chegar ao parâmetro $n = 3$ (número de repetições) onde alcança seu valor máximo: 1.

4.4.2 Valor Esperado

Também chamado de *esperança* da varíavel aleatória X o valor esperado é definido, no caso das variáveis binomiais, como:

$$E(X) = \sum_{x=0}^{n} xP(X = x). \tag{8}$$

Considere uma varíavel aleatória X binomial com parâmetros n e p. Vamos mostrar como Ross [2] nas páginas 175-176 de seu livro calcula o valor esperado de X:

$$
\begin{aligned}
E\left(X^k\right) &= \sum_{i=0}^{n} i^k \binom{n}{i} p^i (1-p)^{n-i} \\
&= 0^k \binom{n}{0} p^0 (1-p)^{n-0} + \sum_{i=1}^{n} i^k \binom{n}{i} p^i (1-p)^{n-i} \\
&= \sum_{i=1}^{n} i^{k-1} \frac{n!i}{i!(n-i)!} p^i (1-p)^{n-i} \\
&= \sum_{i=1}^{n} i^{k-1} \frac{n(n-1)!}{(i-1)!(n-i)!} p^i (1-p)^{n-i} \\
&= \sum_{i=1}^{n} i^{k-1} n \frac{(n-1)!}{(i-1)![n-1-(i-1)]!} p^i (1-p)^{n-i} \\
&= \sum_{i=1}^{n} i^{k-1} n \binom{n-1}{i-1} p^i (1-p)^{n-i} \\
&= np \sum_{i=1}^{n} i^{k-1} \binom{n-1}{i-1} p^{i-1} (1-p)^{n-i} \\
&= np \sum_{j=0}^{n-1} (j+1)^{k-1} \binom{n-1}{j} p^j (1-p)^{n-j-1} \text{ com } j = i-1
\end{aligned}
$$

$$
E\left(X^k\right) = npE\left[(Y+1)^{k-1}\right] \text{ onde } Y \backsim binomial(n-1;p). \tag{9}
$$

Para $k = 1$ na equação 9 temos, finalmente, que

$$
E(X) = np. \tag{10}
$$

O resultado da equação (10) não significa que certamente teremos np sucessos para todo experimento binomial. Na verdade, se realizarmos este experimento binomial um grande número de vezes e observarmos em cada repetição o número de sucessos, então a média aritmética do número de sucessos convergirá com probabilidade 1 para np (Lei Forte dos Grandes Números) . Também podemos calcular a esperança da variável aleatória X, por um caminho alternativo, através da função geradora de momentos (p.42) da seguinte maneira:

$$
\begin{aligned}
M_X(t) &= E\left[e^{Xt}\right] \\
&= \sum_{n=1}^{n} e^{kt} P(X=k) \\
&= \sum_{n=1}^{n} e^{kt} \binom{n}{k} p^k (1-p)^{n-k} \\
&= \sum_{n=1}^{n} \binom{n}{k} \left(pe^t\right)^k (1-p)^{n-k} \text{ ; Teorema Binomial (p.21)} \\
M_X(t) &= \left(pe^t + 1 - p\right)^n
\end{aligned}
$$

Derivando $M_X(t)$ em t, temos

$$\frac{\partial M_X}{\partial t}(t) = n\left(pe^t + 1 - p\right)^{n-1} pe^t$$

Como $E[X] = \frac{\partial M_X}{\partial t}(0)$, segue que

$$
\begin{aligned}
E[X] &= n\left(pe^0 + 1 - p\right)^{n-1} pe^0 \\
&= n(p + 1 - p)^{n-1} p \\
&= np
\end{aligned}
$$

4.4.3 Variância

Agora vamos analisar como se comporta outra propriedade importante no estudo de variáveis aleatórias: a variância. Vamos aos cálculos: do resultado (5) da página (36) deste trabalho temos,

$$
\begin{aligned}
Var(X) &= E\left[X^2\right] - (E[X])^2 \\
&= npE[Y+1] - (np)^2 \text{; recorrendo á equação 9 para } k = 2 \\
&= np(E[Y] + E[1]) - (np)^2 \\
&= np[(n-1)p + 1] - (np)^2 \\
&= np[np - p + 1] - (np)^2 \\
&= (np)^2 - np^2 + np - (np)^2 \\
&= np(1-p).
\end{aligned}
$$

Deste modo constata-se que

$$Var(X) = np(1-p). \tag{11}$$

Utilizando a mesma igualdade e nossos conhecimentos sobre FGM podemos demonstrar a variância da variável binomial X assim:

$$\frac{\partial^2 M_X}{\partial t^2}(t) = n(n-1)\left(pe^t + 1 - p\right)^{n-2}\left(pe^t\right)^2 + n\left(pe^t + 1 - p\right)^{n-1} pe^t$$

Para $t = 0$ resulta,

$$\begin{aligned}
\frac{\partial^2 M_X}{\partial t^2}(0) &= n(n-1)\left(pe^0 + 1 - p\right)^{n-2}\left(pe^0\right)^2 + n\left(pe^0 + 1 - p\right)^{n-1} pe^0 \\
&= n(n-1)(p + 1 - p)^{n-2} p^2 + n(p + 1 - p)^{n-1} p \\
&= n(n-1)p^2 + np
\end{aligned}$$

Recorrendo a $\frac{\partial^2 M_X}{\partial t^2}(0) = E[X^2]$ e aplicando o resultado em $Var(X) = E[X^2] - (E[X])^2$ temos

$$\begin{aligned}
Var(X) &= E\left[X^2\right] - (E[X])^2 \\
&= n(n-1)p^2 + np - (np)^2 \\
&= n^2p^2 - np^2 + np - n^2p^2 \\
Var(X) &= np(1-p)
\end{aligned}$$

4.5 Variáveis Aleatórias Binomiais Negativas (Pascal)

As variáveis que possuem uma distribuição de probabilidade binomial são úteis a uma grande variedade de situações práticas, no entanto, existe um outro modelo que se baseia em ensaios de Bernoulli diferente e fora do alcance das variáveis binomiais convencionais. Trata-se da situação em que temos repetições indepedentes de ensaios de Bernoulli, com mesma probabilidade p, até haver um certo número r de sucessos. Apesar de haver uma certa semelhança ao caso binomial e ambas serem discretas, veremos que as variáveis aleatórias que modelam esse experimento tem várias características diferentes das primeiras. Comecemos por demonstrar sua função de probabilidade.

É claro que agora o número de repetições não é mais fixo, como acontece em experimentos binomiais. Ao contrário, o número de sucessos r é fixo se tornando agora um dos parâmetros juntamente com p. Sendo assim, vamos definir o evento S_i : "sucesso na i-ésima repetição do ensaio de Bernoulli com probabilidade p" e a partir daí encontrar sua distribuição.

Suponha que houveram $k \geq r$ ensaios onde os $k-r$ primeiros foram fracassados e os últimos r bem sucedidos. Então o evento ocorrido foi:

$$\left(\bigcap_{i=1}^{k-r} (S_i)^c\right) \cap \left(\bigcap_{i=k-r+1}^{k} S_i\right).$$

e sua probabilidade é,

$$\begin{aligned}
P\left[\left(\bigcap_{i=1}^{k-r} (S_i)^c\right) \cap \left(\bigcap_{i=k-r+1}^{k} S_i\right)\right] &= \prod_{i=1}^{k-r} P\left[(S_i)^c\right] \prod_{i=k-r+1}^{k} P\left(S_i\right); \text{ (independência)} \\
&= \prod_{i=1}^{k-r} (1-p) \prod_{i=k-r+1}^{k} p; \text{ (mesma probabilidade)} \\
&= (1-p)^{k-r} p^{k-(k-r)} \\
P\left[\left(\bigcap_{i=1}^{k-r} (S_i)^c\right) \cap \left(\bigcap_{i=k-r+1}^{k} S_i\right)\right] &= p^r (1-p)^{k-r}.
\end{aligned}$$

Podemos dizer ainda, sob essas hipóteses, que todas as combinações com $k-r$ fracassos e r sucessos onde a última repetição é, sempre, bem sucedida possui a mesma probabilidade de ocorrência, uma vez que o produto é cumutativo. Assim, todos os eventos onde são necessários k repetições para se obter r sucessos totalizam $\binom{k-1}{k-r}$ eventos com probabilidade $p^r (1-p)^{k-r}$.

Agora, considere a variável X : "número de ensaios de Bernoulli até obter r sucessos". Logo, como todos esses eventos são disjuntos temos,

$$\begin{aligned}
P(X=k) &= \sum_{i=1}^{N} p^r (1-p)^{k-r}; \text{ onde } N = \binom{k-1}{k-r} \\
P(X=k) &= \binom{k-1}{k-r} p^r (1-p)^{k-r}; \ k = r, r+1, ... \qquad (12)
\end{aligned}$$

Agora, vamos determinar os valores da variância e esperança para $X \backsim BN(r;p)$.

$$
\begin{aligned}
E\left[X^k\right] &= \sum_{i=r}^{\infty} i^k \binom{i-1}{i-r} p^r (1-p)^{i-r}\,; \\
&= \sum_{i=r}^{\infty} i^{k-1} \left[i \binom{k-1}{k-r} \right] p^r (1-p)^{i-r}\,;\ \text{como } i\binom{i-1}{i-r} = r\binom{i}{r} \\
&= \frac{r}{p} \sum_{i=r}^{\infty} i^{k-1} \binom{k}{r} p^{r+1} (1-p)^{i-r}\,;\ \text{fazendo } j = i+1 \\
&= \frac{r}{p} \sum_{j=r+1}^{\infty} (j+1)^{k-1} \binom{k}{r} p^{r+1} (1-p)^{j-(r+1)} \\
&= \frac{r}{p} E\left[(Y-1)^{k-1} \right];\ \text{onde } Y \backsim BN(r+1; p). \qquad (13)
\end{aligned}
$$

Fazendo $k = 1$ em (13) temos

$$
E[X] = \frac{r}{p}. \qquad (14)
$$

Para a variância fazemos recorremos à equação (5, p.36) que é,

$$
\begin{aligned}
Var(X) &= E\left[X^2\right] - (E[X])^2 \\
&= \frac{r}{p} E[Y-1] - \left(\frac{r}{p}\right)^2;\ \text{como } Y \backsim BN(r+1; p) \\
&= \frac{r}{p} \left(\frac{r+1}{p} - 1 \right) - \left(\frac{r}{p}\right)^2 \\
&= \frac{r^2 + r - rp - r^2}{p^2} \\
Var(X) &= \frac{r(1-p)}{p^2}. \qquad (15)
\end{aligned}
$$

5 Proposta de aplicação

O seguinte problema é uma proposta de aplicação da teoria abordada no decurso deste trabalho, o abordaremos de forma ampla e posteriormente em uma versão adaptada para o Ensino Médio.

5.1 O problema sobre transmissão de informações

Considere uma sequência de centrais numeradas de 0 a n. Usaremos a terminologia Central i para a i-ésima central mais distante da Central 0 (dizemos que esta central é a origem). Suponha que cada central só possa emitir informação diretamente para a central mais próxima à sua direita, isto é, um sistema de transmissão de informação em série. Assim, a Central i só emite informação diretamente para a Central $i+1$. Admita que no instante inicial a Central 0 deseja emitir m informações ao longo do sistema com as seguintes hipóteses:

I) Cada informação é emitida de forma independente das demais ao longo do sistema;

II) A probabilidade da Central $i-1$ emitir uma informação sem falha para a Central i é p_i.

III) A probabilidade p_i é particular às características da conexão existente entre as centrais $i-1$ e i. Esta probabilidade não depende da informação emitida.

Teorema 5.1. *Seja a variável aleatória X_i : "número de informações que chegam corretamente à i-ésima central partindo da central de origem". Então, $X_i \backsim$ binomial $\left(m; \prod_{k=1}^{i} p_k\right)$ e sua função de probabilidade é*

$$P(X_i = x) = \binom{m}{x} \left(\prod_{k=1}^{i} p_k\right)^{x} \left(1 - \prod_{k=1}^{i} p_k\right)^{m-x} \quad ; \qquad x = 0, 1, ..., m.$$

Demonstração: De fato, para que uma informação qualquer seja transmitida partindo de C_0 (Central de origem) até C_i (Central i-ésima) devemos ter o seguinte quadro:

$$C_0 \curvearrowright C_1 \curvearrowright C_2 \curvearrowright ... \curvearrowright C_{i-1} \curvearrowright C_i.$$

Isto significa que a transmissão da informação deve ser bem sucedida em todas as centrais anteriores. Considere o evento S_k : "sucesso na transmissão de uma informação de C_{k-1} para C_k". Então, é necessário que o evento $\bigcap_{k=1}^{i} S_k$ ocorra para que sucesso na transmissão até C_i. Logo, pelo teorema da multiplicação temos que a probabilidade de sucesso na transmissão de uma informação da central C_0 para a central C_i é:

$$
\begin{aligned}
P\left(\bigcap_{k=1}^{i} S_k\right) &= P\left(S_1\right)\prod_{k=2}^{i} P\left(S_k|S_1 \cap S_2 \cap ... \cap S_{k-1}\right) \ \text{(Proposição 3.4)} \\
P\left(\bigcap_{k=1}^{i} S_k\right) &= \prod_{k=1}^{i} p_k.
\end{aligned}
$$

Por hipótese, a probabilidade de sucesso no envio de uma informação da central C_0 para a central C_i não depende do tipo de informação a ser enviada, logo percebemos que observar o sucesso ou fracasso no envio de uma informação da central C_0 para a central C_i é, isolamente, um *ensaio de Bernoulli.* Além disso, cada informação é emitida de forma independente das demais ao longo do sistema. Então, para cada informação a ser transmitida temos um *ensaio de Bernoulli* com mesma probabilidade de sucesso dada por $\prod_{k=1}^{i} p_k$, e como existem m informações a serem transmitidas são m repetições. Portanto, $X_i \backsim binomial\left(m; \prod_{k=1}^{i} p_k\right)$.

Recorrendo à página 55 deste trabalho concluímos que a função de probabilidade de X_i é:

$$
P(X_i = x) = \binom{m}{x}\left(\prod_{k=1}^{i} p_k\right)^{x}\left(1-\prod_{k=1}^{i} p_k\right)^{m-x}; \qquad x = 0, 1, ..., m. \tag{16}
$$

□

Vamos estudar como ficaria o problema e a distribuição de probabilidade quando $n \to \infty$. Usaremos o termo *Central Remota* para designar este caso. Neste caso, cabe algumas indagações tais como quais são as condições sobre p_i para que haja probabilidade não-nula de uma informação saindo da origem chegar à Central Remota?

Contudo antes de abordarmos esta questão em si propomos um lema que auxíliará no estudo do problema proposto.

Lema 5.1. *Seja* $\{a_n\}_{n\in\mathbb{N}}$ *uma sequência. Se a série* $\sum_{k=1}^{\infty} a_k \to L$, *então* $\lim_{k\to\infty} a_k = 0$ *e* $\lim_{k\to\infty} \frac{a_{k+1}}{a_k} = R \leq 1$.

Demonstração: Note que,

$$\begin{aligned}
\sum_{i=1}^{k} a_i &= \sum_{k=1}^{k-1} a_i + a_k \\
&\implies a_k = \sum_{i=1}^{k} a_i - \sum_{k=1}^{k-1} a_i \\
&\implies \lim_{k\to\infty} a_k = \lim_{k\to\infty}\left(\sum_{i=1}^{k} a_i - \sum_{k=1}^{k-1} a_i\right) \\
\lim_{k\to\infty} a_k &= \lim_{k\to\infty}\sum_{i=1}^{k} a_i - \lim_{k\to\infty}\sum_{k=1}^{k-1} a_i \\
&= L - L \\
&= 0.
\end{aligned}$$

Suponha que exista $n \in \mathbb{N}$ tal que $\forall \quad k > n \implies \frac{a_{k+1}}{a_k} > 1$. Então, $a_{k+1} > a_k > 0$. Absurdo! Pois, $\lim_{k\to\infty} a_k = 0$. Portanto, $\lim_{k\to\infty} \frac{a_{k+1}}{a_k} = R \leq 1$.

□

Proposição 5.1. *Considere a transmissão de uma informação da central de origem à Central Remota. Então, a probabilidade de transmissão da central de origem à central remota é não-nula, se e somente se,*

$$\sum_{k=1}^{\infty} (1 - p_k) < \infty. \tag{17}$$

Demonstração: Primeiramente, suponha a probabilidade de transmissão seja não nula. Isto é,

$$\begin{aligned}
0 &< \prod_{k=1}^{\infty} p_k \\
&\implies 0 < \prod_{k=1}^{\infty} [1 - (1 - p_k)] \\
&\implies \sum_{k=1}^{\infty} (1 - p_k) < \infty. \qquad \text{veja (Júnior, 2010, p.65)[4]}
\end{aligned}$$

Considere, agora, $\sum_{k=1}^{\infty}(1-p_k)$ finito. Logo, para um arbitrário $\epsilon > 0$ existe $n_0 \in \mathbb{N}$ tal que $\sum_{k=n_0}^{\infty}(1-p_k) < \epsilon$. Assim,

$$
\begin{aligned}
\prod_{k=1}^{\infty} p_k &= \prod_{k=1}^{\infty}[1-(1-p_k)] \\
&= \prod_{k=1}^{n_o-1}[1-(1-p_k)] \prod_{k=n_0}^{\infty}[1-(1-p_k)] \\
&\geq \prod_{k=1}^{n_o-1}[1-(1-p_k)] \left[1-\sum_{k=n_0}^{\infty}(1-p_k)\right] \\
&> \prod_{k=1}^{n_o-1}[1-(1-p_k)](1-\epsilon) \\
\prod_{k=1}^{\infty} p_k &> 0.
\end{aligned}
$$

□

Isto significa que $p_k = 1 - a_k$ onde $\sum_{k=1}^{\infty} a_k$ é uma série convergente. Agora, utilizando o Lema 5.1, concluímos que duas condições necessárias para que haja probabilidade de sucesso na transmissão de informações à central remota são:

i) $p_k \to 1$. Ou seja, para qualquer $\delta \in (0,1)$ deve existir $i \in \mathbb{N}$ tal que $1 \geq p_i > \delta$.Veja:

$$
\begin{aligned}
\lim_{k\to\infty} a_k &= 0 \\
&\Longrightarrow \lim_{k\to\infty}(1-p_k) = 0 \\
&\Longrightarrow \lim_{k\to\infty} 1 - \lim_{k\to\infty} p_k = 0 \\
&\Longrightarrow \lim_{k\to\infty} p_k = \lim_{k\to\infty} 1 = 1
\end{aligned}
$$

ii) Além disso, a sequência $\{p_i\}_{i\in\mathbb{N}}$ deve ser estritamente crescente. Isto é, $p_{k+1} > p_k \quad \forall \quad k > n \in \mathbb{N}$, pois

$$
\begin{aligned}
a_{k+1} &< a_k \\
&\Longrightarrow 1 - a_{k+1} > 1 - a_k \\
&\Longrightarrow p_{k+1} > p_k.
\end{aligned}
$$

Devemos ressaltar que as duas condições acima demonstradas apesar de necessárias não garantem sozinhas que haja probabilidade de sucesso na transmissão de uma informação à central remota, mas nos oferecerem características importante sobre essa probabilidade. Observe o seguinte exemplo:

Exemplo 5.1. *Considere a probilidade $p_k = 1 - \frac{1}{k}$ para o caso da central remota.*

Note que $\lim p_k = 1$ e $p_{k+1} > p_k$. Mas,

$$\begin{aligned}
\lim_{n\to\infty} \sum_{k=1}^{n} (1 - p_k) &= \lim_{n\to\infty} \sum_{k=1}^{n} \left[1 - \left(1 - \frac{1}{k}\right)\right] \\
\lim_{n\to\infty} \sum_{k=1}^{n} (1 - p_k) &= \lim_{n\to\infty} \sum_{k=1}^{n} \frac{1}{k}
\end{aligned}$$

A série $\sum_{k=1}^{\infty} \frac{1}{k}$ diverge, pois se trata da série harmônica. Então, a probabilidade de alguma informação chegar à central remota é nula.

Seguindo em frente vamos caracterizar a função de probabilidade que modela a quantidade de informações que, partindo da central de origem, chegam corretamente à Central Remota.

Proposição 5.2. *A função de probabilidade que modela a quantidade de informações que, partindo da central de origem, chegam corretamente à Central Remota é*

$$P(X = x) = \left[\binom{m}{x}\left(\lim_{i\to\infty}\prod_{k=1}^{i} p_k\right)^{x}\left(1 - \lim_{i\to\infty}\prod_{k=1}^{i} p_k\right)^{m-x}\right]; \qquad x = 0, 1, ..., m.$$

Demonstração: Seja a variável aleatória $X = \lim_{i\to\infty} X_i$. Do teorema 5.1 (p.65) temos que

$$\begin{aligned}
P(X_i &= x) = \binom{m}{x}\left(\prod_{k=1}^{i} p_k\right)^{x}\left(1 - \prod_{k=1}^{i} p_k\right)^{m-x}; \qquad x = 0, 1, ..., m \\
&\Longrightarrow \; P(X = x) = \lim_{i\to\infty}\left[\binom{m}{x}\left(\prod_{k=1}^{i} p_k\right)^{x}\left(1 - \prod_{k=1}^{i} p_k\right)^{m-x}\right]; \qquad x = 0, 1, ..., m \\
&\Longrightarrow \; P(X = x) = \left[\binom{m}{x}\left(\prod_{k=1}^{\infty} p_k\right)^{x}\left(1 - \prod_{k=1}^{\infty} p_k\right)^{m-x}\right]; \qquad x = 0, 1, ..., m.
\end{aligned}$$

Onde p_i segue os critérios da proposição 5.1 $\forall \quad i \in \mathbb{N}$.

$\square$

Exemplo 5.2. *Considere* $p_i = 1 - \frac{1}{(i+1)^2}$. *Vamos mostrar que* p_i *assim definida garante probabilidade não-nula para o caso da central remota e determinar a distribuição de probabilidade para este exemplo. Considere a probabilidade* p *de uma informação partindo da central de origem chegar à central remota. Então,*

$$
\begin{aligned}
p &= \lim_{n\to\infty}\left[\prod_{i=1}^{n} p_i\right] \\
&= \lim_{n\to\infty}\left[\prod_{i=1}^{n}\left(1 - \frac{1}{(i+1)^2}\right)\right] \\
&= \lim_{n\to\infty}\left[\prod_{i=1}^{n}\left(\frac{(i+1)^2 - 1}{(i+1)^2}\right)\right] \\
&= \lim_{n\to\infty}\left[\prod_{i=1}^{n}\left(\frac{i^2 + 2i + 1 - 1}{(i+1)^2}\right)\right] \\
&= \lim_{n\to\infty}\left[\prod_{i=1}^{n}\left(\frac{i\,(i+2)}{(i+1)^2}\right)\right]
\end{aligned}
$$

$$
\begin{aligned}
&= \lim_{n\to\infty}\left[\frac{\prod_{i=1}^{n} i \prod_{i=1}^{n}(i+2)}{\prod_{i=1}^{n}(i+1)\prod_{i=1}^{n}(i+1)}\right] \\
&= \lim_{n\to\infty}\left[\frac{n!\frac{(n+2)!}{2}}{(n+1)!\,(n+1)!}\right] \\
&= \lim_{n\to\infty}\frac{n+2}{2\,(n+1)} \\
&= \lim_{n\to\infty}\frac{1+\frac{2}{n}}{2\left(1+\frac{1}{n}\right)} \\
p &= \frac{1}{2}.
\end{aligned}
$$

Para este exemplo, onde $p = \lim_{i\to\infty}\prod_{k=1}^{i} p_k = \frac{1}{2}$ temos que a probabilidade da variável X está assim distribuída:

$$P(X = x) = \binom{m}{x}\left(\frac{1}{2}\right)^m; \qquad x = 0, 1, ..., m.$$

Uma outra questão interessante relativa a este problema é como se comporta o número de informações transmitidas entre centrais vizinhas, entendendo este termo como aquelas centrais que não possuem alguma central entre elas.

Proposição 5.3. *A Caracterização da Distribuição de Probabilidade da variável aleatória* $Y_{n-1,n}$: *"número de informações transmitidas entre a central* $n-1$ *e a central* n*"é dada por*

$$P(Y_{n-1,n} = x) = \binom{X_{n-1}}{x}(p_n)^x (1-p_n)^{X_{n-1}-x}; \qquad x = 0, 1, ..., X_{n-1}.$$

onde X_{n-1} *é o número de informações que chegam à central* $n-1$.

Demonstração: Sejam X_{n-1} :"o número de informações que chegaram a central $n-1$" e $Y_{n-1,n}$: "o número de informações que passam da central $n-1$ para a central n". Então $Y_{n-1,n}$ é uma variável binomial com parâmetros X_{n-1} e p_n. Aqui, percebemos que o número de repetições de ensaios de bernoulli é aleatório. Portanto,

$$P(Y_{n-1,n} = x) = \binom{X_{n-1}}{x}(p_n)^x (1-p_n)^{X_{n-1}-x}; \qquad x = 0, 1, ..., X_{n-1}.$$

□

Sabemos que os fenômenos aleatórios não podem ser preditos com absoluta certeza, mas podemos inferir com certo grau de certeza condições necessárias à ocorrência de certo evento. Assim, imaginemos a seguinte situação.

Exemplo 5.3. *Suponha que a central da origem deseja enviar cem informações à última central. Vamos apresentar condições sobre* p_i *para que todas essas cem informações sejam emitidas à n-ésima central com probabilidade próxima a* 95% *de um acúmulo de fracassos menor que* 10%*?*

Neste caso, podemos utilizar a uma variável binomial negativa X :"número de tentativas de emissão de informação da central de origem à n-ésima central até ocorrer 100 sucessos". Assim, sob as hipóteses do enunciado, segue que $X \backsim BN(100; p^n)$ onde $p = \sqrt[n]{\prod_{i=1}^{n} p_i}$. Recorrendo à equação 12 nosso problema se resume a:

$$\sum_{k=100}^{110} \binom{k-1}{k-100}(p^n)^r (1-p^n)^{k-r} \backsimeq 0,95.$$

Veja o gráfico de simulações da figura 5.1:

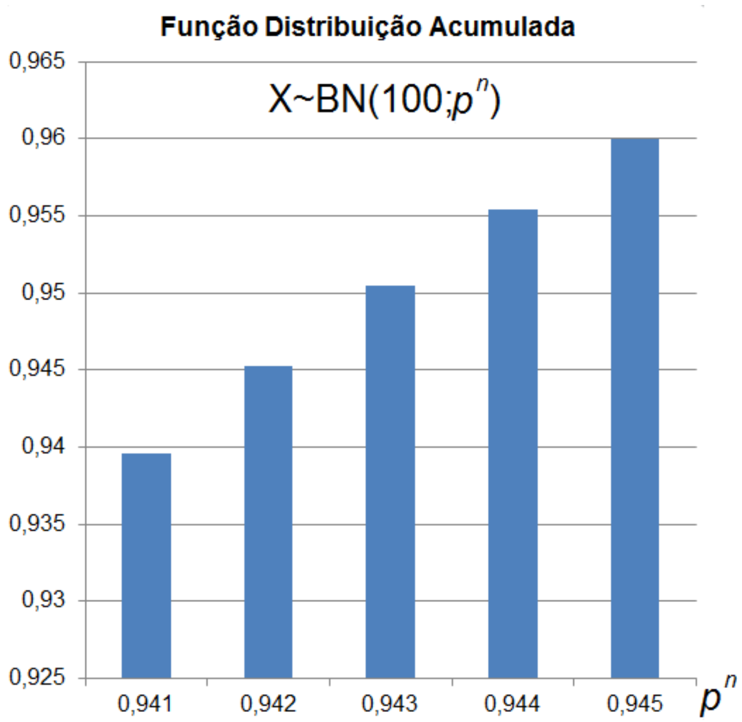

Figura 5.1.

Percebemos através da simulação que, com aproximação de três casas decimais, $p^n \backsimeq 0,943$. Considere, então,

$$\begin{aligned} p^n &\geq 0,943 \\ p &\geq \sqrt[n]{0,943}. \end{aligned}$$

A título de exemplo vamos observar o comportamento da média geométrica p em função de n quando colocado em progressão geométrica de primeiro termo 8 e razão 2.

$n =$	8	16	32	64	128
$p \geq$	0,9927	0,9963	0,9982	0,9990	0,9995

Diante da tendência observada, constatamos que uma margem segura às condições do problema é atrelar as probabilidades $\{p_i\}_{i\in\mathbb{N}}$ ao número máximo de centrais que o sistema terá. Fazendo a sequência de probabilidade tão próxima de um quanto maior seja o número de centrais do sistema. Para o caso da Central Remota, deve-se exigir os mesmos critérios expostos na Proposição 5.1.

Também há questões pertinentes ao problema relativas ao comportamento do tempo de transmissão dessas informações que podemos abordar com o auxílio da teoria da probabilidade. No entanto, como o tempo de transmissão é uma variável aleatória contínua, para isso faremos uso de resultados relativos ao caso contínuo.

Teorema 5.2. *Suponha que* $V_{n,j}$ *é o evento "sucesso na chegada da j-ésima informação à n-ésima central". Admita que o tempo de emissão da j-ésima informação entre duas centrais consecutivas - Centais* $i-1$ *e* i*, é uma variável aleatória* $T_{i,j}$ *tal que* $E\left[T_{i,j}|V_{i,j}\right] < \infty$ *e* $0 < Var\left[T_{i,j}|V_{i,j}\right] < \infty$*. Definida* $S_{n,j} = T_{1,j} + T_{2,j} + ... + T_{n,j}$ *como o tempo que uma informação* j *leva para chegar até a n-ésima central. Então,*

$$\frac{S_{n,j} \mid V_{n,j} - nE\left[T_{1,j} \mid V_{n,j}\right]}{\sqrt{nVar\left[T_{1,j} \mid V_{n,j}\right]}} \xrightarrow{d} N(0;1).$$

Demonstração: Seja $T_{i,j}$ o tempo gasto para transmissão da j-ésima informação da central $i-1$ para a central i. Então, $S_{n,j} = T_{1,j} + T_{2,j} + ... + T_{n,j}$ é o tempo até o informação j chegar à n-ésima central.

Assim, considere os eventos aleatórios $V_{n,j}$:"a informação j chega a n-ésima central"$j = 1, 2, ..., m$.

Supondo $\forall \quad i, k \in \mathbb{N} \quad T_{i,j}$ e $T_{k,j}$ são identicamente distribuídas, então, pelo Teorema Central do Limite (Ross, 2010, p.463)[2], se considerarmos a variável $\widetilde{S}_{n,j}$ cuja distribuição é idêntica à de $S_{n,j} \mid V_{n,j}$ para $n \to \infty$, a função de distribuição de probabilidade de

$$\frac{\widetilde{S}_{n,j} - E\left[\widetilde{S}_{n,j}\right]}{\sqrt{Var\left[\widetilde{S}_{n,j}\right]}} \xrightarrow{d} N(0;1). \tag{18}$$

Mas, $E\left[\widetilde{S}_{n,j}\right] = E\left[S_{n,j} \mid V_{n,j}\right] = \sum_{k=1}^{n} E\left[T_{k,j} \mid V_{n,j}\right] = nE\left[T_{1,j} \mid V_{n,j}\right] = nE\left[\widetilde{T}_{1,j}\right]$ onde $\widetilde{T}_{1,j}$ e $T_{1,j} \mid V_{n,j}$ são identicamente distribuídas. Também, $Var\left[\widetilde{S}_{n,j}\right] = Var\left[S_{n,j} \mid V_{n,j}\right] = \sum_{k=1}^{n} Var\left[T_{k,j} \mid V_{n,j}\right] = nVar\left[T_{1,j} \mid V_{n,j}\right] = nVar\left[\widetilde{T}_{1,j}\right]$.

Logo,

$$\frac{\widetilde{S}_{n,j} - nE\left[\widetilde{T}_{1,j}\right]}{\sqrt{nVar\left[\widetilde{T}_{1,j}\right]}} \xrightarrow{d} N(0;1).$$

□

Corolário 5.1. *Se* $\widetilde{T}_{1,j} \backsim \exp(\lambda)$*. Então,*

$$\frac{\widetilde{S}_{n,j} - \frac{n}{\lambda}}{\frac{1}{\lambda}\sqrt{n}} \xrightarrow{d} N(0;1).$$

Demonstração: Como $\widetilde{T}_{1,j} \backsim \exp(\lambda)$, então $E\left[\widetilde{T}_{1,j}\right] = \frac{1}{\lambda}$ e $Var\left[\widetilde{T}_{1,j}\right] = \frac{1}{\lambda^2}$. Consequentemente, $E\left[\widetilde{S}_{n,j}\right] = \frac{n}{\lambda}$ e $Var\left[\widetilde{S}_{n,j}\right] = \frac{n}{\lambda^2}$. Assim,

$$\begin{aligned}
P\left(S_{n,j} \leq s \mid V_{n,j}\right) &= P\left(\widetilde{S}_{n,j} \leq \widetilde{s}\right) \\
P\left(S_{n,j} \leq s \mid V_{n,j}\right) &= P\left(\widetilde{S}_{n,j} - \frac{n}{\lambda} \leq \widetilde{s} - \frac{n}{\lambda}\right) \\
P\left(S_{n,j} \leq s \mid V_{n,j}\right) &= P\left(\frac{\widetilde{S}_{n,j} - \frac{n}{\lambda}}{\frac{1}{\lambda}\sqrt{n}} \leq \frac{\widetilde{s} - \frac{n}{\lambda}}{\frac{1}{\lambda}\sqrt{n}}\right).
\end{aligned}$$

Agora aplicando o Teorema Central do Limite (Ross, 2010, p.463)[2] temos que,

$$P\left(\widetilde{S}_{n,j} \leq \widetilde{s}\right) \backsimeq \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{\widetilde{s}-\frac{n}{\lambda}}{\frac{1}{\lambda}\sqrt{n}}} e^{\frac{-x^2}{2}} dx.$$

Neste caso, $S_{n,j} \leq s \mid V_{n,j} \backsim gama\left(n;\lambda\right).$

$\square$

Note que o Corolário 5.1 nos dá uma aproximação da distribuição gama pela normal.

Exemplo 5.4. *Sendo o tempo medido em horas e* $\widetilde{T}_{1,j} \backsim \exp(20)$ *temos que o tempo gasto para a informações* j *chegar à centésima central satisfaz*

$$P\left(\widetilde{S}_{100,j} \leq \widetilde{s}\right) \backsimeq \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{2\widetilde{s}-10} e^{\frac{-x^2}{2}} dx.$$

Nosso trabalho neste exemplo se restinge a ajustar o limite de integração

$$\begin{aligned}
\frac{\widetilde{s} - \frac{n}{\lambda}}{\frac{1}{\lambda}\sqrt{n}} &= \frac{\frac{\lambda\widetilde{s}-n}{\lambda}}{\frac{1}{\lambda}\sqrt{n}} \\
&= \frac{\lambda\widetilde{s} - n}{\lambda} \cdot \frac{\lambda}{\sqrt{n}} \\
&= \frac{\lambda\widetilde{s} - n}{\sqrt{n}}
\end{aligned}$$

Fazendo $n = 100$ e $\lambda = 20$ temos

$$\frac{\widetilde{s} - \frac{n}{\lambda}}{\frac{1}{\lambda}\sqrt{n}} = 2\widetilde{s} - 10$$

Logo,

$$P\left(\widetilde{S}_{100,j} \leq \widetilde{s}\right) \backsimeq \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{2\widetilde{s}-10} e^{\frac{-x^2}{2}} dx.$$

Em particular,

$$P\left(\widetilde{S}_{100,j} \leq 5\right) \backsimeq \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} e^{\frac{-x^2}{2}} dx$$
$$P\left(\widetilde{S}_{100,j} \leq 5\right) \backsimeq 0,5$$

Isto significa que a informação j tem aproximadamente 50% de chance de chegar à centésima central num período de, no máximo, 5 horas.

5.1.1 Questões para o Ensino Médio

Para adaptar o problema sobre transmissão de informações ao nível médio de ensino, utilizamos critérios que visam adequar o problema e relacioná-lo ao conhecimento matemático de um aluno que cursou pelo menos o primeiro ano do Ensino Médio brasileiro. Também, decidimos alterar o nível de linguagem procurando acrescentar uma pitada de "ludismo"e bom-humor. Além de requerer do aluno respostas mais simples ao problema inicial. Pensamos então em relacionar a situação-problema inicial à brincadeira popular conhecida como "telefone-sem-fio"a fim de levar o problema em uma linguagem mais acessível ao alunado. Acrescentamos a hipótese de probabilidade constante na transmissão da informação a fim de criar uma relação entre probabilidade, equações exponenciais e progressões geométricas através do problema e, por fim, adaptamos as questões de interesse ao contexto do Ensino Médio. Os números fixados podem ser alterados livremente sem descaracterizar o problema.

Problema do telefone-sem-fio Um grupo de 12 amigos está brincando de "telefone-sem-fio". Ou seja, a primeira pessoa passa uma informação à segunda que repassa à terceira, e assim por diante. A probabilidade de sucesso na transmissão da informação de uma pessoa qualquer ao seu vizinho é constante igual a 80%. A transmissão da informação de pessoa para pessoa é independente.

Questões de interesse:

a) Qual a probabilidade da *quinta* pessoa receber a informação corretamente?

b) O jogo é considerado bem sucedido quando a informação emitida pelo primeiro jogador chega corretamente ao último. Qual a probabilidade do jogo ser bem sucedido exatamente uma vez se for repetido 5 vezes?

c) Considere agora que n pessoas estão jogando. Qual o valor de n para que a probabilidade de sucesso do jogo seja próxima a 10%? Considere $\log 8 = 0,903$.

A questão a), como o leitor pode observar, pode ser facilmente solucionada através da fórmula geral do termo de uma P.G uma vez que a transmissão da informação de pessoa para pessoa é independente. Optamos por fazer uma solução genérica.Vamos à solução:

Seja S_i : "sucesso na transmissão de uma informação do i-ésimo jogador para seu vizinho da direita"e $p_i \in \mathbb{R} \cap (0,1)$ a probabilidade de ocorrer S_i. De fato, para que uma informação qualquer seja transmitida partindo da primeiro jogador até a i-ésimo jogador temos o seguinte quadro:

$$J_1 \curvearrowright J_2 \curvearrowright ... \curvearrowright J_{i-1} \curvearrowright J_i$$

Para que a i-ésima pessoa receba corretamente a informação que partiu da primeira é necessário que a transmissão da informação seja bem sucedida entre todos os jogadores anteriores a esta pessoa. Isto significa que o evento $\bigcap\limits_{k=1}^{i} S_k$ deve ocorrer. Logo, pelo teorema da multiplicação temos que a probabilidade de sucesso na transmissão de uma informação da central C_0 para a central C_i é:

$$\begin{aligned} P\left(\bigcap_{k=1}^{i-1} S_k\right) &= P(S_1)\prod_{k=2}^{i-1} P(S_k|S_1 \cap S_2 \cap ... \cap S_{k-1}) \\ P\left(\bigcap_{k=1}^{i-1} S_k\right) &= \prod_{k=1}^{i-1} p_k. \end{aligned}$$

Note, que

$$\frac{P\left(\bigcap\limits_{k=1}^{i} S_k\right)}{P\left(\bigcap\limits_{k=1}^{i-1} S_k\right)} = p_i \qquad \forall \qquad i \in \mathbb{N}_*.$$

Acrescentando a hipótese $p_1 = p_2 = ... = p_i = p \qquad \forall \qquad i \in \mathbb{N}_*$, temos

$$\frac{P\left(\bigcap\limits_{k=1}^{i} S_k\right)}{P\left(\bigcap\limits_{k=1}^{i-1} S_k\right)} = p.$$

Então, o que temos é uma progressão geométrica de razão p. Deste modo, podemos calcular a probabilidade da i-ésima pessoa receber a informação corretamente utilizando a fórmula:

$$a_i = p^{i-1}.$$

Em particular, para arrematar a questão a) basta fazer $i = 5$ e $p = 0,8$ resultando $a_5 = 0,4096$ que é, aproximadamente, 41%.

O item b) poderia ser facilmente solucionado pela definição da variável aleatória X :"número de jogo bem sucedidos em 5 tentativas"e a posterior prova de que X é uma variável aleatória binomial com parâmetros $n = 5$ e $p = 0,8$, ou ainda, definindo X como "número de tentativas até que se obtenha o primeiro sucesso"o que faria de X uma variável binomial negativa. Para o primeiro caso, bastaria calcular $P(X = 1)$ ao passo que para $X \backsim BN(1; 0,8)$ a resposta seria $P(X = 5)$. No entanto, esses elementos não pertencem ao currículo próprio ao Ensino Médio o que nos exige um caminho alternativo. Iremos, então, utilizar conceitos e propriedades mais básicas, acessíveis ao estudante do nível médio.

Sabemos que a probabilidade do jogo entre n jogadores ser bem sucedido é p^{n-1}. Logo, será mal sucedido com probabilidade $1 - p^{n-1}$. Para k tentativas deseja-se a acertos (jogos bem sucedidos). Assim, pelo conhecimento sobre combinações simples sabemos que são $\binom{k}{a}$ maneiras de a jogos serem bem sucedidos. Mas para que isso aconteça é necessário que haja $k - a$ tentativas mal sucedidas. Logo, a probabilidade de a rodadas entre k serem bem sucedidas é $\left(p^{n-1}\right)^a \left(1 - p^{n-1}\right)^{k-a}$. Como não existe apenas uma maneira disso ocorrer devemos multiplicar essa probabilidade por $\binom{k}{a}$ o que resulta,

$$\binom{k}{a} \left(p^{n-1}\right)^a \left(1 - p^{n-1}\right)^{k-a}.$$

Agora fazendo $k = 5$, $a = 1$, $n = 12$ e $p = 0,8$ e com o auxílio de uma calculadora simples segue a resposta do item:

$$\begin{aligned}
& \binom{5}{1} \left(0,8^{12-1}\right)^1 \left(1 - 0,8^{12-1}\right)^{5-1} \\
= & \ 5\,(0,0859)\,(0,9141)^4 \\
= & \ 5 \cdot 0,0859 \cdot 0,69819 \\
= & \ 0,29985 \\
\backsimeq & \ 0,3.
\end{aligned}$$

Concluímos, através dos cálculos, que quando o jogo é repetido cinco vezes pelo grupo nas mesma condições descritas no enunciado que a probabilidade de haver uma rodada bem sucedida é de, aproximadamente, 30%.

Para responder o último item também disporemos de ferramentas matemáticas acessíveis ao Ensino Médio em coerência aos itens anteriores. Abaixo está sua solução.

Do item a) temos que a probabilidade de sucesso do jogo com n jogadores é $a_n = p^{n-1}$. É exigido que essa probabilidade seja a mais próxima possível de p_0. Assim, basta fazer

$$p^{n-1} = p_0.$$

Temos, então, uma equação exponencial para resolver e faremos aplicando a função logarítma à igualdade.

$$\begin{aligned} \log\left(p^{n-1}\right) &= \log(p_0) \\ (n-1)\log(p) &= \log(p_0) \\ n &= \frac{\log(p_0)}{\log(p)} + 1. \end{aligned}$$

Assumindo $p = 80\%$ e $p_0 = 10\%$ a resposta específica à questão é

$$\begin{aligned} \frac{\log(0,1)}{\log(0,8)} + 1 &= \frac{\log\left(\frac{1}{10}\right)}{\log\left(\frac{8}{10}\right)} + 1 \\ &= \frac{\log(1) - \log(10)}{\log(8) - \log(10)} + 1 \\ &= \frac{0-1}{0,903-1} + 1 \\ &= 10,30928 + 1 \\ &= 11,30928. \end{aligned}$$

Como n é um número natural concluímos que $n = 11$ visto que o resultado de $\frac{\log(0,1)}{\log(0,8)} + 1$ está mais próximo de 11 que de 12. Isto significa que a probabilidade do jogo será próxima a 10% quando 11 pessoas estiverem jogando.

5.1.2 Na sala de aula

O problema proposto para o ensino médio, apesar de ser uma versão bem simples sobre probabilidade aplicada a sistemas de transmissão de dados pode não ser tão trivial para o aluno do Ensino Médio tendo em vista que para apresentar uma solução correta às questões de interesse é necessário ter o domínio de conteúdos como logarítmo, progressão geométrica, análise combinatória, além dos assuntos próprios à probabilidade como sua definição e o conceito de independência entre mais de dois eventos aleatórios. Por isso, defendemos a idéia de trabalhar dez aulas preliminares antes de apresentar o problema proposto aos alunos a fim de proporcionar-lhes conteúdo básico relativo à números binomiais, definição clássica da probabilidade e independência entre eventos. Depois de revisar esses conteúdos com os alunos passaríamos ao problema proposto em si.

No período inicial da aula disporíamos os alunos em círculo explicando a brincadeira "telefone-sem-fio"e depois toda a sala brincaria. Após o momento da brincadeira é importante estimulá-los ao diálogo a respeito da brincadeira para fazermos ponderações de cunho qualitativo a respeito da matemática intrínseca ao jogo. Perguntas do tipo: "Se acrescentássemos mais uma pessoa à brincadeira o que aconteceria com a probabilidade do jogo ser bem-sucedido? E se uma pessoa sai da brincadeira, o que acontece?"Quando há poucas pessoas brincando a chance de sucesso do jogo é maior ou menor que quando há muitas?". Ao fazer essas perguntas esperaríamos a reação da turma e as suas visões do fenômeno florecer. Depois, fecharíamos o diálogo explicando que a cada pessoa que entra na roda a incerteza de transmissão da informação entre a primeira e a última pessoa aumenta. Isto porque a pessoa que entrou no jogo pode ser aquela que não irá passar a informação correta à diante e, também, porque é necessário uma transmissão a mais para finalizar o jogo. Também poderíamos levá-los a pensar na seguinte situação: imagine que alguém na roda tem problema de dicção ou audição, nesse caso a probabilidade de sucesso aumentaria ou diminuiria? Através de perguntas como essas temos a intenção de estimular o aluno a perceber aspectos qualitativos relacionados à situação. Daí em diante proporiámos o problema do telefone-sem-fio ao alunos com suas questões de interesse e deixaríamos os alunos, em duplas ou trios, construírem as soluções a partir dos conceitos já internalizados por eles. Então, no próximo encontro proporíamos uma solução nesses termos:

a) Qual a probabilidade da *quinta* pessoa receber a informação corretamente?

Para que a quinta pessoa receba a informação é necessário a informação seja transmitida corretamente 4 vezes.

$$J_1 \curvearrowright J_2 \curvearrowright J_3 \curvearrowright J_4 \curvearrowright J_5$$

Sabemos que a probabilidade de sucesso na transmissão da informação entre os jogadores é sempre de 80%. Pelo teorema da multiplicação basta multiplicar as probabilidades $(0,8)^4 \backsimeq 40\%$.

b) Qual a probabilidade do jogo ser bem sucedido exatamente uma vez se for repetido 5 vezes?

Como doze pessoas estão jogando, a probabilidade de sucesso e fracasso são, respectivamente, $(0,8)^{11}$ e $1-$ $(0,8)^{11}$. A questão exige que haja exatamente um sucesso em cinco rodadas e cada uma é independente das demais. Então basta calcular

$$\begin{aligned}
& \binom{5}{1}(0,8)^{11}\left[1-(0,8)^{11}\right]^{4} \\
= \;& 5 \cdot 0,0858993\,[1-0,0858993]^{4} \\
= \;& 0,4295\,(0,9141)^{4} \\
\backsimeq \;& 30\%.
\end{aligned}$$

c) Considere agora que n pessoas estão jogando. Qual o valor de n para que a probabilidade de sucesso do jogo seja próxima a 10%? Considere $\log 8 = 0,903$.

Neste item, é interesse deixar os alunos generalizarem a probabilidade de sucesso. Talvez nesse momento alguém perceba que se trata de uma PG de razão $\frac{8}{10}$. De todo modo, cairíamos na equação

$$\begin{aligned}
\left(\frac{8}{10}\right)^{n-1} &\backsimeq \frac{1}{10} \\
\log\left(\frac{8}{10}\right)^{n-1} &\backsimeq \log\left(\frac{1}{10}\right) \\
(n-1)\left[\log(8)-\log(10)\right] &\backsimeq \log(1)-\log(10) \\
n-1 &\backsimeq \frac{0-1}{0,903-1} \\
n &\backsimeq 11,30928 \\
n &= 11.
\end{aligned}$$

Na etapa final da aula poderíamos interpretar os resultados numéricos dizendo que a probabilidade da quinta pessoa receber a informação corretamente é aproximadamente a metade da probabilidade da segunda pessoa recebê-la, ao passo que a probabilidade do décimo primeiro jogador receber a informação correta é $\frac{1}{8}$ da probabilidade de sucesso entre dois jogadores. Com isso, ligaríamos os números às indagações da aula anterior mostrando que as respostas numéricas confirmam os aspectos qualitativos discutidos acerca da situação.

6 Considerações finais

Sem dúvida precisamos buscar recursos intelectuais para, através de situações de aprendizagem, proporcionar aos nossos alunos meios significativos para o desenvolvimento de suas competências e habilidades cognitivas, sociais e de intervenção em situações-problema. Após a publicação dos Parâmetros Curriculares Nacionais do Ensino Médio - PCNEM, se estabeleceu de maneira formal o foco da ação pegagógica no desenvolvimento das competências pessoais dos estudantes. Assim, temos caminhado baseados em eixos norteadores da ação educacional que podem ser descritos como capacidades de: I) Compreensão e expressão: Uso de diversas linguagens para a compreensão e expressão daquilo que vivenciamos das mais variadas formas que o ato de viver nos proporciona; II) Argumentação e decisão: Diz respeito à persuadir, analisar e tomar decisões complexas articulando informações, estabelecendo relações lógicas entre objetos de discussão e estudo para, então, expor pontos de vistas e tirar conclusões; III) Abstração e contextualização: O desenvolvimento de vários ramos do conhecimento se dá por meio de "idas"e "vindas"entre teoria e prática. Onde a teoria transforma a prática que reformula a teoria que novamente influi na prática num processo de mudanças de paradigmas científicos. A capacidade de abstração aponta para a consideração de situações ideiais que se comunicam com a realidade através da criatividade e imaginação do contexto em que estamos inseridos fazendo esse viés entre teoria e prática na vivência dos alunos. Deste modo, apropriar-se de conhecimentos matemáticos contextualizando-os à realidade é capacidade importante presente nos PCNEM.[7]. O que procuramos fazer através deste trabalho foi encadear o caso discreto da Teoria das Probabilidades às diretrizes nacionais para o Ensino Médio.

Percebendo a necessidade de explorar os conteúdos matemáticos e contextualizá-los aos alunos, visando o desenvolvimento das competências e exercício da cidadânia, realizamos a transposição didática, através do problema "telefone-sem-fio", de modo a contemplar essas diretrizes. Além disso, o caso mais geral apresentado pelo problema das centrais em série objeto deste trabalho nos mostra o quanto a teoria das probabilidades oferece ferramentas eficazes para diversos problemas práticos próprios ao século XXI. E que apesar de ser uma teoria relativamente nova terá um caminho promissor através das novas demandas da sociedade em fenômenos aleatórios. Deste modo, esperamos contribuir ao desenvolvimento dessa disciplina como um forte componente na formação intelectual da juventude moderna brasileira.

7 Bibliografia

Referências

[1] Leithold, L., *O cálculo com geometria analítica* / tradutor: Cyro de Carvalho Patarra. – volume 1 – 3. Ed. – São Paulo: Editora Harbra, 1994.

[2] Ross, S., *Probabilidade: um curso moderno com aplicações* / tradutor: Alberto Resende de Conti. – 8. Ed. – Porto Alegre: Bookman, 2010.

[3] Stirzader, D., *Elementary probability* - 2. Ed. – New York: Cambridge University Press, 2003.

[4] Junior, V. V. *Modelagem de epidemias via sistema de partículas inter-agentes*. Tese de doutoramento – São Paulo: USP, 2010. Disponível em <http://www.teses.usp.br/teses/disponiveis/45/45133/tde-27052013-085717/pt-br.php> acesso: 08-jan-2014.

[5] Morgado, A. C.; et al. *Análise combinatória e probabilidade.* / Augusto César Morgado, João Bosco Pitombeira de Carvalho, Paulo Cezar Pinto Carvalho, Pedro Fernandez. - 9.ed. - Rio de Janeiro: SBM, 1991.

[6] Gnedenko, B. V., *A teoria da probabilidade*, Rio de Janeiro: Editora Ciência Moderna Ltda., 2008.

[7] Ministerio da educacao e cultura. Secretaria da educacao media e tecnologica, *Parâmetros Currículares Nacionais: Ensino Médio*, (MEC/SENTEC, Brasília, 1999).

[8] James, B., *Probabilidade: um curso em nível imtermediário*, 3 ed. Rio de Janeiro: IMPA., 2004.

8 Apêndices

Teorema Central do Limite (evocado na p.72)

Sejam variáveis aleatórias $T_1, T_2, ..., T_n$ identicamente distribuídas tais que $E[T_i] < \infty$ e $o < Var[T_i] < \infty$. Então, de acordo com Ross (2010, p.463)[2], se $n \to \infty$:

$$P\left(\frac{T_1 + T_2 + ... + T_n - E[T_1 + T_2 + ... + T_n]}{\sqrt{Var[T_1 + T_2 + ... + T_n]}} \leq a\right) \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{\frac{-x^2}{2}} dx. \tag{19}$$

Proposição utilizada na página 66

Seja a sequência $\{a_k\}$ tal que $0 < a_k < 1 \quad \forall \quad k \in \mathbb{N}$. Então,

$$\prod_{k=0}^{\infty} (1 - a_k) > 0 \Longleftrightarrow \sum_{k=0}^{\infty} a_k < \infty \tag{20}$$

PLANO DE AULA 1

Autor: Dorgival Fidellis de Souza

Tempo de execução: 3 hora-aula.

Conteúdo

Números binomiais

Pré-requisitos

Definição de fatorial e princípios de análise combinatória.

Objetivos

Entender situações de aplicação de números binomiais;

Entender notação de números binomiais;

Resolver questões simples sobre o tema;

Metodologia

Aula expositiva-dialogada.

Recursos

Quadro e giz;

Procedimentos

Propor exemplo sobre princípio multiplicativo;

Explicar as situações de utilização do número binomial em contagem

Propor situações-problema.

Propor a discussão das questões de interesse entre os alunos.

Avaliação

Participação durante a aula;

Soluções apresentadas;

Referências Bibliográficas

SOUZA, Dorgival Fidellis. **Modelos Binomiais:** caracterização e aplicações. Tese de Mestrado - Goiás : UFG, 2014.

PLANO DE AULA 2

Autor: Dorgival Fidellis de Souza

Tempo de execução: 4 hora-aula.

Conteúdo

Eventos aleatórios e cálculo de suas probabilidades.

Pré-requisitos

Operações com fração.

Objetivos

Diferençiar evento certo, impossível e aleatório;

Entender definição clássica da probabilidade e exemplos;

Calcular probabilidades;

Metodologia

Aula expositiva-dialogada.

Recursos

Quadro e giz;

Procedimentos

Exemplificar e depois definir evento certo, impossível e aleatório.

Apresentar a definição clássica de probabilidade

Propor e resolver exemplos.

Propor exercícios.

Avaliação

Participação durante a aula;

Referências Bibliográficas

SOUZA, Dorgival Fidellis. **Modelos Binomiais:** caracterização e aplicações. Tese de Mestrado - Goiás : UFG, 2014.

PLANO DE AULA 3

Autor: Dorgival Fidellis de Souza

Tempo de execução: 3 hora-aula.

Conteúdo

Probabilidade - teorema da multiplicação e independência entre eventos aleatórios

Pré-requisitos

Definição de probabilidade

Objetivos

Compreender definição de probabilidade condicional;

Entender o teorema da multiplicação;

Entender a definição de independência entre dois ou mais eventos aleatórios;

Propor soluções corretas a situações-problema aplicando o teorema da multiplicação e o conceito de independência;

Metodologia

Aula expositiva-dialogada.

Recursos

Quadro e giz;

Procedimentos

Explicar e exemplificar probabilidades condicionais

Apresentar o teorema da multiplicação;

Explicar o conceito de independência entre eventos aleatórios através de exemplos;

Propor e resolver exemplos.

Propor exercícios sobre o conteúdo.

Avaliação

Participação durante a aula;

Soluções apresentadas;

Referências Bibliográficas

SOUZA, Dorgival Fidellis. **Modelos Binomiais:** caracterização e aplicações. Tese de Mestrado - Goiás : UFG, 2014.

PLANO DE AULA 4

Autor: Dorgival Fidellis de Souza

Tempo de execução: 3 hora-aula.

Conteúdo

Probabilidade

Pré-requisitos

Números binomiais, progressão geométrica, logarítmo, teorema da multiplicação e independência entre mais de dois eventos.

Objetivos

Propor soluções ao problema proposto utilizando o teorema da multiplicação;

Encontrar resposta numérica ao problema;

Perceber a ligação entre teoria da probabilidade e suas aplicações em situações práticas;

Metodologia

Aula expositiva-dialogada.

Recursos

Quadro e giz;

Procedimentos

Revisar as hipóteses de uso do teorema da multiplicação e números binomias através da resolução de questões elementares de aplicação imediata da teoria.

Colocar os alunos em círculo, explicar-lhes a brincadeira popular."telefone-sem-fio"e iniciar a brincadeira.

Fazer observações qualitativas a respeito dos conceitos que estão por detrás da brincadeira.

Apresentar o problema "Telefone-sem-fio".

Propor a discussão das questões de interesse entre os alunos.

Avaliação

Participação durante a aula;

Soluções apresentadas;

Referências Bibliográficas

SOUZA, Dorgival Fidellis. **Modelos Binomiais:** caracterização e aplicações. Tese de Mestrado - Goiás : UFG, 2014.