Sobre as Distribuições Binomial e Multinomial

Jacqueline Patrícia Duarte de Oliveira jackrioacima@yahoo.com.br
Escola Municipal Honorina Giannetti, Rio Acima, MG, Brazil

Telles Timóteo da Silva timoteo@ufsj.edu.br
Universidade Federal de São João Del-Rei, Ouro Branco, MG, Brazil

Resumo

O objetivo deste trabalho é oferecer a alunos e professores de matemática uma fonte de pesquisa que possibilite um estudo aprofundado sobre as distribuições binomial e multinomial e suas propriedades. A Distribuição Binomial costuma ser apresentada em alguns livros didáticos de ensino médio no tópico de Probabilidade, por outro lado, verifica-se que, apesar de ser extremamente interessante e estar intimamente ligada ao tema da Distribuição Binomial, a Distribuição Multinomial é pouco explorada nos livros voltados para o ensino médio. Com o presente texto, contata-se ser possível tratar de tais distribuições sem prejuízo do formalismo matemático, com o rigor necessário ao trabalho do professor de matemática, mas de forma clara e acessível a alunos com conhecimento equivalente ao do ensino médio.

Palavras-chave

Análise Combinatória, Binômio de Newton, Probabilidade, Distribuição Binomial, Distribuição Multinomial.

1 Introdução

A teoria de probabilidade é uma ferramenta matemática extremamente interessante para lidar com problemas sujeitos a incertezas. Ela é de grande utilidade pois, estando presente em vários ramos do conhecimento humano, da economia à biologia, da física à engenharia, permite avaliar, por exemplo, as chances de sucesso ou o risco de fracasso nas apostas em jogos de azar, fornece estimativas sobre o resultado de uma eleição, e auxilia nas previsões meteorológicas. Porém nota-se que tal tópico é raramente tratado no ensino básico e, quando é trabalhado, se faz de forma superficial, limitando-se à probabilidade básica, noção de espaço amostral, evento, união de eventos e probabilidade condicional. Em especial, discussões sobre a distribuição binomial aparecem em alguns livros didáticos de ensino médio, mas o conceito é pouco explorado e, em alguns deles, há a simples apresentação da fórmula e aplicação em exercícios. Já a distribuição multinomial não é tratada no ensino básico e, mesmo no ensino superior, é pouco explorada.

Este trabalho almeja explorar a ideia de distribuição binomial e multinomial e suas propriedades, tratando dos assuntos com rigor, mas ainda de forma acessível a alunos com conhecimento equivalente ao do ensino médio. Para o professor de matemática, este texto pode servir de base para uma melhor compreensão destas distribuições.

©2018 by Periódicos UFOP

Como a compreensão da teoria de probabilidade sobre espaços amostrais discretos envolve conhecimentos de análise combinatória, a Seção 2 trata brevemente deste tema. Nas Seções 2.7 e 2.7.1, discorre-se sobre aspectos gerais do Binômio de Newton e Potências de Polinômios que serão úteis no entendimento das distribuições Binomial e Multinomial. Os aspectos principais sobre Teoria de Probabilidade aparecem na Seção 3. As Seções 4 e 5 tratam do tema principal deste trabalho, e a Seção 6 introduz uma primeira noção sobre Processo Estocástico com o auxílio de propriedades da Distribuição Multinomial.

2 Breve Revisão de Análise Combinatória

É comum termos que contar objetos. Porém, na maioria das vezes, nos deparamos com situações em que enumerar os objetos se torna inviável. O uso da Análise Combinatória nos auxilia na resolução de tais situações. Com os artifícios oferecidos pela teoria da Combinatória, cálculos como a quantidade de placas de automóveis, de número de telefones possíveis, número de comissões e vários outros problemas se tornam triviais. Neste texto iremos nos limitar aos conceitos básicos da Análise Combinatória, necessários ao entendimento das distribuições Binomial e Multinomial. O leitor interessado em aprofundar-se no tema pode consultar as referências [3, 6, 7].

2.1 Princípios fundamentais de contagem

Suponha que uma pessoa queira ir de uma cidade A até uma cidade C, passando por uma cidade B, sendo que há três caminhos distintos para ir de A até B, dois caminhos distintos para ir de B até C e um caminho direto de A até C. Para cada maneira de ir de A até B temos duas maneiras de ir de B até C além de um caminho independente dos demais. Temos então $3 \cdot 2 + 1 = 7$ modos de ir de A até C.

O **Princípio Aditivo** diz que "*se temos m possibilidades de escolha para um evento A e n possibilidades de escolha para um evento B então o número total de possibilidades para escolher A ou B é $m + n$*".

O **Princípio Multiplicativo** diz que "*se temos eventos com invariância de escolha de modo que o número de possibilidades na 1ª etapa é m e o número de possibilidades na 2ª etapa é n então o número total de possibilidades de o evento ocorrer é $m \cdot n$*" [3].

2.2 Fatorial

Em grande parte dos problemas da Análise Combinatória nos deparamos com um produto de números consecutivos que decrescem até o número um. Poderíamos calcular o produto obtendo o resultado do mesmo, porém, além de se tornar um trabalho desgastante, na maioria das vezes é desnecessário. Podemos representar tal produto utilizando a definição de fatorial.

Definição 1 (Fatorial)**.** *Seja $n \in \mathbb{N}$. Definimos o* **fatorial** *de n como sendo o número $n \cdot (n-1) \cdot (n-2) \dots 3 \cdot 2 \cdot 1$, e escrevemos*

$$n! = n \cdot (n-1) \cdot (n-2) \dots 3 \cdot 2 \cdot 1.$$

Devemos observar que não podemos operar com o fatorial como números comuns. De fato:

$$n! + m! \neq (n+m)!$$
$$n! - m! \neq (n-m)!$$
$$n! \cdot m! \neq (n \cdot m)!$$
$$\frac{n!}{m!} \neq (\frac{n}{m})!.$$

2.3 Permutação simples

Quando desejamos organizar n objetos em n posições ordenadas temos uma *permutação*. Permutar é sinônimo de embaralhar, trocar objetos de posição [3].

Teorema 2.1. *Seja $C = \{c_1, c_2, \ldots, c_n\}$ um conjunto com n elementos. O número de n-uplas distintas que podemos formar sem repetir qualquer elemento de C é dado por*

$$P_n = n!. \tag{1}$$

A demonstração segue do Princípio Multiplicativo.

2.4 Permutação com repetição

Teorema 2.2. *Sejam $C = \{c_1, c_2, \ldots, c_n\}$ um conjunto com n elementos e $i_1, i_2, \ldots, i_k$ números naturais tais que*

$$\sum_{j=1}^{k} i_j = n.$$

O número de partições distintas de C em subconjuntos disjuntos de tamanhos $i_1, i_2, \ldots, i_k$ é

$$PR_n^{i_1,\ldots,i_k} = \frac{n!}{i_1! i_2! \ldots i_k!}. \tag{2}$$

A demonstração segue do Princípio Multiplicativo.

2.5 Arranjo

Teorema 2.3. *Sejam $C = \{c_1, c_2, \ldots, c_n\}$ um conjunto com n elementos e p um número natural menor do que n. O número de p-uplas distintas que podemos formar com distintos elementos de C é*

$$A_{n,p} = \frac{n!}{(n-p)!}. \tag{3}$$

Ver a demonstração em [6].

Observação 1. *Nos casos em que o número de objetos é igual ao número de posições a serem ocupadas temos uma permutação, $n! = A_{n,n}$. Percebemos então que permutação é um caso* *especial de arranjo e enseja a discussão sobre o fato de que* $0! = 1$*, pois*

$$n! = A_{n,n} = \frac{n!}{(n-n)!} = \frac{n!}{0!} \implies 0! = 1.$$

2.6 Combinação

Teorema 2.4. *Sejam* $C = \{c_1, c_2, \dots, c_n\}$ *um conjunto com* n *elementos e* p *um número natural* $p \leq n$*. O número de subconjuntos distintos de* p *elementos que podemos fazer com os elementos de* C *é*

$$C_{n,p} = \frac{n!}{(n-p)!p!}. \tag{4}$$

Demonstração: Seja $C_{n,p}$ o número de subconjuntos que podemos formar com p elementos. Para cada subconjunto de p elementos temos $p!$ permutações, logo $C_{n,p} \cdot p!$ fornece o número de arranjos de n elementos tomados p a p:

$$A_{n,p} = C_{n,p} \cdot p!$$

o que implica

$$C_{n,p} = \frac{A_{n,p}}{p!}$$

e portanto

$$C_{n,p} = \frac{n!}{(n-p)!p!}.$$

□

Observação 2. *As combinações* $C_{n,p}$ *podem ser vistas como permutações com repetição, onde, de* n *elementos, um se repete* p *vezes e o outro* $n-p$ *vezes:*

$$C_{n,p} = \frac{n!}{(n-p)!p!} = PR_n^{(n-p),p}.$$

As combinações aparecem tão freqüentemente que elas têm uma representação especial: os **números binomiais**

$$C_{n,p} = \binom{n}{p}.$$

O físico e filósofo Blaise Pascal dispôs os números da forma binomial numa tabela (Tabela 1) o que resultou em seguida na Tabela 2 denominada **Triângulo de Pascal**. Pode-se observar que os números binomiais retornam números naturais que estão associados aos coeficientes do Binômio de Newton a seguir.

Os números binomiais satisfazem a Relação de Stifel, dada na proposição a seguir:

Proposição 2.1 (Relação de Stifel)**.** *Dados os números naturais* k *e* n*, com* $k \leq n$*, então é válida*

$$
\begin{array}{ccccccccccc}
 & & & & & \binom{0}{0} & & & & & \\
 & & & & \binom{1}{0} & & \binom{1}{1} & & & & \\
 & & & \binom{2}{0} & & \binom{2}{1} & & \binom{2}{2} & & & \\
 & & \binom{3}{0} & & \binom{3}{1} & & \binom{3}{2} & & \binom{3}{3} & & \\
 & \binom{4}{0} & & \binom{4}{1} & & \binom{4}{2} & & \binom{4}{3} & & \binom{4}{4} & \\
\binom{5}{0} & & \binom{5}{1} & & \binom{5}{2} & & \binom{5}{3} & & \binom{5}{4} & & \binom{5}{5}
\end{array}
$$

Tabela 1: Números Binomiais.

linha											
0						1					
1					1		1				
2				1		2		1			
3			1		3		3		1		
4		1		4		6		4		1	
5	1		5		10		10		5		1

Tabela 2: Triângulo de Pascal.

a igualdade

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

Demonstração:

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

□

Diversos exemplos desta relação podem ser observados no Triângulo de Pascal.

2.7 Binômio de Newton e Teorema Multinomial

As potências da forma $(a+b)^n$ são chamadas de Binômios de Newton. Ao desenvolvê-las podemos observar que os coeficientes de cada monômio são dados pela n-ésima linha do Triângulo de Pascal.

$$\begin{aligned}
(x+y)^0 &= 1\\
(x+y)^1 &= x+y\\
(x+y)^2 &= x^2+2xy+y^2\\
(x+y)^3 &= x^3+3x^2y+3xy^2+y^3\\
(x+y)^4 &= x^4+4x^3y+6x^2y^2+4xy^3+y^4.
\end{aligned}$$

A expansão de um binômio da forma $(x+y)^n$ é o resultado do teorema a seguir, cuja demonstração se encontra em [4].

Teorema 2.5 (Binômio de Newton)**.** *Se x, y são dois números reais e n um número natural, o desenvolvimento do binômio $(x+y)^n$ fornece:*

$$(x+y)^n = \binom{n}{0}x^n y^0 + \binom{n}{1}x^{n-1}y + \cdots + \binom{n}{p}x^{n-p}y^p + \cdots + \binom{n}{n}x^0 y^n = \sum_{i=0}^{n}\binom{n}{i}x^{n-i}y^i. \tag{5}$$

2.7.1 Potências de polinômios - Teorema Multinomial

Podemos nos perguntar: *como seria o desenvolvimento de uma potência de um polinômio da forma*

$$(x_1+x_2+\cdots+x_k)^n \quad ?$$

Proposição 2.2 (Teorema Multinomial)**.** *Se $x_1, x_2, \ldots, x_k$ são k números reais e n é um número natural, então*

$$\begin{aligned}
&(x_1+x_2+\cdots+x_k)^n \qquad (6)\\
&= \sum_{i_1=0}^{n}\sum_{i_2=0}^{n-i_1}\cdots\sum_{i_{k-1}=0}^{n-i_1-\cdots-i_{k-2}} \frac{n!}{i_1!i_2!\ldots i_{k-1}!(n-i_1-\cdots-i_{k-1})!}x_1^{i_1}x_2^{i_2}\ldots x_{k-1}^{i_{k-1}}x_k^{n-i_1-\cdots-i_{k-1}}.
\end{aligned}$$

Demonstração: Note que o caso $k=1$ é trivial e o caso $k=2$ corresponde ao Binômio de Newton. Vamos usar indução em k, para $k \geq 3$. Comecemos desenvolvendo o polinômio $(x_1+x_2+x_3)^n$. Note que se reconhecermos os termos x_1 e (x_2+x_3), podemos utilizar o binômio de Newton para obter

$$\begin{aligned}(x_1+x_2+x_3)^n &= [x_1+(x_2+x_3)]^n \\ &= \sum_{i_1=0}^{n}\binom{n}{i_1}x_1^{i_1}(x_2+x_3)^{n-i_1}.\end{aligned}$$

Agora na última expressão entre parênteses, utilizamos novamente a expansão do binômio de Newton, obtendo

$$\begin{aligned}(x_1+x_2+x_3)^n &= \sum_{i_1=0}^{n}\binom{n}{i_1}x_1^{i_1}\sum_{i_2=0}^{n-i_1}\binom{n-i_1}{i_2}x_2^{i_2}x_3^{n-i_1-i_2} \\ &= \sum_{i_1=0}^{n}\sum_{i_2=0}^{n-i_1}\binom{n}{i_1}\binom{n-i_1}{i_2}x_1^{i_1}x_2^{i_2}x_3^{n-i_1-i_2} \\ &= \sum_{i_1=0}^{n}\sum_{i_2=0}^{n-i_1}\frac{n!}{i_1!i_2!(n-i_1-i_2)!}x_1^{i_1}x_2^{i_2}x_3^{n-i_1-i_2}.\end{aligned}$$

Agora suponha que seja válida a expansão

$$\begin{aligned}&(x_1+x_2+\cdots+x_k)^n \\ &= \sum_{i_1=0}^{n}\sum_{i_2=0}^{n-i_1}\cdots\sum_{i_{k-1}=0}^{n-i_1-\cdots-i_{k-2}}\frac{n!}{i_1!\dots i_{k-1}!(n-i_1-\cdots-i_{k-1})!}x_1^{i_1}x_2^{i_2}\dots x_{k-1}^{i_{k-1}}x_k^{n-i_1-i_2-\cdots-i_{k-1}},\end{aligned}$$

e vamos provar que

$$\begin{aligned}&(x_1+x_2+\cdots+x_{k+1})^n \\ &= \sum_{i_1=0}^{n}\sum_{i_2=0}^{n-i_1}\cdots\sum_{i_k=0}^{n-i_1-\cdots-i_{k-1}}\frac{n!}{i_1!\dots i_k!(n-i_1-\cdots-i_k)!}x_1^{i_1}x_2^{i_2}\dots x_k^{i_k}x_{k+1}^{n-i_1-\cdots-i_k}.\end{aligned}$$

Note que $(x_1+x_2+\cdots+x_{k+1})^n = [x_1+\cdots+x_{k-1}+(x_k+x_{k+1})]^n$ e podemos usar a hipótese de indução para obter:

$$
\begin{aligned}
&(x_1 + x_2 + \cdots + x_{k+1})^n \\
&= [x_1 + \cdots + x_{k-1} + (x_k + x_{k+1})]^n \\
&= \sum_{i_1=0}^{n} \cdots \sum_{i_{k-1}=0}^{n-i_1-\cdots-i_{k-2}} \frac{n!}{i_1! \ldots i_{k-1}!(n - i_1 - \cdots - i_{k-1})!} x_1^{i_1} x_2^{i_2} \ldots x_{k-1}^{i_{k-1}} (x_k + x_{k+1})^{n-i_1-i_2-\cdots-i_{k-1}} \\
&= \sum_{i_1=0}^{n} \cdots \sum_{i_{k-1}=0}^{n-i_1-\cdots-i_{k-2}} \frac{n!}{i_1! \ldots i_{k-1}!(n - i_1 - \cdots - i_{k-1})!} x_1^{i_1} x_2^{i_2} \ldots x_{k-1}^{i_{k-1}} \times \\
&\qquad \sum_{i_k=0}^{n-i_1-\cdots-i_{k-1}} \binom{n - i_1 - \cdots - i_{k-1}}{i_k} x_k^{i_k} x_{k+1}^{n-i_1-\cdots-i_{k-1}},
\end{aligned}
$$

de onde concluímos o resultado. □

Observação 3. *Note que cada coeficiente de cada monômio em* (6) *pode ser escrito como*

$$\frac{n!}{i_1! i_2! \ldots i_{k-1}! i_k!} = PR_n^{i_1,\ldots,i_k},$$

onde $i_k = n - i_1 - \cdots - i_{k-1}$. *É usual, também, a seguinte notação:*

$$\frac{n!}{i_1! i_2! \ldots i_{k-1}! i_k!} = \binom{n}{i_1, i_2, \ldots, i_k}.$$

3 Noções de Teoria de Probabilidade

Vamos apresentar algumas noções de teoria de probabilidade para que possamos prosseguir no estudo das distribuições binomial e multinomial. Algumas referências para esta seção são [1, 5, 7].

O estudo das probabilidades surgiu da necessidade de quantificar os riscos dos seguros e de avaliar as chances de ganhar em jogos de azar. Matemáticos como Pierre de Fermat e Blaise Pascal contribuíram com o desenvolvimento desta Teoria [2].

3.1 Espaço de probabilidade

Definição 2. *Um* **espaço amostral** Ω *é o conjunto de todos os resultados possíveis para um experimento.*

Observação 4. *Vale ressaltar que consideraremos aqui apenas experimentos cujos espaços amostrais são finitos.*

Definição 3. *Seja* Ω *um espaço amostral finito. Qualquer subconjunto do espaço amostral é chamado de* **evento**.

Devemos notar que um conjunto finito Ω com n elementos tem 2^n subconjuntos. A classe de subconjuntos de um conjunto Ω é conhecida como *partes de* Ω e denotada por $\mathcal{P}(\Omega)$ ou ainda 2^Ω. Assim, no decorrer deste trabalho, os possíveis eventos relacionados com um experimento serão elementos de $\mathcal{P}(\Omega)$, e reciprocamente, cada elemento de $\mathcal{P}(\Omega)$ é um evento.

Definição 4. *Sejam A e B dois eventos de um espaço amostral Ω. Os eventos A e B são disjuntos se $A \cap B = \emptyset$.*

Definição 5. *Seja Ω um um espaço amostral finito e seja $S = \mathcal{P}(\Omega)$. Uma* **medida de probabilidade** *é uma função $\mathbb{P}$ sobre S tal que*

$$\mathbb{P} : S \longrightarrow [0, 1]$$

satisfazendo os seguintes axiomas:
(i) $\mathbb{P}(\Omega) = 1$;
(ii) *Se A e B são disjuntos então $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$.*

Proposição 3.1. *Se A é um evento de Ω, então $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$.*

Demonstração: Sabemos que $A \cup A^c = \Omega$, logo

$$\mathbb{P}(A \cup A^c) = \mathbb{P}(\Omega).$$

Mas $\mathbb{P}(A \cup A^c) = \mathbb{P}(A) + \mathbb{P}(A^c)$, pois A e A^c são disjuntos. Além disso temos que $\mathbb{P}(\Omega) = 1$. Portanto

$$\mathbb{P}(A) + \mathbb{P}(A^c) = 1,$$

donde segue o resultado. □

Proposição 3.2. *Se A e B são eventos de Ω, com $A \subset B$, então $\mathbb{P}(A) \leq \mathbb{P}(B)$.*

Demonstração: Note que B é a união dos eventos disjuntos, A e $B - A$, logo, pelo axioma (ii) da Definição 5, $\mathbb{P}(B) = \mathbb{P}(A) + \mathbb{P}(B - A)$, o que implica $\mathbb{P}(B) \geq \mathbb{P}(A)$. □

Uma generalização direta do axioma 5(ii) para uma união finita de conjuntos é a seguinte

Proposição 3.3. *Sejam $A_1, A_2, \ldots, A_k$ eventos são dois a dois disjuntos de um espaço amostral Ω. Então*

$$\mathbb{P}(A_1 \cup A_2 \cup \ldots \cup A_k) = \sum_{i=1}^{k} \mathbb{P}(A_k).$$

Proposição 3.4. *Sejam $B, A_1, A_2, \ldots, A_k$ eventos de um espaço amostral Ω, onde*

$A_1, A_2, \ldots, A_k$ *são dois a dois disjuntos e tais que*

$$\Omega = \bigcup_{i=1}^{k} A_i.$$

Então

$$\mathbb{P}((B \cap A_1) \cup (B \cap A_2) \ldots \cup (B \cap A_k)) = \mathbb{P}(B).$$

Demonstração: Utilizando a propriedade distributiva da interseção em relação à união de conjuntos, temos

$$\begin{aligned}
\mathbb{P}((B \cap A_1) \cup (B \cap A_2) \ldots \cup (B \cap A_k)) &= \mathbb{P}\left(B \cap \left(\bigcup_{i=1}^{k} A_i\right)\right) \\
&= \mathbb{P}(B \cap \Omega) \\
&= \mathbb{P}(B).
\end{aligned}$$

□

Definição 6. *Definimos* **espaço de probabilidade** *como a tripla* $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ *onde* Ω *é espaço amostral,* $\mathcal{P}(\Omega)$ *é o espaço de eventos e* $\mathbb{P}$ *é a medida de probabilidade.*

Em algumas situações temos informações sobre o experimento que alteram a probabilidade de ocorrência de um evento. A esta situação damos o nome de probabilidade condicional.

Definição 7. *Seja* $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ *um espaço de probabilidade. Dados dois eventos* A *e* B *com* $\mathbb{P}(A) > 0$ *definimos a* **probabilidade condicional** *de* B *dado* A *por*

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(B \cap A)}{\mathbb{P}(A)}. \tag{7}$$

Se $\mathbb{P}(A) = 0$, é conveniente definir

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

Definição 8. *Sejam* A *e* B *eventos de* Ω*. Dizemos que* B **independe** *de* A *se*

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

Note que se $\mathbb{P}(A) = 0$, então B é independente de A.

Proposição 3.5. *Dados dois eventos* A *e* B*, se* B *independe de* A *então*

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

Demonstração: Temos, pela definição de probabilidade condicional, que

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

Como B independe de A, então

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

donde segue o resultado. □

Proposição 3.6. *Dados dois eventos A e B, se B independe de A então A independe de B.*

Demonstração: Suponha que B independa de A, então pela Proposição 3.5, $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$. Caso $\mathbb{P}(B) > 0$, temos que

$$\begin{aligned}\mathbb{P}(A|B) &= \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \\ &= \frac{\mathbb{P}(A)\mathbb{P}(B)}{\mathbb{P}(B)} \\ &= \mathbb{P}(A).\end{aligned}$$

Caso $\mathbb{P}(B) = 0$, então $\mathbb{P}(A|B) = \mathbb{P}(A)$.

Em ambos os casos, concluímos que A independe de B. □

Assim, a relação de independência entre eventos é uma relação simétrica.

3.2 Variável aleatória

Definição 9. *Seja $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ um espaço de probabilidade. Uma* **variável aleatória** *X é uma função*

$$\begin{aligned} X:\ \Omega &\longrightarrow \mathbb{R} \\ \omega\ &\longmapsto X(\omega) \end{aligned} \tag{8}$$

que satisfaz

$$X^{-1}\left(]\,a, b\,[\right) \in \mathcal{P}(\Omega) \quad \forall a, b \in \mathbb{R},\ a < b.$$

Observação 5. *Uma variável aleatória é uma característica numérica de um experimento.*

Observação 6. *Utilizamos as notações*

$$\begin{aligned}[X = x_i] &= \{\omega \in \Omega : X(\omega) = x_i\} \\ [a \leq X \leq b] &= \{\omega \in \Omega : a \leq X(\omega) \leq b\}.\end{aligned}$$

Assim temos, por exemplo,

$$X^{-1}([a,b]) = [a \leq X \leq b]$$
$$X^{-1}(\{x_i\}) = [X = x_i].$$

Observação 7. *Utilizamos a notação simplificada* $\mathbb{P}(X = x_i)$ *para representar* $\mathbb{P}([X = x_i])$.

Observação 8. *Sejam* X *e* Y *variáveis aleatórias sobre o mesmo espaço de probabilidade. A Distribuição Conjunta das variáveis* X *e* Y *é dada por*

$$\mathbb{P}(X = x_i, Y = y_i) = \mathbb{P}([X = x_i] \cap [Y = y_i]).$$

Observação 9. *Dadas duas variáveis aleatórias* X *e* Y *sobre o espaço amostral* Ω, *utilizamos a notação*

$$\mathbb{P}(X = x_i | Y = y_i)$$

para representar a probabilidade de X *assumir o valor* x_i *dado que* Y *assume o valor* y_i, *isto é*

$$\mathbb{P}(X = x_i | Y = y_i) = \mathbb{P}([X = x_i] | [Y = y_i]) = \frac{\mathbb{P}(X = x_i, Y = y_i)}{\mathbb{P}(Y = y_i)}. \tag{9}$$

Definição 10. *Seja* X *uma variável aleatória sobre um espaço amostral* Ω *finito. Definimos a* **esperança matemática** *(ou média) de* X *por*

$$E[X] = \sum_{\omega \in \Omega} X(\omega) \mathbb{P}(\{\omega\}). \tag{10}$$

Proposição 3.7. *Seja* X *uma variável aleatória e suponha que a imagem de X seja um conjunto finito,* $Im(X) = \{x_1, x_2, \ldots, x_n\}$. *A esperança matemática de* X *pode ser calculada por*

$$E[X] = \sum_{i=1}^{n} x_i \mathbb{P}(X = x_i). \tag{11}$$

Demonstração: Para cada $x_i \in Im(X)$, somando os termos $X(\omega)\mathbb{P}(\{\omega\})$ tais que $X(\omega) = x_i$ obtemos

$$\sum_{\{\omega \in \Omega : X(\omega) = x_i\}} X(\omega)\mathbb{P}(\{\omega\}) = x_i \mathbb{P}(X = x_i).$$

Somando sobre os possíveis valores x_i temos

$$\sum_{i=1}^{n} \sum_{\{\omega \in \Omega : X(\omega) = x_i\}} X(\omega)\mathbb{P}(\{\omega\}) = \sum_{i=1}^{n} x_i \mathbb{P}(X = x_i),$$

onde o primeiro membro é igual a $E[X]$. □

A esperança matemática possui a propriedade de linearidade, como atesta a proposição a seguir.

Proposição 3.8. *Se X e Y são variáveis aleatórias sobre Ω e $a \in \mathbb{R}$, então*
(i) $E[aX] = a.E[X]$,
(ii) $E[X+Y] = E[X] + E[Y]$.

Demonstração: (i)

$$\begin{aligned} E[aX] &= \sum_{\omega \in \Omega} aX(\omega)\mathbb{P}(\{\omega\}) \\ &= a \sum_{\omega \in \Omega} X(\omega)\mathbb{P}(\{\omega\}) \\ &= aE[X]. \end{aligned}$$

(ii)

$$\begin{aligned} E[X+Y] &= \sum_{\omega \in \Omega} [X(\omega) + Y(\omega)]\mathbb{P}(\{\omega\}) \\ &= \sum_{\omega \in \Omega} X(\omega)\mathbb{P}(\{\omega\}) + \sum_{\omega \in \Omega} Y(\omega)\mathbb{P}(\{\omega\}) \\ &= E[X] + E[Y]. \end{aligned}$$

□

Definição 11. *Dadas as variáveis aleatórias X_0 e X_1 sobre o espaço de probabilidade $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ tais que:*

$$\begin{aligned} X_0 : \Omega \to \mathbb{R} \qquad & Im(X_0) = \{r_1, \ldots, r_{N_0}\} \\ X_1 : \Omega \to \mathbb{R} \qquad & Im(X_1) = \{s_1, \ldots, s_{N_1}\} \end{aligned}$$

definimos **esperança condicional** *de X_1 dado X_0 como a função*

$$E[X_1|X_0] : \Omega \to \mathbb{R},$$

que satisfaz

$$E[X_1|X_0 = r_i] = \sum_{j=1}^{N_1} s_j \mathbb{P}(X_1 = s_j | X_0 = r_i). \tag{12}$$

Proposição 3.9. *Vale a igualdade*

$$E[E[X_1|X_0]] = E[X_1].$$

Demonstração: De fato, temos

$$\begin{aligned}
E[E[X_1|X_0]] &= \sum_{i=1}^{N_0} E[X_1|X_0 = r_i] \cdot \mathbb{P}(X_0 = r_i) \quad \text{[pela Definição 3.7]} \\
&= \sum_{i=1}^{N_0} \left(\sum_{j=1}^{N_1} s_j \mathbb{P}(X_1 = s_j | X_0 = r_i) \right) \cdot \mathbb{P}(X_0 = r_i) \quad \text{[pela Definição 11]} \\
&= \sum_{j=1}^{N_1} s_j \sum_{i=1}^{N_0} \mathbb{P}(X_1 = s_j | X_0 = r_i) \cdot \mathbb{P}(X_0 = r_i) \\
&= \sum_{j=1}^{N_1} s_j \sum_{i=1}^{N_0} \mathbb{P}(X_1 = s_j, X_0 = r_i) \quad \text{[pela Definição 7 e Observação 9].}
\end{aligned}$$

Note que $[X_0 = r_i]$ e $[X_0 = r_j]$ são disjuntos para $i \neq j$ e além disso,

$$\bigcup_{i=1}^{N_0} [X_0 = r_i] = \Omega,$$

portanto

$$\begin{aligned}
E[E[X_1|X_0]] &= \sum_{j=1}^{N_1} s_j \mathbb{P}(X_1 = s_j) \\
&= E[X_1].
\end{aligned}$$

□

A variância nos fornece, apesar de estar em uma escala diferente, uma indicação do quão afastado da média a distribuição dos valores está.

Definição 12. *Dada uma variável aleatória X sobre um espaço de probabilidade $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$, definimos a* **variância** *de X como*

$$Var(X) = E[X - E(X)]^2. \tag{13}$$

Proposição 3.10. *A variância da variável aleatória X pode ser calculada por*

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

Demonstração: De fato

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

□

A covariância entre duas variáveis aleatórias nos mostra a interdependência entre essas variáveis, ou seja, como elas se relacionam.

Definição 13. *Definimos a* **covariância** *entre duas variáveis X e Y, definidas no mesmo espaço de probabilidade* $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$*, como*

$$Cov(X, Y) = E[(X - E[X])(Y - E[Y])]. \tag{14}$$

Proposição 3.11. *A covariância entre X e Y pode ser calculada por*

$$Cov(X, Y) = E[XY] - E[Y]E[X].$$

Demonstração: De fato, desenvolvendo a Equação (14) acima temos

$$\begin{aligned}
Cov(X, Y) &= E[(X - E[X])(Y - E[Y])] \\
&= E[XY - XE[Y] - YE[X] + E[X]E[Y]] \\
&= E[XY] - E[XE[Y]] - E[YE[X]] + E[E[X]E[Y]] \\
&= E[XY] - E[X]E[Y] - E[Y]E[X] + E[X]E[Y] \\
&= E[XY] - E[Y]E[X].
\end{aligned}$$

□

4 Distribuição Binomial

A Distribuição Binomial é um modelo probabilístico raramente tratado nos livros didáticos de ensino médio. Quando o fazem são muito sucintos, basicamente apresentando a fórmula e aplicando-a. Como exemplo podemos citar *Matemática* de Gelson Iezzi [6], *Matemática: Conceitos & Aplicações* de Luiz Roberto Dante [3] e *Matemática: Novo Olhar* de Joamir Roberto de Souza [8]. Aqui procuraremos apresentar algumas características de variáveis aleatórias binomiais, indo além do tradicionalmente encontrado nos livros didáticos.

Segundo Barbetta [1] um experimento é binomial se:

(a) consiste de n ensaios;

(b) cada ensaio tem apenas dois resultados: sim ou não;

(c) os ensaios são independentes entre si, com probabilidade p de ocorrer sim, sendo p uma constante entre 0 e 1 $(0 < p < 1)$.

Podemos citar como exemplos: no lançamento de uma moeda só ocorre cara ou coroa, numa pesquisa onde se pergunta se as pessoas aprovam ou não um candidato as respostas possíveis são sim ou não. Tais experimentos são chamados de Experimentos Binomiais pois só há duas respostas possíveis para cada ensaio (*etapa*) do experimento.

Exemplo 1. *Uma caixa contém bolas azuis na proporção p e bolas vermelhas na proporção $(1 - p)$. Retira-se uma bola da caixa, observa-se sua cor e a repõe na caixa. Assim, sucessivamente, retira-se n bolas da caixa. Qual a probabilidade de se obter i bolas azuis?*

Observe que a cada retirada da caixa só temos duas possibilidades: bola azul ou bola vermelha. O espaço amostral é

$$\Omega = \{aa \dots a, a \dots av, \cdots , av \dots v, \cdots , vv \dots v\} ,$$

(letra a simboliza bola azul; letra v, bola vermelha). Estamos interessados no evento "sair i bolas azuis".

Então temos

- $\binom{n}{i}$ $\rightarrow$ *formas distintas de escolher i bolas azuis;*
- p $\rightarrow$*proporção das bolas azuis;*
- $(1 - p)$ $\rightarrow$*proporção de bolas vermelhas.*

Logo, a probabilidade de obtermos i bolas azuis é:

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

4.1 Distribuição binomial e variável aleatória binomial

Seja $B = \{b_1, b_2\}$ e fixe $0 < p < 1$. Defina $\Omega = B^n$.

A variável aleatória binomial conta a quantidade de vezes que um determinado evento favorável ocorre (número de sucessos).

Definição 14. *A função $X : \Omega \rightarrow \mathbb{R}$ é uma* **variável aleatória binomial** *se $X(\omega)$ for igual ao número de b_1's em ω.*

Definição 15. *Uma distribuição de probabilidade* $\mathbb{P}$ *segue uma* **distribuição binomial** *se*

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

Proposição 4.1. *A esperança da variável binomial* X *é dada por*

$$E[X] = np. \tag{15}$$

Demonstração: De acordo com a Proposição 3.7 temos $E[X] = \sum_{i=0}^{n} i.\mathbb{P}(X = i)$. Utilizando propriedades de fatorial e somatório temos

$$\begin{aligned}
E[X] &= \sum_{i=0}^{n} i \binom{n}{i} p^i (1-p)^{n-i} \\
&= \sum_{i=1}^{n} \frac{in!}{i!\,(n-i!)} p^i (1-p)^{n-i} \\
&= \sum_{i=1}^{n} \frac{in\,(n-1)!}{i\,(i-1)!\,[n-i-1+1]!} p^i (1-p)^{n-i} \\
&= \sum_{i=1}^{n} \frac{n\,(n-1)!}{(i-1)!\,[(n-1)-(i-1)]!} p^i (1-p)^{n-i} \\
&= n \sum_{i=1}^{n} \frac{(n-1)!}{(i-1)!\,[(n-1)-(i-1)]!} p^i (1-p)^{n-i} \\
&= n \sum_{i=1}^{n} \binom{n-1}{i-1} p^i (1-p)^{n-i} \\
&= n \sum_{k=0}^{n-1} \binom{n-1}{k} p^{k+1} (1-p)^{n-k-1}, \quad \text{onde } k = i-1 \\
&= np \sum_{k=0}^{n-1} \binom{n-1}{k} p^{k} (1-p)^{n-k-1} \\
&= np\,[p + (1-p)]^{n-1} \\
&= np.
\end{aligned}$$

□

Proposição 4.2. *A variância da variável* X *é dada por*

$$Var(X) = np\,(1-p). \tag{16}$$

Demonstração: Utilizando a Proposição 3.10 vamos calcular a variância através da expressão

$$Var(X) = E\left[X^2\right] - \left(E\left[X\right]\right)^2. \tag{17}$$

Primeiramente, para calcular o valor de $E[X^2]$, aplicamos propriedades de fatorial e somatório, obtendo,

$$\begin{aligned}
E\left[X^2\right] &= \sum_{i=0}^{n} i^2 \binom{n}{i} p^i \left(1-p\right)^{n-i} \\
&= \sum_{i=0}^{n} \frac{i^2 \cdot n!}{(n-i)!i!} p^i \left(1-p\right)^{n-i} \\
&= \sum_{i=1}^{n} \frac{i^2 \cdot n!}{(n-i)!i\left(i-1\right)!} p^i \left(1-p\right)^{n-i} \\
&= \sum_{i=1}^{n} \frac{i \cdot n!}{(n-i)!\left(i-1\right)!} p^i \left(1-p\right)^{n-i} \\
&= \sum_{k=0}^{n-1} \frac{(k+1) \cdot n!}{(n-k-1)!k!} p^{k+1} \left(1-p\right)^{n-k-1}, \quad \text{onde } k = i-1 \\
&= \sum_{k=0}^{n-1} \frac{k \cdot n!}{(n-k-1)!k!} p^{k+1} \left(1-p\right)^{n-k-1} + \sum_{k=0}^{n-1} \frac{n!}{(n-k-1)!k!} p^{k+1} \left(1-p\right)^{n-k-1} \\
&= np \sum_{k=0}^{n-1} \frac{k \cdot (n-1)!}{(n-k-1)!k!} p^{k} \left(1-p\right)^{(n-1)-k} + np \sum_{k=0}^{n-1} \frac{(n-1)!}{(n-k-1)!k!} p^{k} \left(1-p\right)^{(n-1)-k} \\
&= np \sum_{k=1}^{n-1} \frac{(n-1)\left(n-2\right)!}{(n-k-1)!\left(k-1\right)!} p^{k} \left(1-p\right)^{(n-1)-k} + np\left[p + \left(1-p\right)\right]^{n-1} \\
&= np\left(n-1\right) \sum_{k=1}^{n-1} \frac{(n-2)!}{\left[n-2-(k-1)\right]!\left(k-1\right)!} p.p^{k-1} \left(1-p\right)^{(n-2)-(k-1)} + np \\
&= np^2.\left(n-1\right) \sum_{k=1}^{n-1} \binom{n-2}{k-1} p^{k-1} \left(1-p\right)^{(n-2)-(k-1)} + np \\
&= n\left(n-1\right)p^2 \sum_{j=0}^{n-2} \binom{n-2}{j} p^{j} \left(1-p\right)^{n-2-j} + np, \quad \text{onde } j = k-1 \\
&= n\left(n-1\right)p^2 \left[p + \left(1-p\right)\right]^{n-2} + np \\
&= n\left(n-1\right)p^2 + np.
\end{aligned}$$

Aplicando este último resultado e a Equação (15) na Expressão (17), temos

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

□

5 Distribuição Multinomial

A distribuição multinomial não recebe tratamento na literatura do ensino básico. Porém este é um tema interessante, relevante, e que não requer conhecimentos matemáticos avançados para seu entendimento. Uma referência para esta seção é o livro de modelagem estocástica de Taylor & Karlson [9].

Uma distribuição é considerada multinomial se:
(a) consiste de n ensaios;
(b) cada ensaio tem um número discreto de resultados possíveis;
(c) os ensaios são independentes entre si, com probabilidade constante de ocorrer um determinado resultado.

Exemplo 2. *Uma caixa contém* 5 *bolas brancas,* 3 *bolas pretas e* 2 *bolas verdes. Retira-se uma bola da caixa, observa-se sua cor e a repõe na caixa. Assim, sucessivamente, retira-se* 4 *bolas da caixa. Qual a probabilidade de se obter* 2 *bolas brancas e* 1 *bola preta?*

Observe que a cada retirada da caixa temos três possibilidades: bola branca, bola preta ou bola verde.

A Tabela 7 nos mostra as configurações possíveis e a quantidade de modos de obter cada uma delas. O espaço amostral é

$$\Omega = \{bbbb, bbbp, \cdots, pppv, \cdots, vvvv\}\,.$$

Desejamos obter uma configuração da forma $bbpv$*. Podemos observar que se trata de uma permutação de quatro letras onde a letra* b *se repete, ou seja,* $PR_4^{2,1,1} = \dfrac{4!}{2!}$*. Devemos observar ainda que a proporção de bolas brancas é* $\dfrac{5}{10}$*, a proporção de bolas pretas é* $\dfrac{3}{10}$ *e a de bolas verdes é* $\dfrac{2}{10}$*. Temos, então,*

$$PR_4^{2,1,1} = \frac{4!}{2!}\left(\frac{5}{10}\right)^2\left(\frac{3}{10}\right)^1\left(\frac{2}{10}\right)^1.$$

nº de bolas brancas	configurações possíveis	modos de organizar a configuração
4	$bbbb$	1
3	$bbbv$ $bbbp$	$4!/3!$ $4!/3!$
2	$bbpp$ $bbpv$ $bbvv$	$4!/(2!2!)$ $4!/2!$ $4!/(2!2!)$
1	$bppp$ $bppv$ $bpvv$ $bvvv$	$4!/3!$ $4!/2!$ $4!/2!$ $4!/3!$
0	$pppp$ $pppv$ $ppvv$ $pvvv$ $vvvv$	$4!$ $4!/3!$ $4!/(2!2!)$ $4!/3!$ 1

Tabela 3: Configurações possíveis no Exemplo 2.

Exemplo 3. *De forma geral podemos considerar uma caixa que contém bolas brancas, bolas pretas e bolas verdes nas proporções b, p e v respectivamente. Retira-se uma bola da caixa, observa-se sua cor e a repõe na caixa. Assim, sucessivamente, retira-se n bolas da caixa. Qual a probabilidade de se obter i bolas brancas e j bolas pretas?*

Quando escolhemos as i bolas bolas brancas e j bolas pretas já estamos determinando que haverá na configuração $(n-i-j)$ bolas verdes. Analogamente ao Exemplo 2 temos

- $PR_n^{i,j,n-i-j} = \dfrac{n!}{i!j!(n-i-j)!} \longrightarrow$ *número de modos de organizar as n bolas sendo i bolas brancas, j bolas pretas e $n-i-j$ bolas verdes;*
- $b \longrightarrow$ *proporção de bolas brancas;*
- $p \longrightarrow$ *proporção de bolas pretas;*
- $v \longrightarrow$ *proporção de bolas verdes.*

Temos então

$$\mathbb{P}(i,j,n-i-j) = \frac{n!}{i!j!(n-i-j)!} \cdot b^i \cdot p^j \cdot v^{n-i-j}.$$

5.1 Distribuição multinomial e variáveis aleatórias multinomiais

Sejam $M = \{m_1, m_2, \ldots, m_t\}$ e $0 < p_1, p_2 \ldots, p_t < 1$, com

$$\sum_{j=1}^{t} p_j = 1.$$

Defina $\Omega = M^n$.

Definição 16. *Para cada $j = 1, \dots, t$, a* **variável aleatória multinomial** X_j *conta a quantidade de vezes que m_j ocorre. Isto é*

$$\begin{aligned} X_j :& \Omega \to \mathbb{R} \\ X_j(\omega) =& \text{ número de } m_j \text{ em } \omega. \end{aligned}$$

Definição 17. *Uma distribuição de probabilidade $\mathbb{P}$ segue uma* **distribuição multinomial** *se*

$$\mathbb{P}(X_1 = n_1, X_2 = n_2, \dots, X_t = n_t) = \frac{n!}{n_1! n_2! \cdots n_t!} \cdot p_1^{n_1} p_2^{n_2} \cdots p_t^{n_t}. \tag{18}$$

Proposição 5.1. *Seja $0 \leq i \leq n$. A probabilidade de se obter exatamente i elementos do tipo m_j é*

$$\binom{n}{i} p_j^i (1 - p_j)^{n-i}. \tag{19}$$

Demonstração: Vamos fazer a prova para $j = 1$, os outros casos são análogos. Seja $A \subset \Omega$ o evento em que todos os vetores ω possuem exatamente i elementos do tipo m_1. Então

$$\begin{aligned} \mathbb{P}(A) &= \sum_{i_2=0}^{n-i} \sum_{i_3=0}^{n-i-i_2} \cdots \sum_{i_{t-2}=0}^{n-i-\cdots-i_{t-3}} \sum_{i_{t-1}=0}^{n-i-\cdots-i_{t-2}} \frac{n!}{i! i_2! \dots i_{t-1}! \, (n - i - \cdots - i_{t-1})!} p_1^i \cdot p_2^{i_2} \dots p_t^{n-i-\cdots-i_{t-1}} \\ &= p_1^i \frac{n(n-1)\dots(n-i+1)}{i!} \\ &\quad \times \sum_{i_2=0}^{n-i} \sum_{i_3=0}^{n-i-i_2} \cdots \sum_{i_{t-1}=0}^{n-i-\cdots-i_{t-2}} \frac{(n-1)!}{i_2! i_3! \dots (n - i - \cdots - i_{t-1})!} p_2^{i_2} \dots p_t^{n-i-\cdots-i_{t-1}} \\ &= p_1^i \frac{n!}{(n-i)! i!} \cdot (p_2 + p_3 + \cdots + p_t)^{n-i} \\ &= p_1^i \frac{n!}{(n-i)! i!} \cdot (1 - p_1)^{n-i} \\ &= \binom{n}{i} p_1^i (1 - p_1)^{n-i}. \end{aligned}$$

□

Note que esta proposição mostra que a distribuição de probabilidade multinomial se reduz ao caso da distribuição binomial, quando se deseja observar se um elemento é de um tipo específico ou não é. Podemos dizer então que a distribuição binomial é um caso particular da distribuição multinomial.

Pela Proposição 5.1, vemos que $\mathbb{P}(X_j = i) = \binom{n}{i} p_j^i (1 - p_j)^{n-i}$, logo a esperança de X_j é dada por

$$E[X_j] = np_j. \tag{20}$$

Decorre também que a variância de X_j deve ser dada por

$$Var(X_j) = np_j(1 - p_j). \tag{21}$$

Para referência futura, também destacamos que

$$E[X_j^2] = n(n-1)p_j^2 + np_j. \tag{22}$$

Proposição 5.2. *Dadas duas variáveis aleatórias multinomiais X_u e X_v, com $u \neq v$. A covariância entre X_u e X_v é dada por*

$$Cov(X_u, X_v) = -np_up_v. \tag{23}$$

Demonstração: Pela Proposição 3.11 a covariância entre X_u e X_v pode ser calculada por

$$Cov(X_u, X_v) = E[X_uX_v] - E[X_u]E[X_v].$$

Vamos calcular o caso em que $u = 1, v = 2$; os outros casos serão análogos.

Inicialmente calculamos $E[X_1X_2]$. Note que as amostras que contribuem de forma não-nula para esta média devem conter obrigatoriamente pelo menos uma coordenada m_1 e uma coordenada m_2. Também se houver i coordenadas contendo m_1, para m_2 haverá no máximo $n - i$ possíveis coordenadas. Além disso, havendo i coordenadas m_1 e j coordenadas m_2, teremos $n - i - j$ coordenadas para os outros elementos m_k, ou seja,

$$\mathbb{P}([X_1 = i, X_2 = j]) = \frac{n!}{i!j!(n-i-j)!}p_1^ip_2^j(1 - p_1 - p_2)^{n-i-j}.$$

Em suma, tudo isto se traduz em

$$\begin{aligned} E[X_1X_2] &= \sum_{i=1}^{n-1}\sum_{j=1}^{n-i} ij\mathbb{P}([X_1 = i, X_2 = j]) \\ &= \sum_{i=1}^{n-1}\sum_{j=1}^{n-i} ij\frac{n!}{i!j!(n-i-j)!}p_1^ip_2^j(1 - p_1 - p_2)^{n-i-j}. \end{aligned}$$

Seguem os cálculos:

$$
\begin{aligned}
E[X_1X_2] &= \sum_{i=1}^{n-1}\sum_{j=1}^{n-i} ij\frac{n!}{i!j!(n-i-j)!}p_1^i p_2^j(1-p_1-p_2)^{n-i-j} \\
&= \sum_{i=1}^{n-1}\frac{i}{i!}p_1^i\sum_{j=1}^{n-i} j\frac{n!}{j!(n-i-j)!}p_2^j(1-p_1-p_2)^{n-i-j} \\
&= \sum_{i=1}^{n-1}\frac{i}{i!}p_1^i\sum_{j=1}^{n-i}\frac{n!}{(j-1)!(n-i-j)!}p_2^j(1-p_1-p_2)^{n-i-j} \\
&= \sum_{i=1}^{n-1}\frac{i}{i!}p_1^i\sum_{j=0}^{n-i-1}\frac{n!}{j!(n-i-(j+1))!}p_2^{j+1}(1-p_1-2_2)^{n-i-(j+1)} \\
&= \sum_{i=1}^{n-1}\frac{in(n-1)\dots(n-i)}{i!}p_1^i p_2\sum_{j=0}^{n-i-1}\binom{n-i-1}{j}p_2^j(1-p_1-p_2)^{n-i-1-j} \\
&= \sum_{i=1}^{n-1}\frac{in(n-1)\dots(n-i)}{i!}p_1^i p_2(1-p_1)^{n-i-1} \\
&= np_2\sum_{i=1}^{n-1}\frac{i(n-1)\dots(n-i)}{i!}p_1^i(1-p_1)^{n-i-1} \\
&= n(n-1)p_2\sum_{i=1}^{n-1}\frac{i(n-2)\dots(n-i)}{i!}p_1^i(1-p_1)^{n-1-i} \\
&= n(n-1)p_2\sum_{i=1}^{n-1}\frac{i(n-2)!}{i!(n-i-1)!}p_1^i(1-p_1)^{n-1-i} \\
&= n(n-1)p_2\sum_{i=0}^{n-2}\frac{(i+1)(n-2)!}{(i+1)!(n-1-(i+1))!}p_1^{i+1}(1-p_1)^{n-1-(i+1)} \\
&= n(n-1)p_1p_2\sum_{i=0}^{n-2}\frac{(n-2)!}{i!(n-2-i)!}p_1^i(1-p_1)^{n-2-i} \\
&= n(n-1)p_1p_2\sum_{i=0}^{n-2}\binom{n-2}{i}p_1^i(1-p_1)^{n-2-i} \\
&= n(n-1)p_1p_2(p_1+1-p_1)^{n-2} \\
&= n(n-1)p_1p_2.
\end{aligned}
$$

Como

$$Cov(X_1X_2) = E[X_iX_j] - E[X_j]E[X_i]$$

segue que

$$\begin{aligned} Cov(X_1X_2) &= n(n-1)p_1p_2 - np_1np_2 \\ &= -np_1p_2. \end{aligned}$$

□

Observe que este resultado nos indica que as variáveis aleatórias X_u e X_v variam em sentidos opostos.

6 Seqüências de variáveis aleatórias multinomiais

Nesta seção vamos introduzir uma primeira noção de processo estocástico considerando um experimento realizado em etapas; um experimento que pode, inclusive, ser repetido em sala de aula. Suponhamos que tenhamos uma caixa C com t tipos de cores de bolas, nas proporções $p_1, p_2, \ldots, p_t$, e uma segunda caixa D vazia. Vamos imaginar também que temos um reservatório R onde podemos encontrar uma quantidade muito grande de todos os t tipos de cores de bolas.

O experimento consiste nos seguintes passos:

1. Escolhemos uma bola da caixa C, observamos sua cor e retornamos a bola à caixa. Do reservatório R de bolas, retiramos uma bola da mesma cor, colocando-a na caixa D. Repetimos este processo n vezes.

2. Esvaziamos a caixa C e derramamos o conteúdo da caixa D dentro da caixa C.

3. Retornamos ao passo 1.

Após repetirmos este processo um número grande de vezes, *qual será o provável conteúdo da caixa C ao final do processo?*

Para tratar este experimento, defina $X_i(0)$ como sendo a quantidade de bolas da cor i presentes na caixa C inicialmente, e seja $X_i(\Gamma)$ a quantidade de bolas da cor i presentes na caixa C ao final do passo 2 na Γ-ésima retirada. Veja a Tabela 8. Note que, como as bolas são escolhidas *com reposição* da caixa C, as variáveis $X_i(\Gamma)$ são *multinomiais*. Acontece, porém, que o possível valor de $X_i(2)$ depende de $X_i(1)$, o valor de $X_i(3)$ depende de $X_i(2)$, de forma geral, o valor de $X_i(\Gamma)$ depende de $X_i(\Gamma-1)$. Ou seja, estas variáveis possuem uma relação *condicional*. Veja a Tabela 9, cujos resultados decorrem das Equações (20) e (21).

Podemos calcular em média o conteúdo final da caixa C, se o conteúdo inicial seguir as proporções $p_1, p_2, \ldots, p_t$. Como

$$E[X_i(\Gamma)|X_i(\Gamma-1)] = X_i(\Gamma-1),$$

Retirada	proporção na caixa C no passo 1	proporção na caixa C no final do passo 2
1	$p_1, p_2, \dots, p_t$	$\frac{X_1(1)}{n}, \frac{X_2(1)}{n}, \dots, \frac{X_t(1)}{n}$
2	$\frac{X_1(1)}{n}, \frac{X_2(1)}{n}, \dots, \frac{X_t(1)}{n}$	$\frac{X_1(2)}{n}, \frac{X_2(2)}{n}, \dots, \frac{X_t(2)}{n}$
3	$\frac{X_1(2)}{n}, \frac{X_2(2)}{n}, \dots, \frac{X_t(2)}{n}$	$\frac{X_1(3)}{n}, \frac{X_2(3)}{n}, \dots, \frac{X_t(3)}{n}$
$\vdots$	$\vdots$	$\vdots$
Γ	$\frac{X_1(\Gamma-1)}{n}, \frac{X_2(\Gamma-1)}{n}, \dots, \frac{X_t(\Gamma-1)}{n}$	$\frac{X_1(\Gamma)}{n}, \frac{X_2(\Gamma)}{n}, \dots, \frac{X_t(\Gamma)}{n}$

Tabela 4: Proporções na caixa C.

Retirada	Esperança Condicional	Variância Condicional
1	$E[X_i(1)\|X_i(0)] = n\frac{X_i(0)}{n} = np_i$	$Var(X_i(1)\|X_i(0)) = n\frac{X_i(0)}{n}\left(1-\frac{X_i(0)}{n}\right) = np_i(1-p_i)$
2	$E[X_i(2)\|X_i(1)] = n\frac{X_i(1)}{n}$	$Var(X_i(2)\|X_i(1)) = n\frac{X_i(1)}{n}\left(1-\frac{X_i(1)}{n}\right)$
$\vdots$	$\vdots$	$\vdots$
Γ	$E[X_i(\Gamma)\|X_i(\Gamma-1)] = n\frac{X_i(\Gamma-1)}{n}$	$Var(X_i(\Gamma)\|X_i(\Gamma-1)) = n\frac{X_i(\Gamma-1)}{n}\left(1-\frac{X_i(\Gamma-1)}{n}\right)$

Tabela 5: Esperança e variância condicionais das variáveis X_i.

e como, pela Proposição (3.9),

$$E[X_i(\Gamma)] = E[E[X_i(\Gamma)|X_i(\Gamma-1)]],$$

então

$$E[X_i(\Gamma)] = E[X_i(\Gamma-1)].$$

Indutivamente, é fácil ver que

$$E[X_i(\Gamma)] = E[X_i(\Gamma-1)] = \ldots = E[X_i(1)] = X_i(0) = np_i. \tag{24}$$

Isto mostra que, em média, se realizarmos o experimento várias vezes, os resultados finais para o tipo de cor de bola restante na caixa aparecerão numa proporção que replica a proporção inicial das cores das bolas.

Contudo, há mais conclusões que podemos tirar. Vamos calcular a variância de $X_i(\Gamma)$, isto é,

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

Temos que

$$\begin{aligned}
E\left[X_i(\Gamma)^2\right] &= E\big[E[X_i(\Gamma)^2|X_i(\Gamma-1)]\big] \\
&= E\left[\frac{n-1}{n}[X_i(\Gamma-1)]^2 + X_i(\Gamma-1)\right] \\
&= \frac{n-1}{n}E\left[X_i(\Gamma-1)^2\right] + E[X_i(\Gamma-1)] \\
&= \frac{n-1}{n}E\left[X_i(\Gamma-1)^2\right] + np_i,
\end{aligned}$$

onde usamos novamente a Proposição 3.9 e as Equações (22) e (24).

Por indução matemática[1], obtemos

$$E\left[X_i(\Gamma)^2\right] = \left(\frac{n-1}{n}\right)^{\Gamma}[X_i(0)]^2 + n\left[1-\left(\frac{n-1}{n}\right)^{\Gamma}\right]X_i(0).$$

Daí a variância de $X_i(\Gamma)$ será

$$\begin{aligned}
Var(X_i(\Gamma)) &= E\left[X_i(\Gamma)^2\right] - (E[X_i(\Gamma)])^2 \\
&= \left(\frac{n-1}{n}\right)^{\Gamma}[X_i(0)]^2 + n\left[1-\left(\frac{n-1}{n}\right)^{\Gamma}\right]X_i(0) - X_i(0)^2 \\
&= \left[\left(\frac{n-1}{n}\right)^{\Gamma} - 1\right]X_i(0)^2 + n\left[1-\left(\frac{n-1}{n}\right)^{\Gamma}\right]X_i(0) \\
&= X_i(0)\left[1-\left(\frac{n-1}{n}\right)^{\Gamma}\right][n - X_i(0)].
\end{aligned}$$

Podemos observar que, se realizarmos o experimento diversas vezes, quanto mais retiradas fizermos ($\Gamma \to \infty$), cada variável aleatória $X_i(\Gamma)$ terá uma variância cuja seqüência será crescente, sem ultrapassar do limite

$$X_i(0)\,[n - X_i(0)]\,.$$

Isto quer dizer que os possíveis valores da proporção da bola de cor i se dispersam cada vez mais em torno da média np_i, mas não irão se dispersar demais.

Em suma:

Após realizarmos o experimento uma vez, é provável obtermos ao final apenas bolas de uma só cor. Porém se fizermos o experimento várias e várias vezes, a cor de bola finalmente obtida irá variar de experimento para experimento, mas os resultados, em certa medida, estarão numa proporção que replica a proporção inicial das cores das bolas.

[1] Ver Apêndice.

7 Conclusão

A Distribuição Multinomial, que generaliza a Distribuição Binomial, é um tema que pode ser abordado em sala de aula de ensino médio, sendo que os conceitos matemáticos envolvidos podem ser desenvolvidos com um nível de rigor acessível aos alunos do ensino médio. No que tange o ensino superior, acreditamos que este trabalho possa se tornar uma fonte de referência para um maior aprofundamento do tema supracitado, tanto para professores de matemática quanto para alunos de graduação.

8 Agradecimentos

Agradecimento à CAPES pelo financiamento da bolsa de mestrado.

9 Apêndice

Vamos definir

$$y(\Gamma) = E\left[X_i(\Gamma)^2\right]$$

para cada $\Gamma = 0, 1, 2, \ldots$. Então

$$y(\Gamma) = \frac{(n-1)}{n} y(\Gamma - 1) + X_i(0).$$

Afirmamos que

$$y(\Gamma) = \left(\frac{n-1}{n}\right)^{\Gamma} [X_i(0)]^2 + n\left[1 - \left(\frac{n-1}{n}\right)^{\Gamma}\right] X_i(0).$$

De fato,
i) A afirmação é verdadeira para $\Gamma = 0$, pois

$$y(0) = E\left[X_i(0)^2\right] = X_i(0)^2.$$

ii) Suponha que a afirmação seja válida para Γ e vamos provar para $\Gamma + 1$. Temos que

$$\begin{aligned}
y(\Gamma+1) &= \frac{n-1}{n} y(\Gamma) + X_i(0) \\
&= \frac{n-1}{n} \left\{ \left(\frac{n-1}{n}\right)^{\Gamma} [X_i(0)]^2 + n \left[1 - \left(\frac{n-1}{n}\right)^{\Gamma}\right] X_i(0) \right\} + X_i(0) \\
&= \left(\frac{n-1}{n}\right)^{\Gamma+1} [X_i(0)]^2 + (n-1) \left[1 - \left(\frac{n-1}{n}\right)^{\Gamma}\right] X_i(0) + X_i(0) \\
&= \left(\frac{n-1}{n}\right)^{\Gamma+1} [X_i(0)]^2 + \left[n - 1 - \left(\frac{(n-1)^{\Gamma+1}}{n^{\Gamma}}\right)\right] X_i(0) + X_i(0) \\
&= \left(\frac{n-1}{n}\right)^{\Gamma+1} [X_i(0)]^2 + n \left[1 - \left(\frac{n-1}{n}\right)^{\Gamma+1}\right] X_i(0).
\end{aligned}$$

Referências

[1] Pedro Alberto Barbetta. *Estatística Aplicada às Ciências Sociais*. Editora da UFSC, São Paulo, 2005.

[2] Carl B. Boyer and Uta C. Merzbach. *História da Matemática*. Blücher, São Paulo, 2012.

[3] Luiz Roberto Dante. *Contexto & Aplicações*. Ática, São Paulo, 2014.

[4] Abramo Hefez. *Elementos de Aritmética*. SBM, Rio de Janeiro, 2011.

[5] Rodolfo Hoffmann. *Estatística para Economistas*. Editora Pioneira Thomson Learning, São Paulo, 3a. edition, 2002.

[6] Gelson Iezzi, Osvaldo Doce, David Degenszajn, and Roberto Périgo. *Matemática*. Atual Editora, São Paulo, 2005.

[7] Elon Lages Lima, Paulo César Pinto Carvalho, Eduardo Wagner, and Augusto César Morgado. *A Matemática do Ensino Médio*, volume 2. SBM, Rio de Janeiro, 6a. edition, 2006.

[8] Joamir Roberto De Souza. *Novo Olhar: matemática*. Editora FTD, São Paulo, 2014.

[9] H. M. Taylor and S. Karlson. *An Introduction to Stochastic Modeling*. Academic Press, Orlando (Florida), 1984.