Universidade Federal Fluminense

Instituto de Matemática e Estatística

Departamento de Estatística uff

Variáveis Aleatórias Discretas

Ana Maria Lima de Farias
Departamento de Estatística

Agosto 2015

Sumário

Capítulo 1

Variáveis Aleatórias

Neste capítulo, você aprenderá um conceito muito importante da teoria de probabilidade: o conceito de *variável aleatória*. Você verá que as variáveis aleatórias e suas distribuições de probabilidade são as ferramentas fundamentais na modelagem de fenômenos aleatórios. Nesse capítulo, definiremos as variáveis aleatórias discretas e contínuas, bem como funções que determinam seu comportamento probabilístico: função de probabilidade para o caso discreto e função densidade de probabildiade para o caso contínuo. Definiremos, ainda, a função de distribuição acumulada, que também caracteriza completamente as variáveis aleatórias, tanto discretas quanto contínuas.

1.1 Variável Aleatória

Consideremos o seguinte experimento aleatório: sorteio de uma amostra de 20 funcionários de uma empresa que tem 500 funcionários. O espaço amostral deste experimento é formado por todas as amostras possíveis e, como a ordem não importa e não deve haver repetição de funcionários, o número total de tais amostras é $n\Omega) = \binom{500}{20}$. Cada elemento desse espaço amostral é formado pela relação dos 20 funcionários sorteados.

Em situações como essa, em geral, o interesse não está no funcionário em si, mas, sim, em alguma característica deste funcionário, por exemplo, sua altura, se tem curso superior ou não, número de dependentes. Dessa forma, poderíamos calcular a altura média dos funcionários da amostra, o número médio de dependentes, a proporção de funcionários com curso superior, etc. Então, a cada amostra possível, ou seja, a cada ponto do espaço amostral associamos um número. Essa é a definição de *variável aleatória*.

DEFINIÇÃO Variável aleatória

Uma **variável aleatória** é uma função real (isto é, que assume valores em $\mathbb{R}$) definida no espaço amostral Ω de um experimento aleatório. Dito de outra forma, uma variável aleatória é uma função que associa um número real a cada evento de Ω.

Por questões de simplicidade, muitas vezes abreviaremos a expressão variável aleatória por v.a. A convenção usual para representar uma v.a. consiste em usar letras maiúsculas como X, Y, etc. Um valor específico, mas genérico, desta variável será representado pela letra minúscula correspondente: x, y, etc.

Continuando com o exemplo da amostra de funcionários, podemos, então, definir as seguintes variáveis aleatórias: X = "altura média em centímetros" e Y = "número máximo de dependentes". Estas variáveis têm naturezas distintas, quando levamos em conta os possíveis valores de cada uma. Para a variável X, os valores possíveis formam um intervalo, por exemplo, $[140, 200]$. Para a variável Y, os valores possíveis são números inteiros, variando de 0 a 20, por exemplo. Isso nos leva à seguinte definição.

DEFINIÇÃO Variáveis aleatórias discretas e contínuas

Uma variável aleatória é **discreta** se sua imagem (ou conjunto de valores que ela assume) for um conjunto finito ou enumerável. Se a imagem for um conjunto não enumerável, dizemos que a variável aleatória é **contínua**.

A questão que se coloca, agora, é: como atribuir probabilidade aos valores, ou intervalo de valores, de uma variável aleatória?

EXEMPLO 1.1 Dois dados

Consideremos o lançamento de dois dados equilibrados. Como já visto, o espaço amostral desse experimento é formado pelos pares ordenados (i, j) em que $i, j = 1, 2, 3, 4, 5, 6$. Esse é um experimento em que o espaço amostral não é formado por números. Suponhamos que nosso interesse esteja no máximo das faces dos dois dados. Neste caso, a v.a. X = "máximo das 2 faces" é uma variável discreta, que pode assumir os valores $1, 2, 3, 4, 5, 6$, conforme ilustrado na Tabela 1.1.

Tabela 1.1 – Variável aleatória X = "máximo das faces de 2 dados"

Pontos do espaço amostral	Valor de X
(1,1)	1
(1,2)(2,2),(2,1)	2
(1,3),(2,3),(3,3),(3,2),(3,1)	3
(1,4),(2,4),(3,4),(4,4),(4,3),(4,2),(4,1)	4
(1,5),(2,5),(3,5),(4,5),(5,5),(5,4),(5,3),(5,2),(5,1)	5
(1,6),(2,6),(3,6),(4,6),(5,6),(6,6),(6,5),(6,4),(6,3),(6,2),(6,1)	6

Podemos ver que o valor $X = 2$ corresponde ao evento $A = \{(1,2), (2,1), (2,2)\}$, enquanto o valor $X = 1$ corresponde ao evento $B = \{(1,1)\}$. Sendo assim, é de se esperar que o valor 2 seja mais provável que o valor 1, uma vez que todos os pares são equiprováveis. Podemos calcular a probabilidade de $X = 2$ usando a seguinte equivalência de eventos:

$$\{X = 2\} \equiv A = \{(1,2), (2,1), (2,2)\}$$

Dessa forma, obtemos

$$\mathrm{P}(X = 2) = \mathrm{P}(A) = \frac{3}{36}$$

De maneira análoga, obtemos

$$\begin{aligned}
\mathrm{P}(\{X = 1\}) &= \frac{1}{36} \\
\mathrm{P}(\{X = 3\}) &= \frac{5}{36} \\
\mathrm{P}(\{X = 4\}) &= \frac{7}{36} \\
\mathrm{P}(\{X = 5\}) &= \frac{9}{36} \\
\mathrm{P}(\{X = 6\}) &= \frac{11}{36}
\end{aligned}$$

Observe que conseguimos estabelecer uma probabilidade para cada valor da variável aleatória. Esse exemplo ilustra o conceito de *função de probabilidade* de uma v.a. discreta, que será apresentado mais adiante.

♦♦

EXEMPLO 1.2 Altura média de uma amostra de funcionários

Considere, agora, que retiremos várias amostras de 20 funcionários da empresa considerada anteriormente e, para cada amostra, registremos a altura média. Na Figura 1.1 temos o histograma e o polígono de frequência para essas alturas. Este histograma foi construído de forma que as áreas de cada retângulo são iguais às frequências relativas das respectivas classes. Sabemos, então, que a soma das áreas dos retângulos é 1.

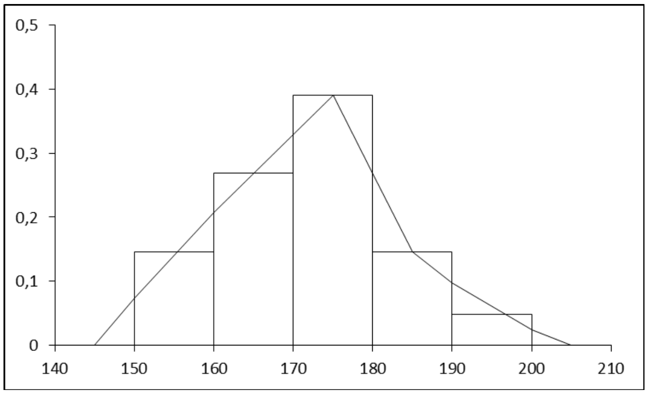

Figura 1.1 – Histograma e polígono de frequência da altura média

Tendo em mente que cada frequência relativa é uma aproximação para a probabilidade de um elemento pertencer à respectiva classe, podemos estimar a probabilidade de a altura média estar entre dois valores quaisquer como a área dos retângulos envolvidos. Veja a Figura 1.2, onde a área sombreada corresponde à frequência (probabilidade) de alturas entre os valores 168 e 178 cm. Esta área pode ser aproximada também pela área sob o polígono de frequência, conforme ilustrado na Figura 1.3. As áreas sombreadas de cinza mais escuro correspondem às diferenças abaixo e acima dopolígono de frequências; note que elas tendem a se compensar.

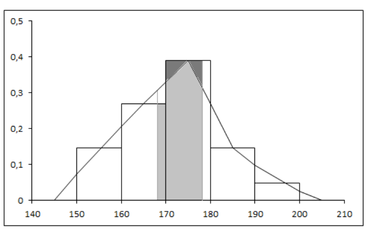

Figura 1.2 – Probabilidade como frequência relativa

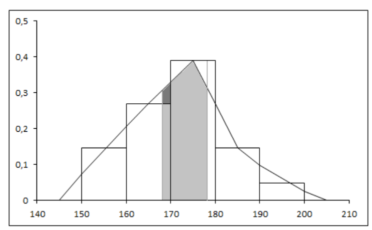

Figura 1.3 – Probabilidade como área sob o polígono de frequência

Como estamos trabalhando com uma variável aleatória contínua, faz sentido pensarmos em reduzir, cada vez mais, o comprimento de classe δ, até a situação limite em que $\delta \to 0$. Nessa situação limite, o polígono de frequências se transforma em uma curva na parte positiva (ou não-negativa) do eixo vertical, tal que a área sob ela é igual a 1. Essa curva será chamada *curva de densidade de probabilidade*.

Na situação limite, a diferença entre as áreas sombreadas mais escuro também tenderá a zero, o que nos permite concluir o seguinte: no limite, quando $\delta \to 0$, podemos estimar a probabilidade de a variável de interesse estar entre dois valores A e B pela área sob a curva de densidade de probabilidade, delimitada por esses pontos. Isso nos permitirá calcular probabilidade de intervalos de valores de qualquer variável aletatória contínua.

♦♦

Iremos, agora, apresentar as definições formais relativas às variáveis aleatórias discretas e contínuas.

1.2 Função de probabilidade

O comportamento de uma variável aleatória discreta fica perfeitamente determinado através da *função de probabilidade*.

DEFINIÇÃO Função de probabilidade

Seja X uma variável aleatória discreta. A **função de probabilidades** de X é a função $f_X(x)$ que associa, a cada valor possível x de X, sua respectiva probabilidade, calculada da seguinte forma: $f_X(x)$ é a probabilidade do evento $\{X = x\}$ que consiste em todos os resultados do espaço amostral que dão origem ao valor x.

$$f_X(x) = \mathrm{P}(\{X = x\}) = \sum_{\omega \in \Omega : X(\omega) = x} \mathrm{P}(\omega) \tag{1.1}$$

Para não sobrecarregar o texto, omitiremos os colchetes oriundos da notação de evento (conjunto) e escreveremos $\mathrm{P}(X = x)$ no lugar de $\mathrm{P}(\{X = x\})$, que seria a forma correta.

Das propriedades (axiomas) da probabilidade resultam os seguintes fatos sobre a função de probabilidades de uma v.a. discreta X:

$$f_X(x) \geq 0 \tag{1.2}$$

$$\sum_x f_X(x) = 1 \tag{1.3}$$

em que $\sum_x$ indica somatório ao longo de todos os possíveis valores de X. Note que a segunda propriedade é decorrente do axioma $\mathrm{P}(\Omega) = 1$, pois os eventos $\{X = x\}$ são mutuamente exclusivos e formam uma partição do espaço amostral. Estas são as *condições definidoras de uma função de probabilidade*.

1.3 Função densidade de probabilidade

O comportamento de uma variável aleatória contínua fica perfeitamente determinado através da *função densidade de probabilidade*.

DEFINIÇÃO Função densidade de probabilidade

Uma **função densidade de probabilidade** é uma função $f(x)$ que satisfaz as seguintes propriedades:

1. $f(x) \geq 0$
2. A área total sob o gráfico de $f(x)$ é igual a 1, isto é, $\int f(x)dx = 1$

Dada uma função $f(x)$ satisfazendo as propriedades acima, então $f(x)$ representa alguma variável aleatória contínua X, de modo que $P(a \leq X \leq b)$ é a área sob a curva limitada pelos pontos a e b (veja a Figura 1.4), isto é

$$P(a \leq X \leq b) = \int_a^b f(x)dx$$

.

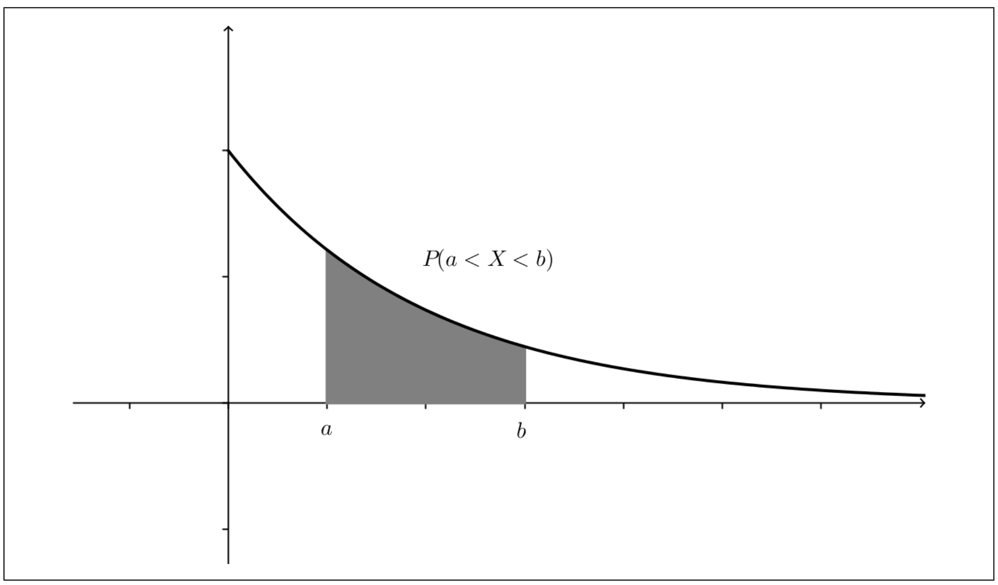

Figura 1.4 – Probabilidade como área sob a curva da função densidade de probabilidade

Uma observação importante que resulta da interpretação geométrica de probabilidade como área sob a curva de densidade de probabilidade é a seguinte: se X é uma v.a. contínua, então a probabilidade do evento $\{X = a\}$ é zero, ou seja, a probabilidade de X ser exatamente igual a um valor específico é nula. Isso pode ser visto na Figura 1.4: o evento $\{X = a\}$ corresponde a um segmento de reta, e tal segmento tem área nula. Lembre-se que $\int_a^a f(x)dx = 0$. Como consequência, são válias as seguintes igualdades:

$$\mathrm{P}(a \leq X \leq b) = \mathrm{P}(a < X \leq b) = \mathrm{P}(a \leq X < b) = \mathrm{P}(a < X < b)$$

Para deixar clara a relação entre a função densidade de probabilidade e a respectiva v.a. X, usaremos a notação $f_X(x)$.

1.4 Função de distribuição acumulada

A função de probabilidade e a função densidade de probabilidade nos dão toda a informação sobre a variável aleatória X. Existe uma outra função com tal característica (na verdade, sob determinadas condições, podemos achar outras funções com essa característica), que é a *função de distribuição acumulada* de X, cuja definição apresentamos a seguir.

DEFINIÇÃO **Função de distribuição acumulada**

Dada uma variável aleatória X, a **função de distribuição acumulada** de X, ou simplesmente **função de distribuição**, é definida por

$$F_X(x) = \mathrm{P}(X \leq x) \qquad \forall x \in \mathbb{R} \tag{1.4}$$

É interessante notar que a função F_X está definida para *todo* número real x.

Os axiomas da probabilidade e as propriedades deles decorrentes nos permitem obter as seguintes propriedades da função de distribuição de uma v.a. X.

1. Como $0 \leq \mathrm{P}(A) \leq 1$ segue que
$$0 \leq F_X(x) \leq 1 \tag{1.5}$$
2. Do axioma $\mathrm{P}(\Omega) = 1$ resulta que
$$\lim_{x \to \infty} F_X(x) = 1 \tag{1.6}$$
Note que o evento $\{X < \infty\}$ corresponde a todos os números reais e, portanto, inclui todos os valores de X.
3. Da propriedade $\mathrm{P}(\varnothing) = 0$ resulta que
$$\lim_{x \to -\infty} F_X(x) = 0 \tag{1.7}$$
Note que o evento $\{X < -\infty\}$ corresponde ao evento impossível.
4. $F_X(x)$ é uma função não decrescente, isto é, se
$$a < b \Rightarrow F_X(a) \leq F_X(b) \tag{1.8}$$
Esse resultado segue do fato de que, se $a < b$, então o evento $\{X \leq a\} \subset \{X \leq b\}$ e, portanto, $\mathrm{P}(\{X \leq a\}) \leq \mathrm{P}(\{X \leq b\})$, ou seja, $F_X(a) \leq F_X(b)$.

5. $F_X(x)$ é uma função contínua à direita, isto é

$$F_X(b) = \lim_{h \to 0} F_X(b+h) \triangleq F_X\left(b^+\right) \tag{1.9}$$

Capítulo 2

Variáveis aleatórias discretas

Nesse capítulo, vamos estudar em mais detalhes as variáveis aleatórias discretas.

2.1 Cálculo da função de probabilidade

Da definição de função de probabilidade, resulta que o seu cálculo se dá em três etapas:

- primeiro, temos que identificar todos os possíveis valores x da v.a. X;
- segundo, temos que identificar os resultados que dão origem a cada valor x e suas respectivas probabilidades;
- finalmente, temos que somar todas essas probabilidades para obter $f_X(x) = \mathrm{P}(X = x)$.

EXEMPLO 2.1 Dois dados: máximo das faces

Considerando novamente a v.a. definida na Tabela 1.1, podemos resumir a sua função de probabilidade na seguinte tabela:

x	1	2	3	4	5	6
$f_X(x)$	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	$\frac{11}{36}$

(2.1)

♦♦

EXEMPLO 2.2 Dois dados: soma das faces

Consideremos, novamente, o lançamento de dois dados, mas agora vamos definir a seguinte v.a. X = "soma das 2 faces". Para facilitar a solução deste problema, vamos construir uma tabela de duas entradas, em que cada dimensão representa o resultado de um dado e em cada cela temos a soma das duas faces.

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Como todos os pontos do espaço amostral são equiprováveis, a função de probabilidade de X é:

x	2	3	4	5	6	7	8	9	10	11	12
$f_X(x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

(2.2)

♦♦

EXEMPLO 2.3 Chaves

Um homem possui quatro chaves em seu bolso. Como está escuro, ele não consegue ver qual a chave correta para abrir a porta de sua casa, que se encontra trancada. Ele testa cada uma das chaves até encontrar a correta.

(a) Defina um espaço amostral para esse experimento.

(b) Defina a v.a. $X =$ número de chaves experimentadas até conseguir abrir a porta (inclusive a chave correta). Quais são os valores de X?

(c) Encontre a função de probabilidade de X.

Solução

(a) Vamos designar por C a chave da porta e por E_1, E_2 e E_3 as outras chaves. Se ele para de testar as chaves depois que acha a chave correta, então o espaço amostral é:

$$\Omega = \left\{ \begin{array}{l} C, \\ E_1C,\ E_2C,\ E_3C, \\ E_1E_2C,\ E_2E_1C,\ E_1E_3C,\ E_3E_1C,\ E_2E_3C,\ E_3E_2C, \\ E_1E_2E_3C,\ E_1E_3E_2C,\ E_2E_1E_3C, \\ E_2E_3E_1C,\ E_3E_1E_2C\ ,E_3E_2E_1C \end{array} \right\}$$

(b) Podemos ver, na listagem de Ω, que os possíveis valores de X são $x = 1, 2, 3, 4$

(c) Note que todas as chaves têm a mesma chance de serem sorteadas e, obviamente, cada chave testada não é colocada de volta no bolso. Feitas essas observações, podemos ver que

$$
\begin{aligned}
P(X=1) &= P(C) = \frac{1}{4} \\
P(X=2) &= P(E_1C \cup E_2C \cup E_3C) \\
&= P(E_1C) + P(E_2C) + P(E_3C) \\
&= \frac{1}{4}\times\frac{1}{3} + \frac{1}{4}\times\frac{1}{3} + \frac{1}{4}\times\frac{1}{3} = \frac{1}{4} \\
P(X=3) &= P(E_1E_2C) + P(E_2E_1C) + P(E_1E_3C) + \\
&\quad P(E_3E_1C) + P(E_2E_3C) + P(E_3E_2C) \\
&= 6\times\frac{1}{4}\times\frac{1}{3}\times\frac{1}{2} = \frac{1}{4} \\
P(X=4) &= P(E_1E_2E_3C) + P(E_1E_3E_2C) + P(E_2E_1E_3C) + \\
&\quad P(E_2E_3E_1C) + P(E_3E_1E_2C) + P(E_3E_2E_1C) \\
&= 6\times\frac{1}{4}\times\frac{1}{3}\times\frac{1}{2}\times 1 = \frac{1}{4}
\end{aligned}
$$

Logo, a função de probabilidade de X é

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

(2.3)

♦♦

EXEMPLO 2.4 Nota média de dois alunos

Dentre os cinco alunos de um curso com coeficiente de rendimento (CR) superior a 8,5, dois serão sorteados para receber uma bolsa de estudos. Os CRs desses alunos são: 8,8; 9,2; 8,9; 9,5; 9,0.

(a) Designando por A, B, C, D, E os alunos, defina um espaço amostral para esse experimento.

(b) Seja X = CR médio dos alunos sorteados. Liste os possíveis valores de X.

(c) Liste o evento $X \geq 9,0$.

(d) Encontre a função de probabilidade de X e calcule $P(X \geq 9)$.

Solução

(a) Note que aqui a ordem não importa; logo, $n(\Omega) = \binom{5}{2} = 10$. Mais especificamente,

$$
\Omega = \left\{ \begin{array}{l} (A,B),(A,C),(A,D),(A,E),(B,C), \\ (B,D),(B,E),(C,D),(C,E),(D,E) \end{array} \right\}
$$

(b) Usando uma tabela de duas entradas, podemos representar os valores de X da seguinte forma:

	$A(8,8)$	$B(9,2)$	$C(8,9)$	$D(9,5)$	$E(9,0)$
$A(8,8)$		$\frac{8,8+9,2}{2} = \mathbf{9,0}$	$8,85$	$\mathbf{9,15}$	$8,90$
$B(9,2)$			$\mathbf{9,05}$	$\mathbf{9,35}$	$\mathbf{9,10}$
$C(8,9)$				$\mathbf{9,20}$	$8,95$
$D(9,5)$					$\mathbf{9,25}$
$E(9,0)$					

(c) $\{X \geq 9\} = \{(A, B), (A, D), (B, C), (B, D), (B, E), (C, D), (D, E)\}$.

(d) Como todos os pontos do espaço amostral são equiprováveis (o sorteio é aleatório), a função de probabilidade de X é:

x	8,85	8,90	8,95	9,00	9,05	9,10	9,15	9,20	9,25	9,35
$\text{P}(X = x)$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$

e $\text{P}(X \geq 9) = \dfrac{7}{10}$.

♦♦

2.2 Função de Distribuição

Vamos calcular a função de distribuição para as variáveis aleatórias definidas nos Exemplos 2.1 a 2.3.

EXEMPLO 2.5 Dois dados: máximo das faces

Considere a função de probabilidade da v.a. X = "máximo das faces de 2 dados", dada em (2.1). Devemos notar inicialmente que nenhum valor menor que 1 é possível. Logo,

$$F_X(x) = 0 \qquad \forall x < 1 \tag{2.4}$$

Para $x = 1$, temos que

$$\begin{aligned} F_X(1) &= \text{P}(X \leq 1) = \text{P}(X < 1) + \text{P}(X = 1) \qquad (2.5) \\ &= 0 + \frac{1}{36} = \frac{1}{36} \end{aligned}$$

Para qualquer valor de x tal que $1 < x < 2$, temos $f_X(x) = 0$. Logo,

$$\begin{aligned} F_X(x) &= \text{P}(X \leq 1) + \text{P}(1 < X < x) \\ &= F_X(1) + 0 = F_X(1) \qquad \forall x : 1 < x < 2 \qquad (2.6) \end{aligned}$$

Juntando os resultados (2.5) e (2.6), obtemos

$$F_X(x) = F_X(1) = \frac{1}{36} \qquad \forall x : 1 \leq x < 2$$

Com raciocínio análogo, obtemos

$$\begin{aligned} F_X(2) &= \mathrm{P}(X \leq 2) \\ &= \mathrm{P}(X \leq 1) + \mathrm{P}(1 < X < 2) + \mathrm{P}(X = 2) \\ &= \frac{1}{36} + 0 + \frac{3}{36} = \frac{4}{36} \end{aligned} \tag{2.7}$$

e para $x \in (2, 3)$

$$F_X(x) = \mathrm{P}(X \leq 2) + \mathrm{P}(2 < X < x) = F_X(2) + 0 = F_X(2) \quad \forall x : 2 < x < 3 \tag{2.8}$$

Usando (2.7) e (2.8), obtemos

$$F_X(x) = F_X(2) = \frac{4}{36} \qquad \forall x : 2 \leq x < 3$$

Continuando, obtemos

$$\begin{aligned} F_X(x) &= F_X(3) = \frac{9}{36} & \forall x : 3 \leq x < 4 \\ F_X(x) &= F_X(4) = \frac{16}{36} & \forall x : 4 \leq x < 5 \\ F_X(x) &= F_X(5) = \frac{25}{36} & \forall x : 5 \leq x < 6 \end{aligned}$$

Para $x \geq 6$ devemos notar que o evento $\{X \leq x\}$ corresponde ao espaço amostral completo; logo

$$F_X(x) = 1 \qquad \forall\, x \geq 6$$

Dessa forma, a função de distribuição de X é

$$F_X(x) = \begin{cases} 0 & x < 1 \\ 1/36 & 1 \leq x < 2 \\ 4/36 & 2 \leq x < 3 \\ 9/36 & 3 \leq x < 4 \\ 16/36 & 4 \leq x < 5 \\ 25/36 & 5 \leq x < 6 \\ 1 & x \geq 6 \end{cases}$$

Na **Figura 2.1**, temos o gráfico de tal função em que a escala vertical está em múltiplos de $\frac{2}{36}$ e a horizontal, em múltiplos de 1. Note que esse gráfico tem a forma de uma "escada", com saltos de descontinuidade nos valores da v.a. X.

A função de probabilidade de X pode ser calculada a partir da função de distribuição da seguinte forma:

$$f_X(x) = F_X(x) - \lim_{\delta \to 0} F_X(x - \delta) \triangleq F_X(x) - F_X\left(x^-\right) \tag{2.9}$$

Isso significa que $f_X(x)$ é igual ao tamanho do "salto" da função de distribuição no ponto x.

A conclusão que podemos tirar é a seguinte: a função de probabilidades e a função de distribuição, ambas nos dão todas as informações sobre a variável aleatória X e a partir de uma podemos obter a outra, de forma inequívoca.

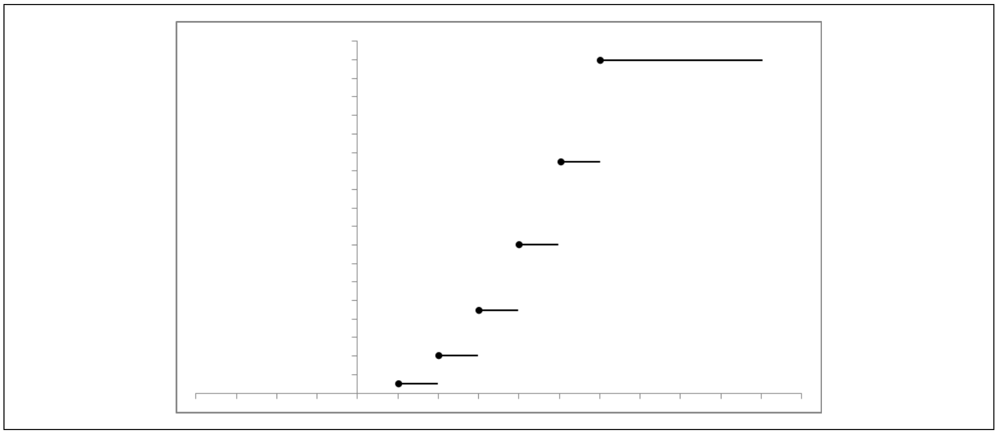

Figura 2.1 – Função de distribuição de X = "máximo das faces de 2 dados"

♦♦

EXEMPLO 2.6 Dois dados: soma das faces

Considere a função de probabilidade da v.a. X = "soma das faces de 2 dados", obtida em (2.2). Devemos notar inicialmente que nenhum valor menor que 2 é possível. Logo,

$$F_X(x) = 0 \qquad \forall x < 2 \tag{2.10}$$

Para $x = 2$, temos que

$$\begin{aligned} F_X(2) &= \mathrm{P}(X \leq 2) = \mathrm{P}(X < 2) + \mathrm{P}(X = 2) \\ &= 0 + \frac{1}{36} = \frac{1}{36} \end{aligned} \tag{2.11}$$

Para qualquer valor de x tal que $2 < x < 3$, temos $f_X(x) = 0$. Logo,

$$\begin{aligned} F_X(x) &= \mathrm{P}(X \leq 2) + \mathrm{P}(2 < X < x) \\ &= F_X(2) + 0 = F_X(2) \quad \forall x : 2 < x < 3 \end{aligned} \tag{2.12}$$

Juntando os resultados (2.11) e (2.12), obtemos

$$F_X(x) = F_X(2) = \frac{1}{36} \qquad \forall x : 2 \leq x < 3$$

Com raciocínio análogo, obtemos

$$\begin{aligned} F_X(3) &= \mathrm{P}(X \leq 3) \\ &= \mathrm{P}(X \leq 2) + \mathrm{P}(2 < X < 3) + \mathrm{P}(X = 3) \\ &= \frac{1}{36} + 0 + \frac{2}{36} = \frac{3}{36} \end{aligned} \tag{2.13}$$

e para $x \in (3, 4)$

$$F_X(x) = \mathrm{P}(X \leq 3) + \mathrm{P}(3 < X < x) = F_X(3) + 0 = F_X(3) \quad \forall x : 3 < x < 4 \tag{2.14}$$

Usando (2.13) e (2.14), obtemos

$$F_X(x) = F_X(3) = \frac{3}{36} \qquad \forall x : 3 \leq x < 4$$

Continuando, obtemos

$$\begin{aligned}
F_X(x) &= F_X(4) = \frac{6}{36} \qquad \forall x : 4 \leq x < 5 \\
F_X(x) &= F_X(5) = \frac{10}{36} \qquad \forall x : 5 \leq x < 6 \\
F_X(x) &= F_X(6) = \frac{15}{36} \qquad \forall x : 6 \leq x < 7 \\
F_X(x) &= F_X(7) = \frac{21}{36} \qquad \forall x : 7 \leq x < 8 \\
F_X(x) &= F_X(8) = \frac{26}{36} \qquad \forall x : 8 \leq x < 9 \\
F_X(x) &= F_X(9) = \frac{30}{36} \qquad \forall x : 9 \leq x < 10 \\
F_X(x) &= F_X(10) = \frac{33}{36} \qquad \forall x : 10 \leq x < 11 \\
F_X(x) &= F_X(11) = \frac{35}{36} \qquad \forall x : 11 \leq x < 12
\end{aligned}$$

Para $x \geq 12$ devemos notar que o evento $\{X \leq x\}$ corresponde ao espaço amostral completo; logo

$$F_X(x) = 1 \qquad \forall\, x \geq 12$$

Dessa forma, a função de distribuição de X é

$$F_X(x) = \begin{cases}
0 & x < 2 \\
1/36 & 2 \leq x < 3 \\
3/36 & 3 \leq x < 4 \\
6/36 & 4 \leq x < 5 \\
10/36 & 5 \leq x6 \\
15/36 & 6 \leq x < 7 \\
21/36 & 7 \leq x < 8 \\
26/36 & 8 \leq x < 9 \\
30/36 & 9 \leq x < 10 \\
33/36 & 10 \leq x < 11 \\
35/36 & 11 \leq x < 12 \\
1 & x \geq 12
\end{cases}$$

Os pontos de descontinuidade são $2, 3, \ldots, 12$, que correspondem aos valores de X. Como antes, podemos obter a função de probabilidade de X em cada um desses pontos pelo "tamanho" do salto. Por exemplo,

$$\mathrm{P}(X = 7) = \mathrm{P}(X \leq 7) - \mathrm{P}(X < 7) = \frac{21}{36} - \frac{15}{36} = \frac{6}{36}$$

♦♦

EXEMPLO 2.7 Chaves

Considere a função de probabilidade da v.a. X = "Chaves de uma porta", dada em (2.3). Seguindo raciocícnio análogo ao adotado nos dois exemplos anteriores, obtemos a seguinte função de distribuição acumulada de X="número de chaves testadas até abrir a porta":

$$F_X(x) = \begin{cases} 0 & x < 1 \\ 1/4 & 1 \leq x < 2 \\ 2/4 & 2 \leq x < 3 \\ 3/4 & 3 \leq x < 4 \\ 1 & x \geq 4 \end{cases}$$

♦♦

EXEMPLO 2.8 Função de distribuição e função de probabilidade

Dada a função

$$F(x) = \begin{cases} 0 & x < 1 \\ 1/2 & 1 \leq x < 2 \\ k & 2 \leq x < 3 \\ 3/4 & 3 \leq x < 4 \\ 1 & x \geq 4 \end{cases}$$

em que k é uma constante, determine os possíveis valores de k para que $F(x)$ seja a função de distribuição acumulada de uma variável aleatória X. Em seguida, determine a função de probabilidade desta v.a. X.

Solução

Como a função de distribuição de qualquer v.a. X tem que ser uma função não decrescente, concluímos que k tem que ser maior ou igual a $\frac{1}{2}$. Pela mesma razão, k tem que ser menor ou igual a $\frac{3}{4}$. Dessa forma, os possíveis valores de k pertencem ao intervalo $\left[\frac{1}{2}, \frac{3}{4}\right]$. Os valores possíveis da v.a. X correspondem aos pontos de descontinuidade da função $F(x)$. Logo, X assume os valores $1, 2, 3, 4$. As probabilidades desses valores são dadas pelo tamanho do "salto" de $F(x)$. Então, $\forall k : \dfrac{1}{2} \leq k \leq \dfrac{3}{4}$, temos:

$$\begin{aligned} \mathrm{P}(X = 1) &= \frac{1}{2} \\ \mathrm{P}(X = 2) &= k - \frac{1}{2} \\ \mathrm{P}(X = 3) &= \frac{3}{4} - k \\ \mathrm{P}(X = 4) &= 1 - \frac{3}{4} = \frac{1}{4} \end{aligned}$$

♦♦

EXEMPLO 2.9 Demanda por produto

A demanda por um certo produto pode ser vista como uma variável aleatória X cuja função de probabilidade $f_X(x)$ é estimada por

Número de unidades demandadas x	1	2	3	4
$f_X(x) = \mathrm{P}(X = x)$	$0,25$	$0,45$	$0,15$	$0,15$

(a) Verifique que $f_X(x)$ realmente define uma função de probabilidade.

(b) Obtenha a função de distribuição acumulada de X.

(c) Usando a função de distribuição calculada no item anterior, calcule $\mathrm{P}(X \leq 3,5)$.

Solução

(a) $0,25+0,45+0,15+0,15 = 1$ e todos os valores são não negativos. Logo, f_X é uma função de probabilidade.

(b)

$$F_X(x) = \begin{cases} 0 & \text{se } x < 1 \\ 0{,}25 & \text{se } 1 \leq x < 2 \\ 0{,}70 & \text{se } 2 \leq x < 3 \\ 0{,}85 & \text{se } 3 \leq x < 4 \\ 1{,}00 & \text{se } x \geq 4 \end{cases}$$

(c) Temos que

$$\mathrm{P}(X \leq 3,5) = F_X(3,5) = 0,85$$

♦♦

EXEMPLO 2.10

Uma variável aleatória discreta X tem a seguinte função de probabilidade

$$f_X(x) = \begin{cases} \frac{k}{(x+2)!} & x = 0,1 \\ \\ 0 & x \neq 0 \text{ e } x \neq 1 \end{cases}$$

onde k é uma constante.

(a) Determine o valor de k.

(b) Calcule a função de distribuição $F_X(x)$.

Solução

(a) Os valores possíveis da v.a. são 0 e 1. Então, temos que ter

$$\begin{aligned} f_X(0) + f_X(1) &= 1 \Rightarrow \frac{k}{2!} + \frac{k}{3!} = 1 \Rightarrow \\ \frac{k}{2} + \frac{k}{6} &= 1 \Rightarrow \frac{3k+k}{6} = 1 \Rightarrow k = \frac{6}{4} = \frac{3}{2} \end{aligned}$$

Logo,

$$\begin{aligned} f_X(0) &= \frac{\frac{3}{2}}{2} = \frac{3}{4} \\ f_X(1) &= \frac{\frac{3}{2}}{6} = \frac{1}{4} \end{aligned}$$

(b) A função de distribuição de X é

$$F_X(x) = \begin{cases} 0 & \text{se } x < 0 \\ \frac{3}{4} & \text{se } 0 \leq x < 1 \\ 1 & \text{se } x \geq 1 \end{cases}$$

♦♦

2.3 Funções de Variáveis Aleatórias

Dada uma v.a. X, podemos obter outras variáveis aleatórias através de funções de X e, da mesma forma que calculamos a função de probabilidade de X, podemos calcular a função de probabilidade dessas novas variáveis.

EXEMPLO 2.11 Função de variável aleatória: $Y = X^2$

Considere a v.a. X cuja função de probabilidade é dada na tabela abaixo:

x	-2	-1	0	1	2	3
$f_X(x)$	0,1	0,2	0,2	0,3	0,1	0,1

Consideremos a função $Y = g(X) = X^2$. Então, Y é uma nova variável aleatória, cujos possíveis valores são $0, 1, 4, 9$. Para calcular as probabilidades desses valores, temos que identificar os valores de X que originaram cada um deles. Temos a seguinte equivalência de eventos:

$$\begin{aligned} \{Y = 0\} &\equiv \{X = 0\} \\ \{Y = 1\} &\equiv \{X = -1\} \cup \{X = 1\} \\ \{Y = 4\} &\equiv \{X = -2\} \cup \{X = 2\} \\ \{Y = 9\} &\equiv \{X = 3\} \end{aligned}$$

(O símbolo $\equiv$ representa "é equivalente a"). Como os eventos são mutuamente exclusivos, segue que

$$\begin{aligned} \mathrm{P}(Y=0) &= \mathrm{P}(X=0)=0,2 \\ \mathrm{P}(Y=1) &= \mathrm{P}(X=-1)+\mathrm{P}(X=1)=0,5 \\ \mathrm{P}(Y=4) &= \mathrm{P}(X=-2)+\mathrm{P}(X=2)=0,2 \\ \mathrm{P}(Y=9) &= \mathrm{P}(X=3)=0,1 \end{aligned}$$

e podemos resumir essa função de probabilidade como

y	0	1	4	9
$f_Y(y)$	0,2	0,5	0,2	0,1

(2.15)

♦♦

DEFINIÇÃO **Função de Variável Aleatória**

Seja X uma variável aleatória discreta com função de probabilidade $f_X(x)$. Se definimos uma nova v.a. $Y = g(X)$, onde g é uma função real qualquer, então a função de probabilidade de Y é calculada como

$$f_Y(y) = \mathrm{P}(Y = y) = \sum_{\{x \mid g(x)=y\}} f_X(x)$$

2.4 Esperança de Variáveis Aleatórias Discretas

No estudo de variáveis aleatórias e suas distribuições de probabilidades, associamos números aos pontos do espaço amostral, ou seja, o resultado é sempre uma variável quantitativa (note que os resultados cara e coroa não definem uma variável aleatória; para tal, temos que associar números, 0 e 1, por exemplo, a esses resultados). Sendo assim, faz sentido perguntar "qual é o valor médio da variável aleatória X?"

DEFINIÇÃO **Esperança de uma variável aleatória discreta**

Seja X uma variável aleatória discreta que assume os valores $x_1, x_2, \ldots$ com probabilidades $p_1, p_2, \ldots$ respectivamente. A **esperança** ou **média** de X é definida como

$$\mathrm{E}(X) = \sum_i p_i x_i = \sum_i x_i \, \mathrm{P}(X = x_i) \tag{2.16}$$

onde o somatório se estende por todos os valores possíveis de X.

Podemos ver, então, que a esperança de X é uma média dos seus valores, ponderada pelas respectivas probabilidades.

EXEMPLO 2.12 Vendas e comissões

Em determinado setor de uma loja de departamentos, o número de produtos vendidos em um dia pelos funcionários é uma variável aleatória P com a seguinte distribuição de probabilidades (esses números foram obtidos dos resultados de vários anos de estudo):

Número de produtos	0	1	2	3	4	5	6
Probabilidade de venda	0,1	0,4	0,2	0,1	0,1	0,05	0,05

Cada vendedor recebe comissões de venda, distribuídas da seguinte forma: se ele vende até dois produtos em um dia, ele ganha uma comissão de R$10,00 por produto vendido. A partir da terceira venda, a comissão passa para R$50,00 por produto. Qual é o número médio de produtos vendidos por cada vendedor e qual a comissão média de cada um deles?

Solução

O número médio de produtos vendidos por funcionário é

$$\begin{aligned} \mathrm{E}(P) &= 0 \times 0,1 + 1 \times 0,4 + 2 \times 0,2 + 3 \times 0,1 \\ &\quad +4 \times 0,1 + 5 \times 0,05 + 6 \times 0,05 \\ &= 2,05 \end{aligned}$$

Com relação à comissão, vamos construir sua função de probabilidade:

Número de produtos P	0	1	2	3	4	5	6
Comissão C	0	10	20	70	120	170	220
Probabilidade de venda	0,1	0,4	0,2	0,1	0,1	0,05	0,05

A partir dessa função de probabilidade, podemos calcular:

$$\begin{aligned} \mathrm{E}(C) &= 0 \times 0,1 + 10 \times 0,4 + 20 \times 0,2 + 70 \times 0,1 + \\ &\quad +120 \times 0,1 + 170 \times 0,05 + 220 \times 0,05 \\ &= 46,5 \end{aligned}$$

ou seja, a comissão média diária de cada vendedor é R$ 46,50.

♦♦

Note que a esperança de X tem a mesma unidade de medida dos valores de X.

2.4.1 Esperança de Funções de Variáveis Aleatórias

Vimos que é possível obter novas variáveis aleatórias a partir de funções $g(X)$ de uma variável X e através da função de probabilidade de X podemos obter a função de probabilidade de Y. Sendo assim, podemos calcular a esperança de Y. Foi exatamente isso o que fizemos no caso das comissões no exemplo anterior, onde tínhamos

$$C = \begin{cases} 10P, & \text{se } P \leq 2 \\ 20 + 50 \times (P-2), & \text{se } P > 2 \end{cases}$$

Analisando atentamente aquele exemplo e notando que, por definição de função, a cada valor de X corresponde um único $Y = g(X)$, obtemos o resultado geral sobre a esperança de funções de variáveis aleatórias.

DEFINIÇÃO Esperança de Funções de uma Variável Aleatória

Seja X uma variável aleatória discreta com função de probabilidade $f_X(x)$. Se definimos uma nova v.a. $Y = g(X)$, então

$$\mathrm{E}(Y) = \mathrm{E}[g(X)] = \sum_x g(x) f_X(x) \tag{2.17}$$

EXEMPLO 2.13

Considere a v.a. X, já analisada no Exemplo 2.11, onde calculamos $\mathrm{E}(X^2)$.

x	-2	-1	0	1	2	3
$f_X(x)$	0,1	0,2	0,2	0,3	0,1	0,1

Naquele exemplo, calculamos a função de probabilidade da v.a. $Y = X^2$, resumida em (2.15), e, a partir dela, podemos calcular:

$$\mathrm{E}(Y) = \mathrm{E}\left(X^2\right) = 0 \times 0,2 + 1 \times 0,5 + 4 \times 0,2 + 9 \times 0,1 = 2,2$$

Usando o resultado (2.17), podemos fazer simplesmente:

$$\begin{aligned} \mathrm{E}\left(X^2\right) &= (-2)^2 \times 0,1 + (-1)^2 \times 0,2 + 0^2 \times 0,2 + \\ &\quad +1^2 \times 0,3 + 2^2 \times 0,1 + 3^2 \times 0,1 = 2,2 \end{aligned}$$

sem necessidade do cálculo da função de probabilidade de Y.

♦♦

2.4.2 Propriedades da Esperança

No que segue, X é uma variável aleatória discreta com função de probabilidade $f_X(x)$ e $a, b \neq 0$ são constantes reais quaisquer. Temos, então, os seguintes resultados, cujas demonstrações são imediatas, a partir da definição de esperança:

$$\mathrm{E}(a) = a \tag{2.18}$$

$$E(X + a) = E(X) + a \tag{2.19}$$

$$E(bX) = b\,E(X) \tag{2.20}$$

$$x_{\min} \leq E(X) \leq x_{\max} \tag{2.21}$$

Nessa última propriedade, $x_{\min}$ e $x_{\max}$ são os valores mínimo e máximo da variável X.

2.5 Variância e desvio-padrão de uma variável aleatória

A esperança de uma variável aleatória X é o centro de gravidade da distribuição de probabilidades. Sendo assim, a esperança é uma medida de *posição*. No entanto, é possível que duas variáveis bem diferentes tenham a mesma esperança, como é o caso das duas distribuições apresentadas na **Figura 2.2**. Nestas duas distribuições, a dispersão dos valores é diferente.

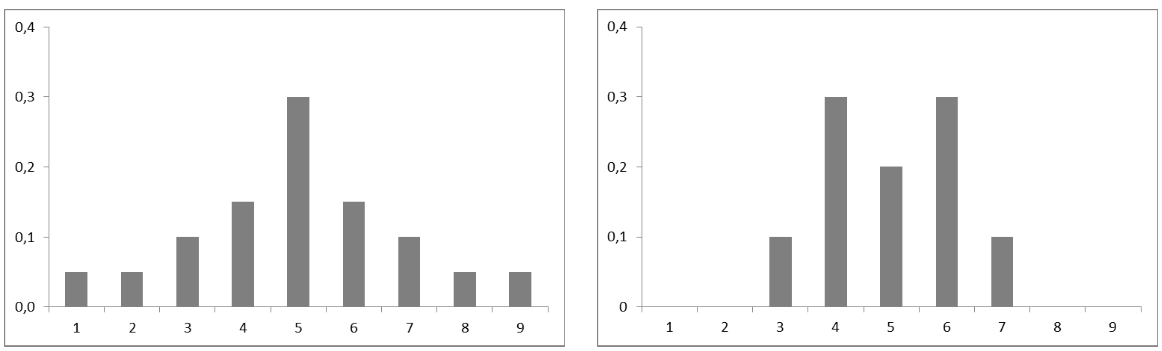

Figura 2.2 – Funções de probabilidade com mesma esperança e diferentes dispersões

A dispersão de uma variável aleatória X será, inicialmente, medida pela sua variância.

DEFINIÇÃO Variância de uma variável aleatória

A **variância** de uma variável aleatória X é definida como

$$\mathrm{Var}\,(X) = E\,[X - E\,(X)]^2 \tag{2.22}$$

O termo $X - E(X)$ é o *desvio em torno da média*. Sendo assim, a variância é a média dos desvios quadráticos em torno da média $E(X)$.

Vamos ver como calcular a variância de uma v.a. discreta. Para isso, vamos definir $g(X) = [X - \mathrm{E}(X)]^2$. Então, usando o resultado dado na equação (2.17), temos que

$$\operatorname{Var}(X) = \mathrm{E}[g(X)] = \sum_x [x - \mathrm{E}(X)]^2 f_X(x)$$

Desenvolvendo o quadrado e usando as propriedades do somatório e da esperança vistas na seção anterior, resulta

$$\begin{aligned}
\operatorname{Var}(X) &= \sum_x \left\{x^2 - 2x\,\mathrm{E}(X) + [\mathrm{E}(X)]^2\right\} f_X(x) = \\
&= \sum_x x^2 f_X(x) - 2\,\mathrm{E}(X) \sum_x x f_X(x) + [\mathrm{E}(X)]^2 \sum_x f_X(x) = \\
&= \sum_x x^2 f_X(x) - 2\,\mathrm{E}(X)\,\mathrm{E}(X) + [\mathrm{E}(X)]^2 \times 1 = \\
&= \sum_x x^2 f_X(x) - 2\,[\mathrm{E}(X)]^2 + [\mathrm{E}(X)]^2 = \\
&= \sum_x x^2 f_X(x) - [\mathrm{E}(X)]^2
\end{aligned}$$

Mas, se definimos $h(X) = X^2$, então $\mathrm{E}[h(X)] = \sum_x x^2 f_X(x)$. Logo, podemos escrever

$$\operatorname{Var}(X) = \mathrm{E}(X^2) - [\mathrm{E}(X)]^2 \tag{2.23}$$

que pode ser lida de maneira mais fácil como *a variância é a esperança do quadrado menos o quadrado da esperança*.

Da definição de variância, resulta que sua unidade de medida é o quadrado da unidade de medida da variável em estudo, sendo assim, uma unidade sem significado físico. Para se ter uma medida de dispersão na mesma unidade dos dados, define-se o *desvio-padrão* como a raiz quadrada da variância.

DEFINIÇÃO **Desvio padrão de uma variável aleatória**

O **desvio-padrão** de uma variável aleatória X é definido como a raiz quadrada de sua variância:

$$\mathrm{DP}(X) = \sqrt{\operatorname{Var}(X)} \tag{2.24}$$

2.5.1 Propriedades da variância e do desvio padrão

Sendo a variância e o desvio-padrão medidas de dispersão, é fácil ver que são válidas as seguintes propriedades, onde $a, b \neq 0$ são constantes quaisquer:

$$\operatorname{Var}(X) \geq 0 \tag{2.25}$$

$$\mathrm{DP}(X) \geq 0 \tag{2.26}$$

$$\operatorname{Var}(a) = 0 \tag{2.27}$$
$$\operatorname{DP}(a) = 0 \tag{2.28}$$

$$\operatorname{Var}(X + a) = \operatorname{Var}(X) \tag{2.29}$$
$$\operatorname{DP}(X + a) = \operatorname{DP}(X) \tag{2.30}$$

$$\operatorname{Var}(bX) = b^2 \operatorname{Var}(X) \tag{2.31}$$
$$\operatorname{DP}(bX) = |b| \operatorname{DP}(X) \tag{2.32}$$

EXEMPLO 2.14

Considere a v.a. Y com função de probabilidade dada por

y	-3	-1	0	2	5	8	9
$f_Y(y)$	$0,25$	$0,30$	$0,20$	$0,10$	$0,07$	$0,05$	$0,03$

e seja $Z = 2Y - 3$. Vamos calcular a esperança e a variância de Y e Z.

Solução

$$\begin{aligned} \mathrm{E}(Y) &= -3 \times 0,25 - 1 \times 0,30 + 0 \times 0,20 + 2 \times 0,10 \\ &\quad +5 \times 0,07 + 8 \times 0,05 + 9 \times 0,03 = 0,17 \end{aligned}$$

$$\mathrm{E}(Z) = 2 \times \mathrm{E}(Y) - 3 = 2 \times 0,17 - 3 = -2,66$$

Vamos calcular agora $\mathrm{E}(Y^2)$:

$$\begin{aligned} \mathrm{E}(Y^2) &= 9 \times 0,25 + 1 \times 0,30 + 0 \times 0,20 + 4 \times 0,10 \\ &\quad +25 \times 0,07 + 64 \times 0,05 + 81 \times 0,03 = 10,33 \end{aligned}$$

Logo

$$\operatorname{Var}(Y) = 10,33 - 0,17^2 = 10,3011$$

Usando as propriedades da variância, temos que

$$\operatorname{Var}(Z) = 2^2 \times \operatorname{Var}(Y) = 41,2044$$

♦♦

EXEMPLO 2.15

Um lojista mantém extensos registros das vendas diárias de certo aparelho. O quadro a seguir, dá a distribuição de probabilidades do número de aparelhos vendidos em uma semana. Se o lucro por unidade vendida é de R$500,00, qual o lucro esperado em uma semana? Qual é o desvio-padrão do lucro?

x= número de aparelhos	0	1	2	3	4	5
$f_X(x)$	0,1	0,1	0,2	0,3	0,2	0,1

Solução

Seja X o número de aparelhos vendidos em uma semana e seja L o lucro semanal. Então, $L = 500X$.

$$\begin{aligned} \mathrm{E}(X) &= 0 \times 0,1 + 1 \times 0,1 + 2 \times 0,2 + 3 \times 0,3 + 4 \times 0,2 + 5 \times 0,1 \\ &= 2,7 \text{ aparelhos} \end{aligned}$$

$$\begin{aligned} \mathrm{E}\left(X^2\right) &= 0^2 \times 0,1 + 1^2 \times 0,1 + 2^2 \times 0,2 + 3^2 \times 0,3 + 4^2 \times 0,2 + 5^2 \times 0,1 \\ &= 10,2 \text{ aparelhos}^2 \end{aligned}$$

$$\mathrm{Var}(X) = 10,2 - (2,7)^2 = 2,91 \text{ aparelhos}^2$$

$$\mathrm{DP}(X) = 1,706 \text{ aparelhos}$$

Com relação ao lucro semanal, temos que

$$\mathrm{E}(L) = 500\,\mathrm{E}(X) = R\$1350,00$$

$$\mathrm{DP}(L) = 500\,\mathrm{DP}(X) = R\$852,94$$

♦♦

EXEMPLO 2.16

Seja uma v.a. X com função de probabilidade dada na tabela a seguir:

x	1	2	3	4	5
$f_X(x)$	p^2	p^2	p	p	p^2

(a) Encontre o valor de p para que $f_X(x)$ seja, de fato, uma função de probabilidade.

(b) Calcule $\mathrm{P}(X \geq 4)$ e $\mathrm{P}(X < 3)$.

(c) Calcule $\mathrm{P}(|X - 3| \geq 2)$.

(d) Calcule $\mathrm{E}(X)$ e $\mathrm{Var}(X)$.

Solução

(a) Como $\sum_x f_X(x) = 1$, temos que ter:

$$3p^2 + 2p = 1 \Rightarrow 3p^2 + 2p - 1 = 0 \Rightarrow$$
$$p = \frac{-2 \pm \sqrt{4+12}}{6} = \frac{-2 \pm 4}{6} \Rightarrow \begin{cases} p = -1 \\ \text{ou} \\ p = \frac{1}{3} \end{cases}$$

Como p é uma probabilidade, temos que ter $p \geq 0$. Logo, o valor correto é $p = \frac{1}{3}$.

(b) $\mathrm{P}(X \geq 4) = \mathrm{P}(X = 4) + \mathrm{P}(X = 5) = p + p^2 = \frac{1}{3} + \frac{1}{9} = \frac{4}{9}$.

$Pr(X < 3) = \mathrm{P}(X = 1) + \mathrm{P}(X = 2) = 2p^2 = \frac{2}{9}$.

(c) Aqui temos que notar o seguinte fato sobre a função módulo, ilustrado na **Figura 2.3**. Valores $y = |x|$ no eixo vertical menores que k (abaixo da linha horizontal sólida) correspondem a valores de x no intervalo $(-k, k)$ e valores y no eixo vertical maiores que k correspondem ou a $x > k$ ou a $x < -k$. Mais precisamente,

$$\begin{array}{lcl} |\,x\,| \geq k & \Leftrightarrow & x \geq k \quad \text{ou} \quad x \leq -k \\ |\,x\,| \leq k & \Leftrightarrow & -k \leq x \leq k \end{array}$$

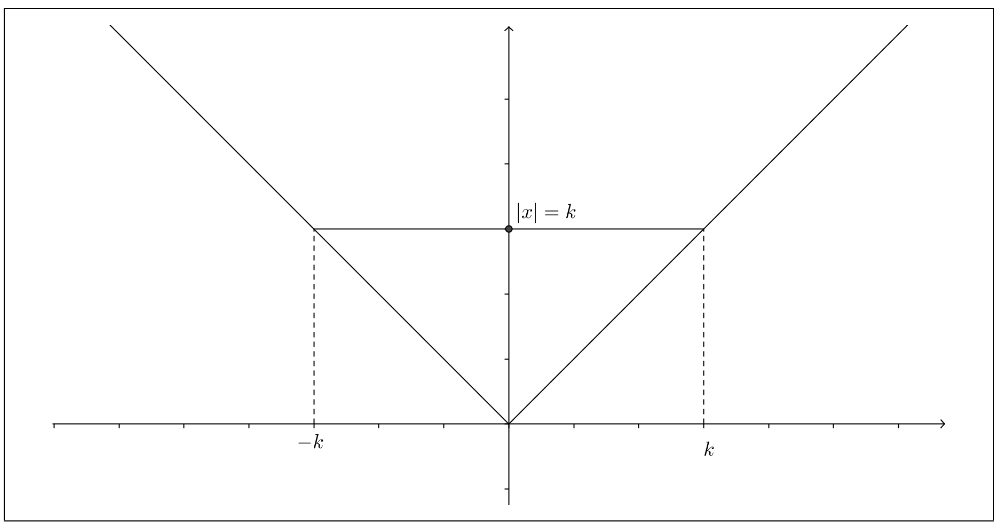

Figura 2.3 – Função módulo

Usando esses fatos, temos que

$$\begin{aligned} \mathrm{P}\left(|X-3| \geq 2\right) &= \mathrm{P}\left(\{X - 3 \leq -2\} \;\cup\; \{X - 3 \geq 2\}\right) = \\ &= \mathrm{P}\left(X - 3 \leq -2\right) + \mathrm{P}\left(X - 3 \geq 2\right) = \\ &= \mathrm{P}\left(X \leq 1\right) + \mathrm{P}\left(X \geq 5\right) = \\ &= \mathrm{P}\left(X = 1\right) + \mathrm{P}(X = 5) = \\ &= 2p^2 = \frac{2}{9} \end{aligned}$$

(d) Temos que

$$\begin{aligned} E(X) &= 1 \times p^2 + 2 \times p^2 + 3 \times p + 4 \times p + 5 \times p^2 \\ &= \frac{1}{9} + \frac{2}{9} + 1 + \frac{4}{3} + \frac{5}{9} \\ &= \frac{29}{9} = 3,2222 \end{aligned}$$

$$\begin{aligned} E(X^2) &= 1^2 \times p^2 + 2^2 \times p^2 + 3^2 \times p + 4^2 \times p + 5^2 \times p^2 \\ &= \frac{1}{9} + \frac{4}{9} + 3 + \frac{16}{3} + \frac{25}{9} \\ &= \frac{105}{9} = \frac{35}{3} \\ \mathrm{Var}(X) &= \frac{35}{3} - \left(\frac{29}{9}\right)^2 = \frac{14}{81} \end{aligned}$$

♦♦

EXEMPLO 2.17 Jogo de dados

Um jogador A paga R$5,00 a B e lança um dado. Se sair face 3, ganha R$20,00. Se sair face 4, 5, ou 6, perde. Se sair face 1 ou 2, tem o direito de jogar novamente. Desta vez, lança dois dados. Se saírem duas faces 6, ganha R$50,00. Se sair uma face 6, recebe o dinheiro de volta. Nos demais casos, perde. Seja L o lucro líquido do jogador A nesse jogo. Calcule a função de probabilidade de L e o lucro esperado do jogador A.

Solução

Sabemos que o dado é honesto e que os lançamentos são independentes. O diagrama de árvore para o espaço amostral desse experimento é dado na **Figura 2.4**.

Para calcular a probabilidade dos eventos associados aos lançamentos dos dois dados (parte inferior da árvore), usamos o fato de que a probabilidade da interseção de eventos independentes é o produto das probabilidades.

No cálculo da probabilidade de uma face 6, multiplicamos por 2, porque a face 6 pode estar em qualquer um dos dois dados.

Vemos que os valores do lucro L são: -5; 0; 15; 45 e a função de probabilidade de L é

Lucro ℓ	-5	0	15	45
$P(L = \ell)$	$\frac{1}{2} + \frac{2}{6} \times \frac{5}{6} \times \frac{5}{6}$	$\frac{2}{6} \times 2 \times \frac{1}{6} \times \frac{5}{6}$	$\frac{1}{6}$	$\frac{2}{6} \times \frac{1}{6} \times \frac{1}{6}$

ou

Lucro ℓ	-5	0	15	45
$P(L = \ell)$	$\frac{158}{216}$	$\frac{20}{216}$	$\frac{36}{216}$	$\frac{2}{216}$

$$E(L) = -5 \times \frac{158}{216} + 15 \times \frac{36}{216} + 45 \times \frac{2}{216} = -\frac{160}{216} = -0,74$$

♦♦

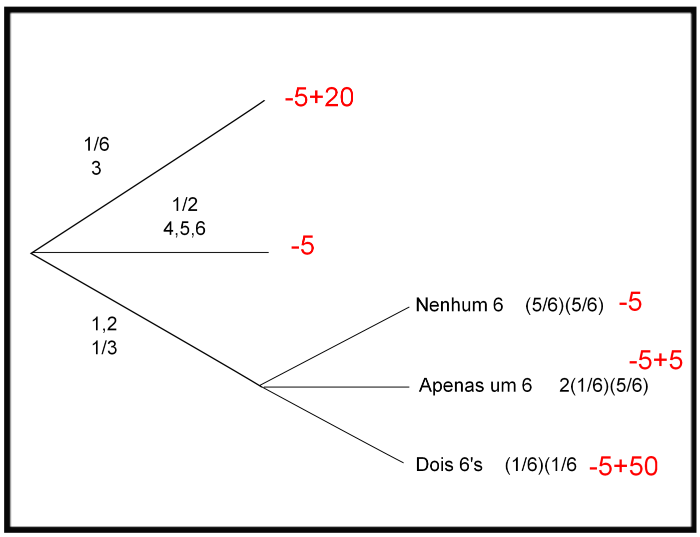

Figura 2.4 – Espaço amostral para o Exemplo 2.17

Capítulo 3

Algumas Distribuições Discretas

3.1 Introdução

Considere as seguintes situações:

1. (a) Lança-se uma moeda viciada e observa-se o resultado obtido e (b) pergunta-se a um eleitor se ele vai votar no candidato A ou B.

2. (a) Lança-se uma moeda n vezes e observa-se o número de caras obtidas e (b) de uma grande população, extrai-se uma amostra de n eleitores e pergunta-se a cada um deles em qual dos candidatos A ou B eles votarão e conta-se o número de votos do candidato A.

3. (a) De uma urna com P bolas vermelhas e Q bolas brancas, extraem-se n bolas sem reposição e conta-se o número de bolas brancas e (b) de uma população com P pessoas a favor do candidato A e Q pessoas a favor do candidato B, extrai-se uma amostra de tamanho n sem reposição e conta-se o número de pessoas a favor do candidato A na amostra.

Em cada uma das situações anteriores, os experimentos citados têm algo em comum: em certo sentido, temos a "mesma situação ", mas em contextos diferentes. Por exemplo, na situação 1, cada um dos experimentos tem dois resultados possíveis e observamos o resultado obtido. Na situação 3, temos uma população dividida em duas categorias e dela extraímos uma amostra sem reposição; o interesse está no número de elementos de uma determinada categoria.

Na prática, existem muitas outras situações que podem se "encaixar" nos modelos acima e mesmo em outros modelos. O que veremos nesse capítulo são alguns *modelos* de variáveis aleatórias discretas que podem descrever situações como as listadas anteriormente. Nesse contexto, um modelo será definido por uma variável aleatória e sua função de probabilidade, explicitando-se claramente as hipóteses de validade. De posse desses elementos, poderemos analisar diferentes situações práticas para tentar "encaixá-las" em algum dos modelos dados.

Neste capítulo, serão descritas as distribuições de probabilidade discretas mais usuais.A introdução de cada uma delas será feita através de um exemplo clássico (moeda, urna, baralho etc.) e, em seguida, serão explicitadas as características do experimento. Tais características são a ferramenta necessária para sabermos qual modelo se aplica a uma determinada situação prática. Definida a distribuição, calculam-se a média e a variância.

3.2 Distribuição Uniforme Discreta

Suponha que seu professor de Estatística decida dar de presente a um dos alunos um livro de sua autoria. Não querendo favorecer qualquer aluno em especial, ele decide sortear aleatoriamente o ganhador, dentre os 45 alunos da turma. Para isso, ele numera os nomes dos alunos que constam do diário de classe de 1 a 45, escreve esses números em pedaços iguais de papel, dobrando-os ao meio para que o número não fique visível, e sorteia um desses papéis depois de bem misturados. Qual é a probabilidade de que você ganhe o livro? Qual é a probabilidade de que o aluno que tirou a nota mais baixa na primeira prova ganhe o livro? E o que tirou a nota mais alta?

O importante a notar nesse exemplo é o seguinte: o professor tomou todos os cuidados necessários para não favorecer qualquer aluno em especial. Isso significa que todos os alunos têm a mesma chance de ganhar o livro. Temos, assim, um exemplo da distribuição uniforme discreta.

DEFINIÇÃO Distribuição uniforme discreta

A variável aleatória discreta X, que assume os valores $x_1, x_2, \ldots, x_n$, tem **distribuição uniforme** se

$$f_X(x_i) = \mathrm{P}(X = x_i) = \frac{1}{n} \qquad \forall i = 1, 2, \ldots, n \tag{3.1}$$

Note que, em uma distribuição discreta uniforme, todos os valores são igualmente prováveis. Além disso, para que uma v.a. X tenha distribuição uniforme discreta, é necessário que X assuma um número *finito* de valores, já que $\sum_x f_X(x) = 1$.

3.2.1 Esperança e Variância

Seja X uma v.a. discreta uniforme que assume valores $x_1, x_2, \ldots, x_n$. Por definição, a esperança de X é

$$\mathrm{E}(X) = \frac{1}{n}x_1 + \frac{1}{n}x_2 + \cdots + \frac{1}{n}x_n = \overline{x},$$

ou seja, $\mathrm{E}(X)$ é a média aritmética dos valores possíveis de X.

Com relação à variância, temos, por definição, que

$$\begin{aligned}
\text{Var}(X) &= \text{E}\left[X - \text{E}(X)\right]^2 \\
&= \frac{1}{n}(x_1 - \overline{x})^2 + \frac{1}{n}(x_2 - \overline{x})^2 + \cdots + \frac{1}{n}(x_n - \overline{x})^2 = \sigma_X^2
\end{aligned}$$

EXEMPLO 3.1 Lançamento de uma moeda

Considere o lançamento de uma moeda. Vamos definir a seguinte variável aleatória X associada a esse experimento:

$$\begin{aligned}
X = 0, &\quad \text{se ocorre cara} \\
X = 1, &\quad \text{se ocorre coroa}
\end{aligned}$$

Para que essa v.a. tenha distribuição uniforme, é necessário supor que a moeda seja honesta e, nesse caso,

$$\begin{aligned}
f_X(0) &= f_X(1) = \frac{1}{2} \\
\text{E}(X) &= \frac{0+1}{2} = \frac{1}{2} \\
\text{Var}(X) &= \frac{1}{2} \times \left(0 - \frac{1}{2}\right)^2 + \frac{1}{2} \times \left(1 - \frac{1}{2}\right)^2 \\
&= \frac{1}{2} \times \frac{1}{4} + \frac{1}{2} \times \frac{1}{4} = \frac{1}{4}
\end{aligned}$$

♦♦

EXEMPLO 3.2 Conserto de máquina

Os defeitos em determinada máquina ocorrem aproximadamente na mesma frequência. Dependendo do tipo de defeito, o técnico leva 1, 2, 3, 4 ou 5 horas para consertar a máquina.

(a) Descreva o modelo probabilístico apropriado para representar a duração do tempo de reparo da máquina.

(b) Qual é o tempo médio de reparo desta máquina? E o desvio-padrão deste tempo de reparo?

(c) São 15 horas e acaba de ser entregue uma máquina para reparo. A jornada normal de trabalho do técnico termina às 17 horas. Qual é a probabilidade de que o técnico não precise fazer hora extra para terminar o conserto desta máquina?

Solução

Seja T = "tempo de reparo, em horas".

(a) Como os defeitos ocorrem na mesma frequência, o modelo probabilístico apropriado é uma distribuição uniforme:

t	1	2	3	4	5
$f_T(t) = \mathrm{P}(T = t)$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$

(b) $\mathrm{E}(T) = \dfrac{1+2+3+4+5}{5} = 3$ horas

$\mathrm{Var}(T) = \dfrac{1^2+2^2+3^2+4^2+5^2}{5} - 9 = 2 \Longrightarrow DP(T) = 1,41$ horas

(c) Seja E o evento "técnico vai ter que fazer hora extra". Então

$$\mathrm{P}(E) = \mathrm{P}(T > 2) = \frac{3}{5} = 0,6$$

Logo, a probabilidade de que ele não tenha que fazer hora extra é 0,4.

♦♦

3.3 Distribuição de Bernoulli

Considere o lançamento de uma moeda. A característica de tal experimento aleatório é que ele possui apenas dois resultados possíveis. Uma situação análoga surge quando da extração da carta de um baralho, em que o interesse está apenas na cor (preta ou vermelha) da carta sorteada.

DEFINIÇÃO Experimento de Bernoulli

Um **experimento de Bernoulli** é um experimento aleatório com apenas dois resultados possíveis; por convenção, um deles é chamado "sucesso" e o outro, "fracasso".

DEFINIÇÃO Variável aleatória de Bernoulli

A **v.a. de Bernoulli** é a v.a. X associada a um experimento de Bernoulli, em que se define

$$X = \begin{cases} 1 & \text{se ocorre sucesso} \\ 0 & \text{se ocorre fracasso} \end{cases}$$

Chamando de p a probabilidade de sucesso ($0 < p < 1$), a **distribuição de Bernoulli** é

x	0	1
$f_X(x)$	$1-p$	p

(3.2)

Obviamente, as condições definidoras de uma função de probabilidade são satisfeitas, uma vez que

$$p > 0, \quad 1 - p > 0 \quad \text{e} \quad p + (1 - p) = 1.$$

O valor de p é o único valor que precisamos conhecer para determinar completamente a distribuição; ele é, então, chamado *parâmetro* da distribuição de Bernoulli. Vamos denotar a distribuição de Bernoulli com parâmetro p por $Bern(p)$.

A função de distribuição acumulada é dada por:

$$F_X(x) = \begin{cases} 0 & \text{se } x < 0 \\ 1 - p & \text{se } 0 \leq x < 1 \\ 1 & \text{se } x \geq 1 \end{cases} \tag{3.3}$$

Na **Figura 3.1**, temos os gráficos da função de probabilidade e da função de distribuição acumulada de uma variável de Bernoulli.

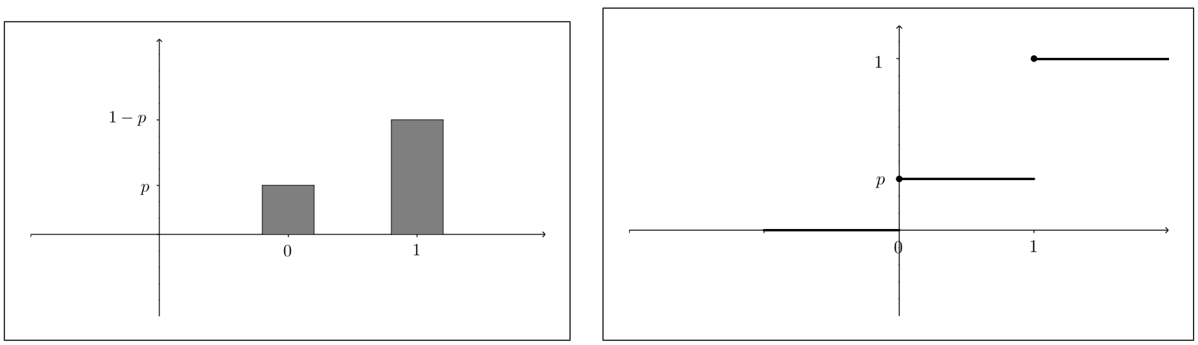

Figura 3.1 – Distribuição de Bernoulli: função de probabilidade e função de distribuição

3.3.1 Esperança e Variância

Seja $X \sim Bern(p)$ (lê-se: a variável aleatória X tem distribuição de Bernoulli com parâmetro p). Então,

$$E(X) = 0 \times (1 - p) + 1 \times p = p$$

$$E(X^2) = 0^2 \times (1 - p) + 1^2 \times p = p$$

$$Var(X) = E(X^2) - [E(X)]^2 = p - p^2$$

Em resumo:

$$X \sim Bern(p) \quad \Rightarrow \quad \begin{cases} E(X) = p \\ \\ Var(X) = p(1 - p) \end{cases} \tag{3.4}$$

É comum denotar a probabilidade de fracasso por q, isto é, $q = 1 - p$.

EXEMPLO 3.3 Lançamento de uma moeda

Considere novamente o lançamento de uma moeda e a seguinte variável aleatória X associada a esse experimento:

$$X = \begin{cases} 1 & \text{se ocorre cara} \\ 0 & \text{se ocorre coroa} \end{cases}$$

Seja p a probabilidade de cara, $0 < p < 1$. Então, X tem distribuição de Bernoulli com parâmetro p. Note que, nesse caso, a Bernoulli com parâmetro $p = 1/2$ é equivalente à distribuição uniforme.

♦♦

EXEMPLO 3.4 Auditoria da Receita Federal

Um auditor da Receita Federal examina declarações de Imposto de Renda de pessoas físicas, cuja variação patrimonial ficou acima do limite considerado aceitável. De dados históricos, sabe-se que 10% dessas declarações são fraudulentas.

Vamos considerar o experimento correspondente ao sorteio aleatório de uma dessas declarações. Esse é um experimento de Bernoulli, em que o sucesso equivale à ocorrência de declaração fraudulenta e o parâmetro da distribuição de Bernoulli é $p = 0,1$.

Esse exemplo ilustra o fato de que "sucesso", nesse contexto, nem sempre significa uma situação feliz na vida real. Aqui, sucesso é definido de acordo com o interesse estatístico no problema. Em uma situação mais dramática, "sucesso" pode indicar a morte de um paciente, por exemplo.

♦♦

3.4 Distribuição Binomial

Vamos introduzir a distribuição binomial, uma das mais importantes distribuições discretas, através de um exemplo. Em seguida, discutiremos as hipóteses feitas e apresentaremos os resultados formais sobre tal distribuição e novos exemplos.

EXEMPLO 3.5 Lançamentos de uma moeda

Considere o seguinte experimento: uma moeda é lançada 4 vezes e sabe-se que $p = \mathrm{P}(\text{cara})$. Vamos definir a seguinte variável aleatória associada a este experimento:

$$X = \text{ número de caras}$$

Como visto antes, cada lançamento da moeda representa um experimento de Bernoulli e como o interesse está no número de caras, vamos definir sucesso = cara.

Para encontrar a função de probabilidade de X, o primeiro fato a notar é que os valores possíveis de X são: 0, que equivale à ocorrência de nenhuma cara e, portanto, de 4 coroas; 1, que equivale à ocorrência de apenas 1 cara e, portanto, 3 coroas; 2, que equivale à ocorrência de 2 caras e, portanto, 2 coroas; 3, que equivale à ocorrência de 3 caras e 1 coroa e, finalmente, 4, que equivale à ocorrência de 4 caras e nenhuma coroa. Assim, os possíveis valores de X são

$$X = 0, 1, 2, 3, 4$$

Vamos, agora, calcular a probabilidade de cada um desses valores, de modo a completar a especificação da função de probabilidade de X. Para isso, vamos representar por K_i o evento "cara no i-ésimo lançamento" e por C_i o evento "coroa no i-ésimo lançamento".

- $X = 0$

  Temos a seguinte equivalência de eventos:

  $$\{X = 0\} \equiv C_1 \cap C_2 \cap C_3 \cap C_4$$

  É razoável supor que os lançamentos da moeda sejam eventos independentes, ou seja, o resultado de um lançamento não interfere no resultado de qualquer outro lançamento.

  Dessa forma, os eventos C_i e K_j são independentes para $i \neq j$. (Note que os eventos C_i e K_i são mutuamente exclusivos e, portanto, não são independentes – se sair cara em um lançamento específico, não é possível sair coroa nesse mesmo lançamento e vice-versa).

  Analogamente, os eventos C_i e C_j são independentes para $i \neq j$, bem como os eventos K_i e K_j, $i \neq j$. Pela regra da probabilidade da interseção de eventos independentes, resulta que

  $$\begin{aligned} \mathrm{P}\left(C_1 \cap C_2 \cap C_3 \cap C_4\right) &= \mathrm{P}(C_1) \times \mathrm{P}(C_2) \times \mathrm{P}(C_3) \times \mathrm{P}(C_4) \\ &= (1-p) \times (1-p) \times (1-p) \times (1-p) \\ &= (1-p)^4 \end{aligned}$$

- $X = 1$

  O evento $X = 1$ corresponde à ocorrência de 1 cara e, consequentemente, de 3 coroas. Uma sequência possível de lançamentos é

  $$K_1 \cap C_2 \cap C_3 \cap C_4.$$

  Vamos calcular a probabilidade desse resultado. Como antes, os lançamentos são eventos independentes e, portanto,

  $$\begin{aligned} \mathrm{P}(K_1 \cap C_2 \cap C_3 \cap C_4) &= \mathrm{P}(K_1) \times \mathrm{P}(C_2) \times \mathrm{P}(C_3) \times \mathrm{P}(C_4) \\ &= p \times (1-p) \times (1-p) \times (1-p) \\ &= p(1-p)^3 \end{aligned}$$

  Mas qualquer sequência com 1 cara resulta em $X = 1$, ou seja, a face cara pode estar em qualquer uma das quatro posições e todas essas sequências resultam em $X = 1$. Além disso, definida a posição da face cara, as posições das faces coroas já estão determinadas – são as posições restantes. Então, temos a seguinte equivalência:

  $$\begin{aligned} \{X = 1\} \equiv\ & \{K_1 \cap C_2 \cap C_3 \cap C_4\} \cup \{C_1 \cap K_2 \cap C_3 \cap C_4\} \cup \\ & \{C_1 \cap C_2 \cap K_3 \cap C_4\} \cup \{C_1 \cap C_2 \cap C_3 \cap K_4\} \end{aligned}$$

Mas os eventos que aparecem no lado direito da expressão anterior são eventos mutuamente exclusivos. Logo,

$$\begin{aligned}
\mathrm{P}(X=1) &= \mathrm{P}(K_1 \cap C_2 \cap C_3 \cap C_4) \\
&\quad + \mathrm{P}(C_1 \cap K_2 \cap C_3 \cap C_4) \\
&\quad + \mathrm{P}(C_1 \cap C_2 \cap K_3 \cap C_4) \\
&\quad + \mathrm{P}(C_1 \cap C_2 \cap C_3 \cap K_4) \\
&= p \times (1-p) \times (1-p) \times (1-p) \\
&\quad +(1-p) \times p \times (1-p) \times (1-p) \\
&\quad +(1-p) \times (1-p) \times p \times (1-p) \\
&\quad +(1-p) \times (1-p) \times (1-p) \times p \\
&= 4p(1-p)^3
\end{aligned}$$

- $X = 2$

  O evento $X = 2$ corresponde à ocorrência de 2 caras e, consequentemente, de 2 coroas. Qualquer uma dessas sequências tem probabilidade $p^2(1-p)^2$.

  As sequências de lançamentos com 2 caras e 2 coroas são as seguintes:

$$\begin{array}{l}
K_1K_2C_3C_4 \\
K_1C_2K_3C_4 \\
K_1C_2C_3K_4 \\
C_1C_2K_3K_4 \\
C_1K_2C_3K_4 \\
C_1K_2K_3C_4
\end{array}$$

  Todas essas 6 sequências têm a mesma probabilidade e correspondem a eventos mutuamente exclusivos. Temos a seguinte equivalência:

$$\begin{aligned}
\{X=2\} \equiv\ & (K_1 \cap K_2 \cap C_3 \cap C_4) \cup (K_1 \cap C_2 \cap K_3 \cap C_4) \cup \\
& (K_1 \cap C_2 \cap C_3 \cap K_4) \cup (C_1 \cap C_2 \cap K_3 \cap K_4) \cup \\
& (C_1 \cap K_2 \cap C_3 \cap K_4) \cup (C_1 \cap K_2 \cap K_3 \cap C_4)
\end{aligned}$$

  e, portanto,

$$\begin{aligned}
\mathrm{P}(X=2) &= \mathrm{P}(K_1 \cap K_2 \cap C_3 \cap C_4) + \mathrm{P}(K_1 \cap C_2 \cap K_3 \cap C_4) + \\
&\quad \mathrm{P}(K_1 \cap C_2 \cap C_3 \cap K_4) + \mathrm{P}(C_1 \cap C_2 \cap K_3 \cap K_4) + \\
&\quad \mathrm{P}(C_1 \cap K_2 \cap C_3 \cap K_4) + \mathrm{P}(C_1 \cap K_2 \cap K_3 \cap C_4) \\
&= 6p^2(1-p)^2
\end{aligned}$$

- $X = 3$ e $X = 4$

Os casos $X = 3$ e $X = 4$ são análogos aos casos $X = 1$ e $X = 0$, respectivamente; basta trocar caras por coroas e vice-versa. Assim,

$$\begin{aligned} \mathrm{P}(X = 3) &= 4p^3(1 - p) \\ \mathrm{P}(X = 4) &= p^4 \end{aligned}$$

É importante notar que a hipótese de independência dos lançamentos da moeda foi absolutamente fundamental na solução do exemplo; foi ela que nos permitiu multiplicar as probabilidades dos resultados de cada lançamento para obter a probabilidade da sequência completa de n lançamentos. Embora essa hipótese seja muito razoável nesse exemplo, ainda assim é uma hipótese "subjetiva".

Outra propriedade utilizada foi a da probabilidade da união de eventos mutuamente exclusivos. Mas aqui essa propriedade é óbvia, ou seja, não há qualquer subjetividade: os eventos $C_1 \cap K_2$ e $K_1 \cap C_2$ são mutuamente exclusivos, pois no primeiro lançamento ou sai cara ou sai coroa; não pode sair cara e coroa no primeiro lançamento, ou seja, cada lançamento é um experimento de Bernoulli.

♦♦

EXEMPLO 3.6 Bolas em uma urna

Uma urna contém quatro bolas brancas e seis bolas verdes. Três bolas são retiradas dessa urna, com reposição, isto é, depois de tirada a primeira bola, ela é recolocada na urna e sorteia-se a segunda, que também é recolocada na urna para, finalmente, ser sorteada a terceira bola. Vamos definir a seguinte variável aleatória associada a esse experimento:

$$X = \text{ "número de bolas brancas sorteadas"}$$

O importante a notar aqui é o seguinte: como cada bola sorteada é recolocada na urna antes da próxima extração, a composição da urna é sempre a mesma e o resultado de uma extração não afeta o resultado de outra extração qualquer. Dessa forma, em todas as extrações a probabilidade de bola branca (e também bola verde) é a mesma e podemos considerar as extrações como independentes. Assim, temos uma situação análoga à do exemplo anterior: temos três repetições de um experimento (sorteio de uma bola), essas repetições são independentes, em cada uma delas há dois resultados possíveis – bola branca (sucesso) ou bola verde (fracasso) – e as probabilidades de sucesso e fracasso são as mesmas. Assim, cada extração equivale a um experimento de Bernoulli e como o interesse está nas bolas brancas, vamos considerar sucesso = bola branca e da observação anterior resulta que

$$\mathrm{P}(sucesso) = \frac{4}{10}$$

Os valores possíveis de X são $0, 1, 2, 3$, uma vez que são feitas três extrações. Vamos calcular a probabilidade de cada um dos valores de X. Como antes, vamos denotar por V_i o evento "bola verde na i-ésima extração" e por B_i o evento "bola branca na i-ésima extração". Da discussão anterior, resulta que, para $i \neq j$, os eventos V_i e B_j são independentes, assim como os eventos B_i e B_j e os eventos V_i e V_j.

- $X = 0$

  Esse resultado equivale à extração de bolas verdes em todas as três extrações.

$$\{X = 0\} \equiv \{V_1 \cap V_2 \cap V_3\}$$

  Logo,

$$\begin{aligned} P(X = 0) &= P(V_1 \cap V_2 \cap V_3) \\ &= P(V_1) \times P(V_2) \times P(V_3) \\ &= \frac{6}{10} \times \frac{6}{10} \times \frac{6}{10} \\ &= \left(\frac{6}{10}\right)^3 \end{aligned}$$

- $X = 1$

  Esse resultado equivale à extração de uma bola branca e, por consequência, duas bolas verdes. A bola branca pode sair em qualquer uma das três extrações e, definida a posição da bola branca, as posições das bolas verdes ficam totalmente estabelecidas. Logo,

$$P(X = 1) = 3\left(\frac{4}{10}\right)\left(\frac{6}{10}\right)^2$$

- $X = 2$ e $X = 3$

  Os casos $X = 2$ e $X = 3$ são análogos aos casos $X = 1$ e $X = 0$, respectivamente; basta trocar bola branca por bola verde e vice-versa. Assim,

$$P(X = 2) = 3\left(\frac{4}{10}\right)^2\left(\frac{6}{10}\right)$$

$$P(X = 3) = \left(\frac{4}{10}\right)^3$$

♦♦

Esses dois exemplos ilustram a *distribuição binomial*, que depende de dois parâmetros: o número de repetições e a probabilidade de sucesso de um experimento de Bernoulli. No Exemplo 3.5, $n = 4$ e temos uma probabilidade de sucesso qualquer p. No Exemplo 3.6, $n = 3$ e $p = \frac{4}{10}$.

3.4.1 A Distribuição Binomial

Nos dois exemplos anteriores, tínhamos repetições de um experimento de Bernoulli que podiam ser consideradas independentes e a probabilidade de sucesso se mantinha constante ao longo de todas as repetições. Essas são as condições definidoras de um *experimento binomial*.

DEFINIÇÃO **Experimento binomial**

Um **experimento binomial** consiste em repetições *independentes* de um experimento de Bernoulli com probabilidade p de sucesso, probabilidade essa que permanece constante em todas as repetições.

A variável aleatória que associamos aos experimentos binomiais dos dois exemplos foi

$$X = \text{“número de sucessos”}$$

Se o experimento binomial consiste em n repetições, então os valores possíveis de X são $0, 1, 2, \cdots, n$. O evento $X = x$ corresponde a todas as sequências de resultados com x sucessos e $n - x$ fracassos. Como as repetições são independentes, cada uma dessas sequências tem probabilidade $p^x(1-p)^{n-x}$. O número total de tais sequências é dado pelo *coeficiente binomial* $\binom{n}{x}$, definido a seguir.

DEFINIÇÃO **Coeficiente binomial**

Para n e k inteiros cm $n \geq k$, define-se o coeficiente binomial como

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \tag{3.5}$$

em que $n!$ representa o *fatorial* de n, definido como

$$n! = n \times (n-1) \times (n-2) \times \cdots \times 2 \times 1 \tag{3.6}$$

Por definição, $0! = 1$.

Temos condições, agora, de definir a *variável aleatória binomial*.

DEFINIÇÃO **Variável aleatória binomial**

Para um experimento binomial consistindo em n repetições independentes de um experimento de Bernoulli com parâmetro p, defina a variável aleatória

$$X = \text{"número de sucessos"}$$

Então, X tem **distribuição binomial** com parâmetros n e p, cuja função de probabilidade é dada por

$$\mathrm{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \qquad k = 0, 1, 2, \ldots, n \tag{3.7}$$

Verificação das condições definidoras de uma função de probabilidade

É imediato ver, da Equação (3.7), que $\mathrm{P}(X = k) \geq 0$. Para mostrar que a soma das probabilidades é 1, usaremos o teorema do binômio de Newton:

TEOREMA 3.1 ***Dados dois números reais quaisquer*** x ***e*** a ***e um inteiro qualquer*** n***, então***

$$(x+a)^n = \sum_{k=0}^{n} \binom{n}{k} a^k x^{n-k} \tag{3.8}$$

■

Aplicando esse teorema à distribuição binomial, temos que

$$\sum_{k=0}^{n} \mathrm{P}(X = k) = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = [1 + (1-p)]^n = 1$$

Assim, a equação (3.7) realmente define uma função de probabilidade. Vamos denotar por $X \sim bin(n, p)$ o fato de a v.a. X ter distribuição binomial com parâmetros n e p.

3.4.2 Esperança e Variância

No Apêndice A prova-se que

$$X \sim bin(n,p) \quad \Rightarrow \quad \begin{cases} \mathrm{E}(X) = np \\ \mathrm{Var}(X) = np(1-p) \end{cases} \tag{3.9}$$

Note que a esperança e a variância da binomial são iguais à esperança e à variância da distribuição de Bernoulli, multiplicadas por n, o número de repetições. Pode-se pensar na distribuição de Bernoulli como uma distribuição binomial com parâmetros $1, p$.

EXEMPLO 3.7 Tiro ao alvo

Um atirador acerta, na mosca do alvo, 20% dos tiros. Se ele dá 10 tiros, qual a probabilidade de ele acertar na mosca no máximo uma vez?

Solução

Podemos pensar os tiros como experimentos de Bernoulli independentes, em que o sucesso é acertar no alvo e a probabilidade de sucesso é 0,20. Então, o problema pede $\mathrm{P}(X \leq 1)$, em que $X =$ número de acertos em 10 tiros. Logo, $X \sim bin(10; 0,20)$ e

$$\begin{aligned}
\mathrm{P}(X \leq 1) &= \mathrm{P}(X=0) + \mathrm{P}(X=1) \\
&= \binom{10}{0}(0,20)^0(0,80)^{10} + \binom{10}{1}(0,20)^1(0,80)^9 \\
&= 0,37581
\end{aligned}$$

♦♦

EXEMPLO 3.8 Partidas de um jogo

Dois adversários A e B disputam uma série de oito partidas de um determinado jogo. A probabilidade de A ganhar uma partida é 0,6 e não há empate. Qual é a probabilidade de A ganhar a série?

Solução

Note que só podem ocorrer vitórias ou derrotas, o que significa que temos repetições de um experimento de Bernoulli com probabilidade 0,6 de sucesso (vitória do jogador A). Assumindo a independência das provas, se definimos $X =$ número de vitórias de A, então $X \sim bin(8; 0,6)$ e o problema pede $\mathrm{P}(X \geq 5)$, isto é, a probabilidade de A ganhar mais partidas que B.

$$\begin{aligned}
\mathrm{P}(X \geq 5) &= \mathrm{P}(X=5) + \mathrm{P}(X=6) + \mathrm{P}(X=7) + \mathrm{P}(X=8) \\
&= \binom{8}{5}(0,6)^5(0,4)^3 + \binom{8}{6}(0,6)^6(0,4)^2 + \\
&\quad + \binom{8}{7}(0,6)^7(0,4)^1 + \binom{8}{8}(0,6)^8(0,4)^0 \\
&= 0,5940864
\end{aligned}$$

♦♦

EXEMPLO 3.9

Em uma distribuição binomial, sabe-se que a média é 4,5 e a variância é 3,15. Encontre os valores dos parâmetros da distribuição.

Solução

Temos que

$$\begin{aligned} np &= 4,5 \\ np(1-p) &= 3,15 \end{aligned}$$

Substituindo a primeira equação na segunda, resulta

$$\begin{aligned} 4,5(1-p) &= 3,15 \Rightarrow \\ 1-p &= 0,7 \Rightarrow \\ p &= 0,3 \end{aligned}$$

Substituindo na primeira equação, obtemos que

$$n = 4,5/0,3 = 15.$$

♦♦

3.5 Distribuição Geométrica

3.5.1 Introdução

Considere as seguintes situações: (i) uma moeda com probabilidade p de cara é lançada até que apareça cara pela primeira vez; (ii) em uma população muito grande (pense na população mundial), p% das pessoas sofrem de uma rara doença desconhecida e portadores precisam ser encontrados para estudos clínicos. Quais são as semelhanças entre essas daus situações? No primeiro caso, matematicamente falando, poderíamos ter que fazer "infinitos" lançamentos. No segundo caso, "muito grande" pode ser uma aproximação para "infinitos". Em ambos os casos, temos repetições de um experimento de Bernoulli. No primeiro caso, as repetições certamente podem ser consideradas independentes. No segundo caso, também podemos assumir independência, desde que esqueçamos os fatores genéticos por um momento. Então, temos repetições independentes de um experimento de Bernoulli e estamos interessados no *número de repetições até a ocorrência do primeiro sucesso* (cara no caso da moeda, pessoa portadora no estudo epidemiológico). Vamos, agora, formalizar a definição de tal variável e obter a sua função de probabilidade.

3.5.2 A Distribuição Geométrica

Consideremos, então, repetições independentes de um experimento de Bernoulli com probabilidade p de sucesso. Vamos definir a variável aleatória que representa o número de repetições até a ocorrência do primeiro sucesso. Os possíveis valores dessa variável são 1 (primeiro sucesso na primeira repetição), 2 (primeiro sucesso na segunda repetição e, portanto, fracasso na primeira), 3 (primeiro sucesso na terceira repetição e, portanto, fracasso nas duas primeiras), etc. Esse é um exemplo de v.a. discreta em que o espaço amostral, enumerável, é infinito.

Para calcular a probabilidade do evento $X = k$, $k = 1, 2, 3, \ldots$, devemos notar que tal evento corresponde à ocorrência de fracassos nas $k-1$ primeiras repetições e sucesso na k-ésima repetição. Denotando por F_i e S_i a ocorrência de fracasso e sucesso na i-ésima repetição respectivamente, temos a seguinte equivalência de eventos:

$$\{X = k\} = \{F_1 \cap F_2 \cap \cdots \cap F_{k-1} \cap S_k\}$$

Como as repetições são independentes, segue que

$$\begin{aligned} \mathrm{P}(X = k) &= \mathrm{P}(F_1 \cap F_2 \cap \cdots \cap F_{k-1} \cap S_k) = \mathrm{P}(F_1) \times \mathrm{P}(F_2) \times \cdots \times \mathrm{P}(F_{k-1}) \times \mathrm{P}(S_k) = \\ &= (1-p) \times (1-p) \times \cdots \times (1-p) \times p \end{aligned}$$

ou seja,

$$\mathrm{P}(X = k) = (1-p)^{k-1} p \qquad k = 1, 2, 3, \cdots$$

DEFINIÇÃO Variável aleatória geométrica

Para repetições independentes de um experimento de Bernoulli com parâmetro p, defina a variável aleatória

$$X = \text{“número de repetições até a ocorrência do primeiro sucesso”}$$

Então, X tem **distribuição geométrica** com parâmetro p cuja função de probabilidade é dada por

$$\mathrm{P}(X = k) = (1-p)^{k-1} p \qquad k = 1, 2, 3, \cdots \tag{3.10}$$

Dizemos que X tem *distribuição geométrica com parâmetro p* (o único valor necessário para especificar completamente a fdp) e vamos representar tal fato por $X \sim Geom(p)$.

As características definidoras desse modelo são: (i) repetições de um mesmo experimento de Bernoulli, o que significa que em todas elas a probabilidade de sucesso (e, portanto, de fracasso) é a mesma e (ii) as repetições são independentes. No caso do lançamento de uma moeda essas hipóteses são bastante plausíveis mas no caso da doença a hipótese de independência pode não ser satisfeita; por exemplo, pode haver um componente de hereditariedade.

Para estudar as propriedades da variável geométrica, faremos uso de alguns resultados sobre séries geométricas que serão apresentados a Seção 3.9.

Para mostrar que (3.10) realmente define uma função de probabilidade, temos que mostrar que a soma das probabilidades, isto é, a probabilidade do espaço amostral é 1 (obviamente, $\mathrm{P}(X = k) \geq 0$).

$$\sum_{k=1}^{\infty} \mathrm{P}(X = k) = \sum_{k=1}^{\infty} (1-p)^{k-1} p = p \sum_{k=1}^{\infty} (1-p)^{k-1}.$$

Fazendo $j = k - 1$, temos que $k = 1 \Rightarrow j = 0$ e $k = \infty \Rightarrow j = \infty$. Portanto,

$$\sum_{k=1}^{\infty} \mathrm{P}(X = k) = p \sum_{j=0}^{\infty} (1-p)^{j}$$

Usando (3.22) com $r = 1 - p$, obtém-se que:

$$\sum_{k=1}^{\infty} \mathrm{P}(X = k) = p \cdot \frac{1}{1-(1-p)} = 1.$$

3.5.3 Esperança e Variância

No Apêndice A prova-se que

$$X \sim Geom(p) \quad \Rightarrow \quad \begin{cases} E\left(X\right) = \dfrac{1}{p} \\[2ex] Var\left(X\right) = \dfrac{1-p}{p} \end{cases} \tag{3.11}$$

3.6 Distribuição binomial negativa

3.6.1 Definição

Consideremos novamente repetições independentes de um experimento de Bernoulli com probabilidade p de sucesso. Vamos considerar agora uma generalização da v.a. geométrica, no seguinte sentido:

$X =$ número de repetições necessárias até a obtenção do r-ésimo sucesso, $r \geq 1$

Note que $r = 1$ corresponde à distribuição geométrica.

Para definir os possíveis valores de X, devemos notar que para ter r sucessos, são necessários no mínimo r repetições. Logo, os possíveis valores de X são $r, r+1, r+2, \ldots$. O evento $X = k$ indica que foram necessárias k repetições para obter r sucessos e, portanto,

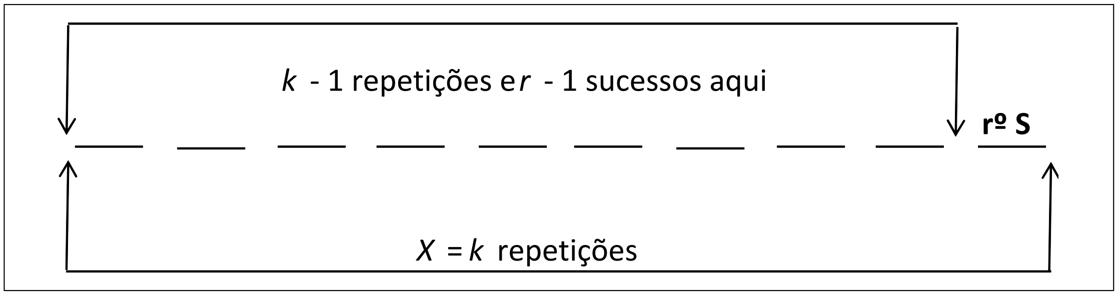

Figura 3.2 – Ilustração dos resultados da binomial negativa

$k - r$ fracassos. Pela definição da variável, a última repetição resultou em sucesso e os outros $r - 1$ sucessos podem estar em quaisquer das $k - 1$ posições restantes (ver figura 3.2).

Uma sequência possível de resultados é ter os $r - 1$ sucessos nas primeiras posições, os $k - r$ fracassos nas posições seguintes e o último sucesso na última posição: $S_1 \ldots S_{r-1} F_r \ldots F_{k-1} S_k$. A probabilidade de tal sequência é dada pelo produto das probabilidades, já que as repetições são independentes, isto é:

$$\begin{aligned}
& \mathrm{P}\left(S_1 \cap \ldots \cap S_{r-1} \cap F_r \cap \ldots F_{k-1} \cap S_k\right) = \\
= \; & \mathrm{P}\left(S_1\right) \times \cdots \times \mathrm{P}\left(S_{r-1}\right) \times \mathrm{P}\left(F_r\right) \times \cdots \times \mathrm{P}\left(F_{k-1}\right) \times \mathrm{P}\left(S_k\right) \\
= \; & p \times \cdots \times p \times (1-p) \times \cdots \times (1-p) \times p = p^r\,(1-p)^{k-r}
\end{aligned}$$

Mas existem $\binom{k-1}{r-1}$ maneiras de arrumar $r - 1$ sucessos em $k - 1$ posições e as sequências resultantes têm todas a probabilidade acima. Como elas constituem eventos mutuamente exclusivos, a probabilidade da união delas, que é $\mathrm{P}(X = k)$, é a soma das probabilidades, ou seja:

$$\mathrm{P}(X = k) = p^r\,(1-p)^{k-r} + p^r\,(1-p)^{k-r} + \cdots + p^r\,(1-p)^{k-r}$$

em que o número de parcelas é $\binom{k-1}{r-1}$. Logo

$$\mathrm{P}(X = k) = \binom{k-1}{r-1} p^r (1-p)^{k-r} \qquad k \geq r$$

DEFINIÇÃO Variável aleatória binomial negativa

Para repetições independentes de um experimento de Bernoulli com parâmetro p, defina a variável aleatória

$$X = \text{“número de repetições até a ocorrência do } r-\text{ésimo sucesso”}$$

Então, X tem **distribuição binomial negativa** com parâmetros r e p cuja função de probabilidade é dada por

$$\mathrm{P}(X = k) = \binom{k-1}{r-1} p^r (1-p)^{k-r} \qquad k \geq r \tag{3.12}$$

Essa distribuição é também conhecida como *distribuição de Pascal*. Se X tem tal distribuição, vamos representar esse fato por $X \sim BinNeg(r, p)$.

Como $\mathrm{P}(X = k) \geq 0$, para mostrar que (3.12) realmente define uma fdp, fica faltando mostrar que

$$\sum_{k=r}^{\infty} \mathrm{P}(X = k) = \sum_{k=r}^{\infty} \binom{k-1}{r-1} p^r (1-p)^{k-r} = 1.$$

Fazendo $k - r = j$, temos que $k = r \Rightarrow j = 0$ e $k = r + j$. Logo

$$\sum_{k=r}^{\infty} \mathrm{P}(X = k) = \sum_{j=0}^{\infty} \binom{r+j-1}{r-1} p^r (1-p)^j = p^r \sum_{j=0}^{\infty} \binom{r-1+j}{r-1} (1-p)^j$$

Usando o resultado dado na equação (3.24) da Seção 3.9 com $k = r - 1$ e $r = 1 - p$, obtemos que:

$$\sum_{k=r}^{\infty} \mathrm{P}(X = k) = p^r \times \frac{1}{[1-(1-p)]^{r-1+1}} = 1$$

3.7 Distribuição hipergeométrica

3.7.1 Introdução

Considere a situação do Exemplo 3.6, em que 3 bolas eram retiradas de uma urna composta por 4 bolas brancas e 6 bolas verdes. Naquele exemplo, as extrações eram feitas com reposição. Vamos supor, agora, que as extrações sejam feitas sem reposição. O que muda? A primeira observação é a de que as probabilidades em cada extração dependem das extrações anteriores, ou seja, não temos mais independência ou probabilidades constantes. A segunda observação é que cada subconjunto de 3 bolas é igualmente provável, já que as extrações são aleatorias. O número total de subconjuntos de 3 bolas retiradas das 10 bolas da urna é $\binom{10}{3}$ e, portanto, cada subconjunto tem probabilidade $\frac{1}{\binom{10}{3}}$.

Vamos considerar, novamente, a variável aleatória X que representa o número de bolas brancas extraídas. Como há 6 bolas verdes na urna, é possivel que todas as três bolas extraídas sejam verdes, isto é, $X = 0$ é o valor mínimo. No outro extremo, como há 4 brancas na urna, é possível que todas as bolas extraídas sejam brancas, ou seja, $X = 3$ é o valor máximo. É possível obter, também, 1 ou 2 bolas brancas. Logo, os valores possíveis de X são 0, 1, 2, 3. Vamos calcular a probabilidade de cada um desses valores.

- $X = 0$

  Ter 0 bola branca na amostra, significa que todas as 3 bolas são verdes. Como há 6 bolas vedes, o número de maneiras que podemos retirar 3 bolas verdes é dado por $\binom{6}{3}$. Logo,

  $$\mathrm{P}(X = 0) = \frac{\binom{6}{3}}{\binom{10}{3}}$$

- $X = 1$

  Ter 1 bola branca na amostra, significa que as outras 2 são verdes. Como há 4 bolas brancas, o número de maneiras que podemos retirar 1 bola branca é dado por $\binom{4}{1}$. Analogamente, como há 6 bolas verdes, o número de maneiras que podemos retirar 2 bolas verdes é $\binom{6}{2}$. Pelo Princípio Fundamental da Multiplicação, o número de maneiras que podemos retirar 1 bola branca *e* 2 bolas verdes é $\binom{4}{1} \times \binom{6}{2}$. Logo,

$$P(X = 1) = \frac{\binom{4}{1} \times \binom{6}{2}}{\binom{10}{3}}$$

- $X = 2$ e $X = 3$

  Analogamente,

$$P(X = 2) = \frac{\binom{4}{2} \times \binom{6}{1}}{\binom{10}{3}}$$

$$P(X = 3) = \frac{\binom{4}{3} \times \binom{6}{0}}{\binom{10}{3}}$$

Suponhamos que nossa amostra seja, agora, de 5 bolas. Como só há 4 bolas brancas, não é possivel obter uma amostra formada apenas por bolas brancas. Mas, vamos pensar, por um momento, que pudéssemos ter $X = 5$. Seguindo o raciocíonio anterior, teríamos que

$$P(X = 5) = \frac{\binom{4}{5} \times \binom{6}{0}}{\binom{10}{5}}$$

Lembre-se de que o coeficiente binomial $\binom{n}{k}$ foi definido para $n \geq k$; no entanto, se definirmos

$$\binom{n}{k} = 0 \qquad \text{se } n < k \tag{3.13}$$

isso resulta numa probabilidade nula para o evento impossível $X = 5$. Essa observação será útil na definição da distribuição hipergeométrica, apresentada a seguir.

3.7.2 A Distribuição Hipergeométrica

A generalização do contexto do exemplo anterior é a seguinte: temos uma população de tamanho N (a urna com as bolas) dividida em 2 classes (duas cores), uma composta de r "sucessos" e a outra composta de $N - r$ "fracassos". Dessa população, vamos extrair uma amostra de tamanho *n sem reposição*.

A variável aleatória de interesse é

$$X = \text{número de sucessos na amostra}$$

Valores possíveis de X

Para determinar os valores possíveis de X, temos que considerar diferentes possibilidades para a composição da população em termos dos números de sucessos e fracassos relativos ao tamanho da amostra.

- Se houver sucessos e fracassos suficientes, isto é, se $r \geq n$ e $N - r \geq n$ então os possíveis valores de X variam de 0 (amostra formada só por fracassos) a n (amostra formada só por sucessos).

- Se houver sucessos suficientes, mas não fracassos, isto é, se $r > n$ e $N - r < n$, não é possível ter uma amostra formada só por fracassos; o número máximo de fracassos na amostra é $N - r$ e, portanto, o número mínimo de sucessos é $n - (N - r)$. Logo, os valores de X variam de $n - (N - r)$ a n.

  Ilustrando: $N = 6$, $n = 3$, $r = 4$, $N - r = 2 < n$: como só há dois fracassos, a amostra tem que conter pelo menos 1 sucesso, ou seja, podemos ter $(1S, 2F)$ ou $(2S, 1F)$ ou $(3S, 0F)$.

- Se houver fracassos suficientes, mas não sucessos, isto é, se $N - r > n$ e $r < n$, não é possível ter uma amostra formada só por sucessos; o número mínimo de sucessos na amostra é 0 e o número máximo é r.

  Ilustrando: $N = 6$, $n = 3$, $r = 2 < n$, $N - r = 4 > n$: como só há dois sucessos, o número máximo possível de sucessos na amostra é 2, ou seja, podemos ter $(0S, 4F)$, $(1S, 3F)$ e $(2S, 2F)$.

Cálculo das probabilidades

O número total de amostras de tamanho n que podem ser extraídas de uma população de tamanho N, sem reposição, é $\binom{N}{n}$ (note que isso equivale ao número de subconjuntos de tamanho n do conjunto universo de tamanho N).

Para calcular a probabilidade de k sucessos na amostra, $\mathrm{P}(X = k)$, vamos considerar as 3 situações anteriores:

- Sucessos e fracassos suficientes

  Há $\binom{r}{k}$ maneiras de tirarmos k sucessos e $\binom{N-r}{n-k}$ maneiras de tirar os fracassos que completam a amostra. Logo,

$$\mathrm{P}(X = k) = \frac{\dbinom{r}{k}\dbinom{N-r}{n-k}}{\dbinom{N}{n}} \tag{3.14}$$

  e nesse caso, os valores de k vão de 0 a n.

- Sucessos suficientes, mas não fracassos

Vimos que os valores de X variam de $n-(N-r)$ a n. Para qualquer valor $k < n-(N-r)$, o coeficiente binomial $\binom{N-r}{n-k}$ será 0, pela definição 3.13. Por exemplo, se $k = n-(N-r)-1$, resulta

$$\binom{N-r}{n-k} = \binom{N-r}{n-[n-(N-r)-1]} = \binom{N-r}{N-r+1} = 0$$

Então, ainda podemos usar a Equação 3.14 para calcular as probabilidades, considerando k variando de 0 a n.

- Fracassos suficientes, mas não sucessos

  Vimos que os valores X variam de 0 a r. Para qualquer valor $k > r$, o coeficiente binomial $\binom{r}{k}$ será 0, pela definição 3.13. Então, ainda podemos usar a Equação 3.14 para calcular as probabilidades, considerando k variando de 0 a n.

Resumindo:

$$\mathrm{P}(X = k) = \frac{\binom{r}{k}\binom{N-r}{n-k}}{\binom{N}{n}} \qquad k = 0, \ldots, n \tag{3.15}$$

DEFINIÇÃO **Variável aleatória hipergeométrica**

De uma poupulação de tamanho N, formada por r sucessos e $N-r$ fracassos, extrai-se uma amostra de tamanho n sem reposição. Se a variável de interesse é

$$X = \text{número de sucessos na amostra}$$

então

$$\mathrm{P}(X = k) = \frac{\binom{r}{k}\binom{N-r}{n-k}}{\binom{N}{n}} \qquad k = 0, \ldots, n \tag{3.16}$$

onde se adota a convenção $\binom{n}{k} = 0$ se $k > n$.

Essa é a distribuição hipergeométrica com parâmetros N, r e n. Notação: $X \sim hiper(N, r, n)$.

Verificação das condições definidoras de uma função de probabilidade

No Apêndice A, prova-se que a Equação 3.16 realmente define uma função de probabilidade. Obviamente $\mathrm{P}(X = K) \geq 0$ e com base em argumentos combinatórios, prova-se que $\sum_{k=0}^{n} \mathrm{P}(X = k) = 1$.

3.7.3 Esperança e Variância

No Apêndice A prova-se que

$$X \sim hiper(N, r, n) \quad \Rightarrow \begin{cases} E(X) = n\dfrac{r}{N} \\ \\ Var(X) = n\dfrac{r}{N}\dfrac{N-r}{N}\dfrac{N-n}{N-1} \end{cases} \tag{3.17}$$

3.7.4 Distribuição binomial *versus* distribuição hipergeomtrica

Vamos fazer agora algumas comparações entre as distribuições binomial e hipergeométrica, considerando que elas descrevem a extração de amostra de tamanho n. No contexto da binomial, a amostra é retirada *com* reposição, enquanto na hipergeométrica as extrações são feitas *sem* reposição.

A esperança da binomial é igual ao produto do tamanho da amostra pela probabilidade de sucesso; na hipergeométrica, a esperança também é o produto do tamanho da amostra pela probabilidade de sucesso, probabilidade essa tomada apenas na primeira extração.

A variância da binomial é igual ao produto do tamanho da amostra pelas probabilidades de sucesso e fracasso. Na hipergeométrica, considerando apenas a primeira extração, a variância é igual a esse produto, mas corrigido pelo fator $\dfrac{N-n}{N-1}$.

Em pesquisas estatésticas por amostragem, normalmente lidamos com amostragem sem reposição. No entanto, os resultados teóricos sobre amostragem com reposição são bem mais simples, pois envolvem variáaveis independentes; assim, costuma-se usar uma aproximação, sempre que possível. Ou seja, quando a população (tamanho N) é suficientemente grande (de modo que podemos encará-la como uma população infinita) e o tamanho da amostra é relativamente pequeno, podemos "ignorar" o fato de as extrações serem feitas sem reposição. Lembre-se que a probabilidade em extrações sucessivas são $\frac{1}{N}, \frac{1}{N-1}, \ldots, \frac{1}{N-n}$. Então, se N é "grande" e n é pequeno, temos que $N \approx N-1 \approx \cdots \approx N-n$. Nessas condições, extrações com e sem reposição podem ser consideradas como equivalentes. O termo que aparece na variância da hipergeométrica, $\frac{N-n}{N-1}$, é chamado correção para populações finitas, exatamente porque, se a população é pequena, não podemos ignorar o fato de as extrações estarem sendo feitas sem reposição.

3.8 A distribuição de Poisson

3.8.1 Aproximação da binomial

Suponhamos que estejamos observando um determinado fenômeno de interesse por um certo período de tempo de comprimento t com o interesse de contar o número de vezes X que determinado evento ocorre.

Vamos fazer as seguintes hipóteses sobre a forma como esse evento ocorre:

H1) Em um intervalo de tempo suficientemente curto, apenas 0 ou 1 evento ocorre, ou seja, 2 ou mais ocorrências não podem acontecer simultaneamente. Então, em cada um desses intervalos temos um *experimento de Bernoulli*.

H2) A probabilidade de exatamente 1 ocorrência nesse pequeno intervalo de tempo, de comprimento Δt, é proporcional a esse comprimento, ou seja, é $\lambda \Delta t$. Logo, a probabilidade de nenhuma ocorrência é $1 - \lambda \Delta t$.

H3) As ocorrências em intervalos pequenos e disjuntos são experimentos de Bernoulli *independentes*.

Estamos interessados na v.a. X = número de ocorrências do evento no intervalo $(0, t]$. Particionando esse intervalo em n pequenos subintervalos de comprimento Δt, temos que o número total de ocorrências será a soma do número de ocorrências em cada subintervalo. Mas em cada subintervalo podemos aplicar as hipóteses acima. Logo, X é uma variável binomial com parâmetros $n = \dfrac{t}{\Delta t}$ (note que $\Delta t = \dfrac{t}{n}$) e probabilidade de sucesso igual a $\lambda \Delta t$ pela hipótese 2 acima. Então, para $k = 0, 1, 2, \ldots, n$ temos que:

$$\begin{aligned}
\mathrm{P}(X &= k) = \binom{n}{k} (\lambda \Delta t)^k (1 - \lambda \Delta t)^{n-k} = \binom{n}{k} \left(\frac{\lambda t}{n}\right)^k \left(1 - \frac{\lambda t}{n}\right)^{n-k} = \\
&= \frac{n!}{k!(n-k)!} \times \frac{1}{n^k} \times (\lambda t)^k \times \left(1 - \frac{\lambda t}{n}\right)^n \times \left(1 - \frac{\lambda t}{n}\right)^{-k} = \\
&= \frac{n(n-1)\cdots(n-k+1)(n-k)!}{(n-k)!} \times \frac{1}{n^k} \times \frac{(\lambda t)^k}{k!} \times \left(1 - \frac{\lambda t}{n}\right)^n \times \left(1 - \frac{\lambda t}{n}\right)^{-k} = \\
&= \frac{n(n-1)\cdots(n-k+1)}{n^k} \times \frac{(\lambda t)^k}{k!} \times \left(1 - \frac{\lambda t}{n}\right)^n \times \left(1 - \frac{\lambda t}{n}\right)^{-k} = \\
&= \frac{n}{n} \times \frac{n-1}{n} \times \cdots \times \frac{n-k+1}{n} \times \frac{(\lambda t)^k}{k!} \times \left(1 - \frac{\lambda t}{n}\right)^n \times \left(1 - \frac{\lambda t}{n}\right)^{-k}
\end{aligned}$$

Consideremos, agora, a situação em que $\Delta t \to 0$, ou equivalentemente, $n \to \infty$. Nesse caso, a v.a. X pode assumir qualquer valor inteiro não negativo e

$$\begin{aligned}
\lim_{n\to\infty} \mathrm{P}(X &= k) = \lim_{n\to\infty} \left[\frac{n}{n} \times \frac{n-1}{n} \times \cdots \times \frac{n-k+1}{n} \times \frac{(\lambda t)^k}{k!} \times \left(1 - \frac{\lambda t}{n}\right)^n \times \left(1 - \frac{\lambda t}{n}\right)^{-k} \right] \\
&= 1 \times 1 \times \cdots \times 1 \times \frac{(\lambda t)^k}{k!} \times \lim_{n\to\infty} \left(1 - \frac{\lambda t}{n}\right)^n \times 1 = \frac{(\lambda t)^k}{k!} \exp(-\lambda t)
\end{aligned}$$

Aqui fêz-se uso do resultado (3.30) da seção 3.9, que resulta em

> **! Aproximação da binomial pela Poisson**
>
> Sejam eventos gerados de acordo com as hipóteses H1 a H3 acima. Se X é o número de eventos em um intervalo de tempo de comprimento t, então a função de probabilidade de X é
>
> $$\mathrm{P}(X = k) = \frac{(\lambda t)^k}{k!} \exp(-\lambda t) \qquad k = 0, 1, 2, \ldots \tag{3.18}$$

Diz-que que X tem distribuição de Poisson com parâmetro $\lambda t : X \sim Poi(\lambda t)$.

Para mostrar que (3.18) realmente define uma função de probabilidade, temos que provar que $\sum_k \mathrm{P}(X = k) = 1$. De fato, usando o resultado (3.31) da Seção 3.9, temos que:

$$\sum_{k=0}^{\infty} \mathrm{P}(X = k) = \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} \exp(-\lambda t) = \exp(-\lambda t) \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} = \exp(-\lambda t) \exp(\lambda t) = 1$$

Esperança e variância

Vamos agora calcular a esperança e a variância de tal distribuição.

$$\begin{aligned}
E(X) &= \sum_{k=0}^{\infty} k \frac{(\lambda t)^k}{k!} \exp(-\lambda t) = \sum_{k=1}^{\infty} k \frac{(\lambda t)^k}{k(k-1)!} \exp(-\lambda t) = \\
&= \sum_{k=1}^{\infty} \frac{(\lambda t)^k}{(k-1)!} \exp(-\lambda t) = \exp(-\lambda t) \sum_{k=1}^{\infty} \frac{(\lambda t)(\lambda t)^{k-1}}{(k-1)!} = \\
&= (\lambda t) \exp(-\lambda t) \sum_{k=1}^{\infty} \frac{(\lambda t)^{k-1}}{(k-1)!} = (\lambda t) \exp(-\lambda t) \sum_{j=0}^{\infty} \frac{(\lambda t)^j}{j!} = \\
&= (\lambda t) \exp(-\lambda t) \exp(\lambda t)
\end{aligned}$$

Logo,

$$X \sim Poi(\lambda t) \Rightarrow E(X) = \lambda t \tag{3.19}$$

$$\begin{aligned}
E(X^2) &= \sum_{k=0}^{\infty} k^2 \frac{(\lambda t)^k}{k!} \exp(-\lambda t) = \sum_{k=1}^{\infty} k^2 \frac{(\lambda t)^k}{k(k-1)!} \exp(-\lambda t) = \\
&= \sum_{k=1}^{\infty} k \frac{(\lambda t)^k}{(k-1)!} \exp(-\lambda t) = \exp(-\lambda t) \sum_{k=1}^{\infty} k \frac{(\lambda t)(\lambda t)^{k-1}}{(k-1)!} = \\
&= (\lambda t) \exp(-\lambda t) \sum_{k=1}^{\infty} k \frac{(\lambda t)^{k-1}}{(k-1)!} = (\lambda t) \exp(-\lambda t) \sum_{j=0}^{\infty} (j+1) \frac{(\lambda t)^j}{j!} =
\end{aligned}$$

$$
\begin{aligned}
&= (\lambda t)\left[\sum_{j=0}^{\infty} j\frac{(\lambda t)^j}{j!}\exp(-\lambda t) + \exp(-\lambda t)\sum_{j=0}^{\infty}\frac{(\lambda t)^j}{j!}\right] = \\
&= (\lambda t)\left[E(X) + \exp(-\lambda t)\exp(\lambda t)\right] = (\lambda t)^2 + (\lambda t)
\end{aligned}
$$

Aqui fêz-se uso dos resultados (3.19) e (3.31) da seção 3.9. Logo,

$$\text{Var}(X) = (\lambda t)^2 + (\lambda t) - (\lambda t)^2$$

ou

$$X \sim Poi(\lambda t) \Rightarrow \text{Var}(X) = \lambda t \tag{3.20}$$

A interpretação desses resultados nos dá que o número médio de ocorrências do evento em um intervalo de comprimento t é λt, proporcional ao comprimento. Fazendo $t = 1$, obtém-se o número médio de ocorrências em um intervalo unitário. Note que a esperança e a variância são iguais!

3.8.2 A distribuição de Poisson

Vamos apresentar, agora, a definição geral da distribuição de Poisson, usando uma outra parametrização.

DEFINIÇÃO **Distribuição de Poisson**

Diz-se que uma variável aleatória X tem distribuição de Poisson com parâmetro λ se sua função de probabilidade é dada por

$$\text{P}(X = k) = \frac{\lambda^k}{k!}e^{-\lambda} \qquad k = 0, 1, 2, \ldots$$

Nesse caso, a esperança e a variância de X são dadas por:

$$\text{E}(X) = \text{Var}(X) = \mu \tag{3.21}$$

Pelos resultados anteriores, μ é o número médio de ocorrências do evento de interesse em um intervalo unitário e o número de ocorrências num intervalo qualquer é proporcional ao comprimento do intervalo.

3.9 Alguns resultados de cálculo

3.9.1 Séries geométricas

Recordemos inicialmente a progressão geométrica (pg) de razão r, dada por $\left(1, r, r^2, r^3, \ldots, r^n\right)$. Para calcular a soma $S_n = 1 + r + r^2 + \ldots + r^n$ dos n primeiros termos de uma pg, note que, se $r \neq 1$, podemos escrever:

$$\begin{aligned} (1-r)S_n &= (1-r)(1+r+r^2+\ldots+r^n) \\ &= (1+r+r^2+\ldots+r^n) - (r+r^2+\ldots+r^{n+1}) = 1 - r^{n+1} \end{aligned}$$

Logo,

$$(1+r+r^2+\ldots+r^n) = \frac{1-r^{n+1}}{1-r} \qquad r \neq 1$$

Além disso, se $|r| < 1$, $\lim\limits_{n\to\infty} r^{n+1} = 0$. Logo, a soma dos termos da pg converge, quando $n \to \infty$, isto é:

$$\lim_{n\to\infty} S_n = \frac{1}{1-r}$$

Mas S_n nada mais é do que a $n-$ésima soma parcial da série geométrica $\sum\limits_{i=0}^{\infty} r^i$. Resulta, assim, que a série geométrica converge, ou seja

$$\sum_{i=0}^{\infty} r^i = \frac{1}{1-r} \qquad \text{se } |\,r\,| < 1. \tag{3.22}$$

Note que podemos escrever (atenção aos índices dos somatórios!):

$$\sum_{i=0}^{\infty} r^i = 1 + \sum_{i=1}^{\infty} r^i \Rightarrow \frac{1}{1-r} = 1 + \sum_{i=1}^{\infty} r^i \quad \text{se } |\,r\,| < 1$$

ou

$$\sum_{i=1}^{\infty} r^i = 1 - \frac{1}{1-r} = \frac{r}{1-r} \qquad \text{se } |\,r\,| < 1 \tag{3.23}$$

Usando resultados sobre diferenciação de funções representadas por séries de potências, obtemos o seguinte resultado:

$$\sum_{i=0}^{\infty} \binom{k+i}{i} r^i = \sum_{i=0}^{\infty} \binom{k+i}{k} r^i = \frac{1}{(1-r)^{k+1}} \tag{3.24}$$

Mas note que

$$\binom{k+i}{k} = \frac{1 \cdot 2 \cdot \cdots \cdot i \cdot (i+1) \cdot (i+2) \cdots (i+k)}{i!k!} = \frac{(i+1) \cdot (i+2) \cdots (i+k)}{k!} \tag{3.25}$$

e, portanto

$$\sum_{i=0}^{\infty} \binom{k+i}{k} r^i = \sum_{i=0}^{\infty} \frac{(i+1)\cdot(i+2)\cdots(i+k)}{k!} r^i = \frac{1}{(1-r)^{k+1}} \tag{3.26}$$

ou equivalentemente,

$$\sum_{i=0}^{\infty} (i+1)\cdot(i+2)\cdots(i+k) r^i = \frac{k!}{(1-r)^{k+1}} \tag{3.27}$$

3.9.2 O número e (base dos logaritmos naturais)

1.
$$\lim_{n\to\infty} \left(1+\frac{1}{n}\right)^n = e \tag{3.28}$$

2.
$$\lim_{h\to 0} (1+h)^{\frac{1}{h}} = e \tag{3.29}$$

3. A partir desses resultados obtém-se que:
$$\lim_{n\to\infty} \left(1+\frac{x}{n}\right)^n = e^x \tag{3.30}$$

   De fato: fazendo $t = \frac{x}{n}$, resulta que

$$\lim_{n\to\infty} \left(1+\frac{x}{n}\right)^n = \lim_{t\to 0} (1+t)^{\frac{x}{t}} = \left[\lim_{t\to 0} (1+t)^{\frac{1}{t}}\right]^x = e^x$$

4.
$$\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{\lambda} \quad \forall \lambda \in \mathbb{R} \tag{3.31}$$

Capítulo 4

Exercícios

4.1 Enunciados

1. Cinco cartas são extraídas de um baralho comum (52 cartas, 13 de cada naipe) sem reposição. Defina a v.a. X = número de cartas vermelhas sorteadas.

   (a) Quais são os possíveis valores de X?

   (b) Encontre a fdp de X.

   (c) Calcule a esperança de X.

2. Numa urna há sete bolas brancas e quatro bolas verdes. Cinco bolas são extraídas dessa urna sem reposição. Defina a v.a. X = número de bolas verdes.

   (a) Quais são os possíveis valores de X?

   (b) Encontre a fdp de X.

   (c) calcule a esperança e a variância de X.

3. Repita o exercício anterior para o caso de extrações com reposição.

4. Uma variável aleatória discreta X tem a seguinte função de distribuição de probabilidade

$$f_X(x) = \begin{cases} \frac{k}{(x+2)!} & x = 0, 1 \\ \\ 0 & x \neq 0 \text{ e } x \neq 1 \end{cases}$$

   onde k é uma constante.

   (a) Determine o valor de k para que f_X seja uma fdp.

   (b) Calcule a função de distribuição $F_X(x)$.

   (c) Calcule a esperança e a variância de X.

5. Considere o lançamento de três moedas e denote por K a ocorrência de cara e por C a ocorrência de coroa. Se ocorre o evento CCC, dizemos que temos uma sequência, ao passo que se ocorre o evento CKC temos três sequências. Defina a v.a. X = "número de caras obtidas" e Y = "número de sequências obtidas".

(a) Obtenha as distribuições de X e Y.

(b) Calcule a esperança e a variância de X e Y.

6. Um vendedor de serviços de informática visita diariamente uma ou duas empresas, com probabilidades 0,6 e 0,4. Em cada visita, ele pode ser malsucedido e não conseguir fechar negócio com probabilidade 0,6 ou ser bem sucedido e conseguir fechar um contrato médio no valor de 5.000 reais ou um contrato grande no valor de 20.000 reais com probabilidades 0,3 e 0,1, respectivamente. Determine o valor esperado das vendas diárias desse vendedor.

7. Uma empresa de aluguel de carros tem em sua frota 4 carros de luxo e ela aluga esses carros por dia, segundo a seguinte função de distribuição de probabilidade:

No. de carros alugados/dia	0	1	2	3	4
Probabilidade de alugar	0,10	0,30	0,30	0,20	0,10

O valor do aluguel é de R$2000,00 por dia; a despesa total com manutenção é de R$500,00 por dia quando o carro é alugado e de R$200,00 por dia quando o carro não é alugado. Calcule:

(a) o número médio de carros de luxo alugados por dia, bem como o desvio padrão;

(b) a média e o desvio padrão do lucro diário com o aluguel dos carros de luxo.

8. As chamadas diárias recebidas por um batalhão do Corpo de Bombeiros apresentam a seguinte distribuição:

Número de chamadas/dia	0	1	2	3	4	5
Percentual (%) de dias	10	15	30	25	15	5

(a) Calcule o número médio de chamadas por dia, bem como o desvio padrão do número de chamadas diárias.

(b) Em um ano de 365 dias, qual é o número total de chamadas?

9. As probabilidades de que haja 1, 2, 3, 4 ou 5 pessoas em cada carro que se dirige ao Barra Shopping em um sábado são, respectivamente, 0,05; 0,20; 0,40; 0,25 e 0,10.

(a) Qual o número médio de pessoas por carro?

(b) Se chegam ao shopping 50 carros por hora, qual o número esperado de pessoas no período das 13 às 18 horas?

10. Na manufatura de certa peça, é sabido que uma entre dez peças é defeituosa. Uma amostra de tamanho quatro é retirada com reposição, de um lote da produção. Qual a probabilidade de que a amostra contenha

(a) nenhuma defeituosa?

(b) pelo menos uma defeituosa?

(c) exatamente uma defeituosa?

Na solução desse exercício, é importante que você identifique o experimento, a variável aleatória de interesse e sua respectiva fdp.

11. Um supermercado faz a seguinte promoção: o cliente, ao passar pelo caixa, lança um dado. Se sair a face 6 tem um desconto de 30% sobre o total de sua conta. Se sair a face 5, o desconto é de 20%. Se sair a face 4, o desconto é de 10% e se ocorrerem as faces 1, 2 ou 3, o desconto é de 5%. Seja X = desconto concedido.

    (a) Encontre a função de distribuição de probabilidade de X.

    (b) Calcule o desconto médio concedido.

    (c) Calcule a probabilidade de que, num grupo de cinco clientes, pelo menos um consiga um desconto maior que 10%.

    (d) Calcule a probabilidade de que o quarto cliente seja o primeiro a receber 30% de desconto.

12. Um atirador acerta na mosca do alvo 20% dos tiros.

    (a) Qual é a probabilidade de ele acertar na mosca pela primeira vez no décimo tiro?

    (b) Se ele dá 10 tiros, qual é a probabilidade de ele acertar na mosca exatamente uma vez?

4.2 Solução

1. (a) No baralho há 26 cartas vermelhas, 13 de ouros e 13 de copas. Logo, os possíveis valores de X são 0, 1, 2, 3, 4, 5.

    (b) O espaço amostral desse experimento consiste em todos os possíveis subconjuntos de 5 cartas. Como a ordem não interessa, o número de elementos do espaço amostral é $n\Omega = \binom{52}{5}$.

$$
\begin{aligned}
P(X=0) &= P(5 \text{ pretas}) = \frac{\binom{26}{5}}{\binom{52}{5}} \\
&= \frac{26 \times 25 \times 24 \times 23 \times 22}{52 \times 51 \times 50 \times 49 \times 48} \\
&= \frac{23 \times 22}{2 \times 51 \times 2 \times 49 \times 2} = 0,0253
\end{aligned}
$$

$$
\begin{aligned}
P(X=1) &= P(4 \text{ pretas}, 1 \text{ vermelha}) = \frac{\binom{26}{4}\binom{26}{1}}{\binom{52}{5}} \\
&= \frac{\dfrac{26 \times 25 \times 24 \times 23}{4 \times 3 \times 2} \times 26}{\dfrac{52 \times 51 \times 50 \times 49 \times 48}{5 \times 4 \times 3 \times 2}} = \frac{26 \times 25 \times 23 \times 26}{52 \times 51 \times 10 \times 49 \times 2} \\
&= \frac{5 \times 23 \times 13}{2 \times 51 \times 2 \times 49} = \frac{65 \times 23}{4 \times 51 \times 49} = 0,1496
\end{aligned}
$$

$$
\begin{aligned}
P(X=2) &= P(3 \text{ pretas}, 2 \text{ vermelhas }) = \frac{\binom{26}{3}\binom{26}{2}}{\binom{52}{5}} \\
&= \frac{\dfrac{26 \times 25 \times 24}{3 \times 2} \times \dfrac{26 \times 25}{2}}{\dfrac{52 \times 51 \times 50 \times 49 \times 48}{5 \times 4 \times 3 \times 2}} = \frac{26 \times 25 \times 4 \times 13 \times 25}{52 \times 51 \times 5 \times 49 \times 4} \\
&= \frac{5 \times 13 \times 25}{2 \times 51 \times 49} = \frac{65 \times 25}{2 \times 51 \times 49} = 0,3251
\end{aligned}
$$

Como o número de cartas pretas e vermelhas é o mesmo, resulta que

$$P(X=3) = P(2 \text{ pretas,}3 \text{ vermelhas }) = \frac{\binom{26}{2}\binom{26}{3}}{\binom{52}{5}} = 0,3251$$

$$P(X=4) = P(1 \text{ preta,}4 \text{ vermelhas}) = \frac{\binom{26}{1}\binom{26}{4}}{\binom{52}{5}} = 0,1496$$

$$P(X=5) = P(5 \text{ vermelhas}) = \frac{\binom{26}{5}}{\binom{52}{5}} = 0,0253$$

Logo, a fdp de X é

x	0	1	2	3	4	5
$p_X(x)$	0,0253	0,1496	0,3251	0,3251	0,1496	0,0253

(c) Analisando a fdp de X, vemos que ela é simétrica em torno do valor 2,5. Como a esperança (ou média) é o centro de gravidade da distribuição, resulta que $E(X) = 2,5$. Vamos fazer os cálculos para confirmar:

$$E(X) = 0 \times 0,0253 + 1 \times 0,1496 + 2 \times 0,3251 + 3 \times 0,3251 + 4 \times 0,1496 + 5 \times 0,0253 = 2,5$$

2. Note que temos bolas brancas em quantidade suficiente para podermos tirar todas brancas ($X = 0$), mas não temos bolas verdes suficientes para tirar todas verdes.

(a) Como há apenas 4 verdes, os valores de X são $0, 1, 2, 3, 4$.

(b) O número de elementos do espaço amostral é

$$\#\Omega = \binom{11}{5} = \frac{11 \times 10 \times 9 \times 8 \times 7}{5 \times 4 \times 3 \times 2} = 11 \times 6 \times 7 = 462$$

$$
\begin{aligned}
P(X=0) &= P(5 \text{ brancas}) = \frac{\binom{7}{5}}{\binom{11}{5}} \\
&= \frac{7}{11} \times \frac{6}{10} \times \frac{5}{9} \times \frac{4}{8} \times \frac{3}{7} = \frac{1}{22} = \frac{3}{66}
\end{aligned}
$$

$$
\begin{aligned}
P(X=1) &= P(1 \text{ verde, } 4 \text{ brancas}) = \frac{\binom{4}{1}\binom{7}{4}}{11 \times 6 \times 7} \\
&= \frac{4 \times \dfrac{7 \times 6 \times 5}{3 \times 2}}{11 \times 6 \times 7} = \frac{10}{33} = \frac{20}{66}
\end{aligned}
$$

$$\begin{aligned}
\mathrm{P}(X=2) &= \mathrm{P}(2 \text{ verdes, } 3 \text{ brancas}) = \frac{\binom{4}{2}\binom{7}{3}}{11\times 6\times 7} \\
&= \frac{\frac{4\times 3}{2}\times\frac{7\times 6\times 5}{3\times 2}}{11\times 6\times 7} = \frac{5}{11} = \frac{30}{66}
\end{aligned}$$

$$\begin{aligned}
\mathrm{P}(X=3) &= \mathrm{P}(3 \text{ verdes, } 2 \text{ brancas}) = \frac{\binom{4}{3}\binom{7}{2}}{11\times 6\times 7} \\
&= \frac{4\times\frac{7\times 6}{2}}{11\times 6\times 7} = \frac{2}{11} = \frac{12}{66}
\end{aligned}$$

$$\begin{aligned}
\mathrm{P}(X=4) &= \mathrm{P}(4 \text{ verdes, } 1 \text{ branca}) = \frac{\binom{4}{4}\binom{7}{1}}{11\times 6\times 7} \\
&= \frac{1\times 7}{11\times 6\times 7} = \frac{1}{66}
\end{aligned}$$

Logo, a fdp de X é

x	0	1	2	3	4
$p_X(x)$	$\frac{3}{66}$	$\frac{20}{66}$	$\frac{30}{66}$	$\frac{12}{66}$	$\frac{1}{66}$

(c)

$$\begin{aligned}
\mathrm{E}(X) &= 1\times\frac{20}{66}+2\times\frac{30}{66}+3\times\frac{12}{66}+4\times\frac{1}{66} = \frac{120}{66} = 1,8181\overline{81} \\
\mathrm{E}(X^2) &= 1^2\times\frac{20}{66}+2^2\times\frac{30}{66}+3^2\times\frac{12}{66}+4^2\times\frac{1}{66} = \frac{264}{66} \\
\mathrm{Var}(X) &= \frac{264}{66}-\left[\frac{120}{66}\right]^2 = \frac{264\times 66-120^2}{66^2} = 0,694214876
\end{aligned}$$

3. (a) Se as extrações são feitas com reposição, em cada extração podemos tirar bola branca ou verde. Logo, os possíveis valores de X são $0,1,2,3,4,5$.

   (b) Com reposição, sempre temos na urna 7 brancas e 4 verdes e em cada extração, temos que $\mathrm{P}(\text{branca}) = \frac{7}{11}$ e $\mathrm{P}(\text{verde}) = \frac{4}{11}$. Como as extrações são independentes, resulta que

$$\mathrm{P}(X=0) = \mathrm{P}(5 \text{ brancas}) = \left(\frac{7}{11}\right)^5 = \frac{16807}{11^5}$$

$$\begin{aligned}
\mathrm{P}(X=1) &= \mathrm{P}(1 \text{ verde, } 4 \text{ brancas}) \\
&= \binom{5}{1}\left(\frac{7}{11}\right)^4\left(\frac{4}{11}\right)^1 = \frac{48020}{161051}
\end{aligned}$$

$$\begin{aligned}
\mathrm{P}(X=2) &= \mathrm{P}(2 \text{ verdes, } 3 \text{ brancas}) \\
&= \binom{5}{2}\left(\frac{7}{11}\right)^3\left(\frac{4}{11}\right)^2 = \frac{54880}{161051}
\end{aligned}$$

$$\begin{aligned} P(X=3) &= P(3 \text{ verdes, } 2 \text{ brancas}) \\ &= \binom{5}{3}\left(\frac{7}{11}\right)^2\left(\frac{4}{11}\right)^3 = \frac{31360}{161051} \end{aligned}$$

$$\begin{aligned} P(X=4) &= P(4 \text{ verdes, } 1 \text{ branca}) \\ &= \binom{5}{4}\left(\frac{7}{11}\right)^1\left(\frac{4}{11}\right)^4 = \frac{8960}{161051} \end{aligned}$$

$$P(X=5) = P(5 \text{ verdes}) = \left(\frac{4}{11}\right)^5 = \frac{1024}{161051}$$

Logo, a fdp de X é:

x	0	1	2	3	4	5
$p_X(x)$	$\frac{16807}{161051}$	$\frac{48020}{161051}$	$\frac{54880}{161051}$	$\frac{31360}{161051}$	$\frac{8960}{161051}$	$\frac{1024}{161051}$

A razão de multiplicarmos pelos números combinatórios se deve ao fato de que a(s) bola(s) verde(s) pode(m) sair em qualquer uma das extrações.

(c)

$$\begin{aligned} E(X) &= 1 \times \frac{48020}{161051} + 2 \times \frac{54880}{161051} + 3 \times \frac{31360}{161051} + 4 \times \frac{8960}{161051} + 5 \times \frac{1024}{161051} \\ &= \frac{292820}{161051} = 1,8181\overline{81} \\ E(X^2) &= 1^2 \times \frac{48020}{161051} + 2^2 \times \frac{54880}{161051} + 3^2 \times \frac{31360}{161051} + 4^2 \times \frac{8960}{161051} + 5^2 \times \frac{1024}{161051} \\ &= \frac{718740}{161051} \\ Var(X) &= \frac{718740}{161051} - \left[\frac{292820}{161051}\right]^2 = \frac{161051 \times 718740 - 292820^2}{161051^2} = 1,15702479 \end{aligned}$$

4. (a) Os valores possíveis da v.a. são 0 e 1. Então, temos que ter

$$\begin{aligned} f_X(0) + f_X(1) &= 1 \Rightarrow \frac{k}{2!} + \frac{k}{3!} = 1 \Rightarrow \\ \frac{k}{2} + \frac{k}{6} &= 1 \Rightarrow \frac{3k+k}{6} = 1 \Rightarrow k = \frac{6}{4} = \frac{3}{2} \end{aligned}$$

Logo,

$$\begin{aligned} f_X(0) &= \frac{\frac{3}{2}}{2} = \frac{3}{4} \\ f_X(1) &= \frac{\frac{3}{2}}{6} = \frac{1}{4} \end{aligned}$$

(b) A fda de X é

$$F_X(x) = \begin{cases} 0 & \text{se } x < 0 \\ \frac{3}{4} & \text{se } 0 \leq x < 1 \\ 1 & \text{se } x \geq 1 \end{cases}$$

(c)

$$\begin{aligned} \mathrm{E}(X) &= 0 \times \frac{3}{4} + 1 \times \frac{1}{4} = \frac{1}{4} \\ \mathrm{E}(X^2) &= 0^2 \times \frac{3}{4} + 1^2 \times \frac{1}{4} = \frac{1}{4} \\ \mathrm{Var}(X) &= \frac{1}{4} - \left[\frac{1}{4}\right]^2 = \frac{3}{16} \end{aligned}$$

5. (a) Na tabela a seguir, listam-se todos os elementos do espaço amostral, bem como os valores das variáveis aleatórias.

w	$\mathrm{P}(w)$	X	Y
CCC	$\frac{1}{8}$	3	1
CCK	$\frac{1}{8}$	2	2
CKC	$\frac{1}{8}$	2	3
KCC	$\frac{1}{8}$	2	2
CKK	$\frac{1}{8}$	1	2
KCK	$\frac{1}{8}$	1	3
KKC	$\frac{1}{8}$	1	2
KKK	$\frac{1}{8}$	0	1

Logo, as fdp's de X e Y são:

x	0	1	2	3
$f_X(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

y	1	2	3
$f_Y(y)$	$\frac{2}{8}$	$\frac{4}{8}$	$\frac{2}{8}$

(b) As duas distribuições são simétricas. Logo,

$$\begin{aligned} \mathrm{E}(X) &= 1,5 = \frac{3}{2} \\ \mathrm{E}(Y) &= 2 \\ \mathrm{E}(X^2) &= 1^2 \times \frac{3}{8} + 2^2 \times \frac{3}{8} + 3^2 \times \frac{1}{8} = \frac{24}{8} = 3 \\ \mathrm{Var}(X) &= 3 - \left[\frac{3}{2}\right]^2 = \frac{3}{4} \\ \mathrm{E}(Y^2) &= 1^2 \times \frac{2}{8} + 2^2 \times \frac{4}{8} + 3^2 \times \frac{2}{8} = \frac{36}{8} = \frac{9}{2} \\ \mathrm{Var}(Y) &= \frac{9}{2} - 2^2 = \frac{1}{2} \end{aligned}$$

6. Veja a Figura 4.1 com o espaço amostral deste experimento. Daí, podemos ver que o vendedor pode

- não vender qualquer projeto,
- vender apenas um projeto médio,
- vender apenas um projeto grande,
- vender um projeto médio e um projeto grande,
- vender dois projetos médios ou
- vender dois projetos grandes.

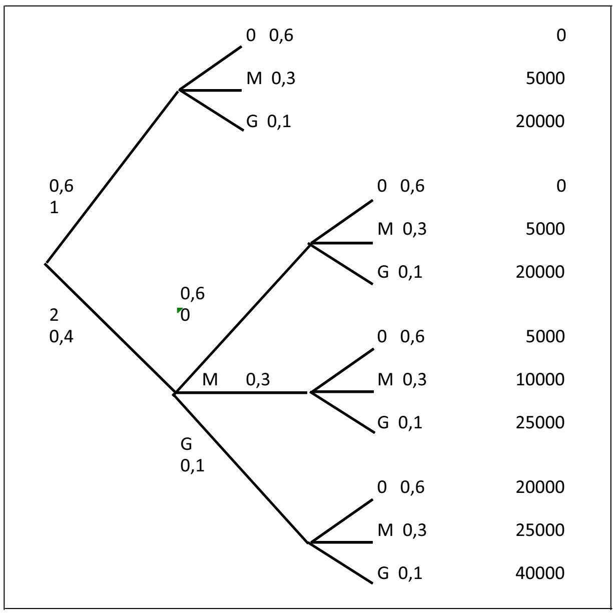

Figura 4.1 – Espaço amostral para o exercício 6

Seja V a v.a. "valor das vendas diárias". Usando a regra da multiplicação, que diz que $\mathrm{P}(A \cap B) = \mathrm{P}(A)\,\mathrm{P}(B|A)$ e o axioma da probabilidade da união de eventos mutuamente exclusivos, podemos calcular:

$$
\begin{aligned}
\mathrm{P}(V = 0) &= 0,6 \times 0,6 + 0,4 \times 0,6 \times 0,6 = 0,504 \\
\mathrm{P}(V = 5000) &= 0,6 \times 0,3 + 2 \times 0,4 \times 0,6 \times 0,3 = 0,324 \\
\mathrm{P}(V = 10000) &= 0,4 \times 0,3 \times 0,3 = 0,036 \\
\mathrm{P}(V = 20000) &= 0,6 \times 0,1 + 2 \times 0,4 \times 0,6 \times 0,1 = 0,108 \\
\mathrm{P}(V = 25000) &= 2 \times 0,4 \times 0,3 \times 0,1 = 0,024 \\
\mathrm{P}(V = 40000) &= 0,4 \times 0,1 \times 0,1 = 0,004
\end{aligned}
$$

$$
\begin{aligned}
E(V) &= 5000 \times 0.324 + 10000 \times 0.036 + 20000 \times 0.108 \\
&\quad + 25000 \times 0.024 + 40000 \times 0.004 = 4900
\end{aligned}
$$

7. Na tabela a seguir temos os resultados pertinentes para a solução do problema.

Número de carros alugados/dia	Probabilidade de alugar	Lucro por dia
0	0,10	$4 \times (-200) = -800$
1	0,30	$2000 - 500 - 3 \times 200 = 900$
2	0,30	$4000 - 2 \times 500 - 2 \times 200 = 2600$
3	0,20	$6000 - 3 \times 500 - 200 = 4300$
4	0,10	$8000 - 4 \times 500 = 6000$

Sejam X = "número de carros de luxo alugados por dia" e L = "lucro diário com aluguel de carros de luxo".

(a)

$$
\begin{aligned}
\mathrm{E}(X) &= 1 \times 0,30 + 2 \times 0,30 + 3 \times 0,20 + 4 \times 0,10 = 1,9 \text{ carros por dia} \\
\mathrm{E}(X^2) &= 1^2 \times 0,30 + 2^2 \times 0,30 + 3^2 \times 0,20 + 4^2 \times 0,10 = 4,9 \\
\mathrm{Var}(X) &= \mathrm{E}(X^2) - [\mathrm{E}(X)]^2 = 4,9 - (1,9)^2 = 1,29 \\
\mathrm{DP}(X) &= 1,1356 \text{ carros por dia}
\end{aligned}
$$

(b)

$$
\begin{aligned}
\mathrm{E}(L) &= (-800) \times 0.10 + 900 \times 0.30 + 2600 \times 0.30 + 4300 \times 0.20 + \\
&\quad +6000 \times 0.10 = 2430 \text{ reais} \\
\mathrm{E}(L^2) &= (-800)^2 \times 0.10 + 900^2 \times 0.30 + 2600^2 \times 0.30 + 4300^2 \times 0.20 \\
&\quad +6000^2 \times 0.10 = 9633000 \\
\mathrm{Var}(L) &= \mathrm{E}(L^2) - [\mathrm{E}(L)]^2 = 9633000 - 2430^2 = 3728100 \\
\mathrm{DP}(L) &= 1930,83 \text{ reais}
\end{aligned}
$$

8. (a) Seja X = "número de chamadas por dia".

$$
\begin{aligned}
\mathrm{E}(X) &= 1 \times 0,15 + 2 \times 0,30 + 3 \times 0,25 + 4 \times 0,15 + 5 \times 0,05 = 2,35 \text{ chamadas por dia} \\
\mathrm{Var}(X) &= 1 \times 0,15 + 4 \times 0,30 + 9 \times 0,25 + 16 \times 0,15 + \\
&\quad +25 \times 0,05 - (2,35)^2 = 1,7275 \\
\mathrm{DP}(X) &= 1,3143 \text{ chamadas por dia}
\end{aligned}
$$

(b) Seja T = "número de chamadas em um ano". Então, $T = 365X$ e $\mathrm{E}(T) = 2,35 \times 365 = 857,75$

9. (a) Seja X = "número de pessoas em cada carro". Então, sua fdp é dada por

x	1	2	3	4	5
p	0,05	0,20	0,40	0,25	0,10

e

$$\mathrm{E}(X) = 0,05 + 0,40 + 1,20 + 1,0 + 0,5 = 3,15 \text{ pessoas por carro}$$

(b) Seja Y = "número de pessoas em 50 carros em 5 horas de contagem". Então, $Y = 50 \times 5 \times X = 250X$ e

$$\mathrm{E}(Y) = 250\,\mathrm{E}(X) = 250 \times 3,15 = 787,5 \text{ pessoas}$$

10. Temos uma variável aleatória de Bernoulli, a saber:

$$X = \begin{cases} 1 \text{ se peça é defeituosa} \\ 0 \text{ se peça é não defeituosa} \end{cases}$$

e $P(X = 1) = 0,10$, o que implica que $P(X = 0) = 0,9$.

Seja Y = "número de peças defeituosas na amostra de tamanho 4". Como as peças são sorteadas com reposição, resulta que as extrações são independentes e a probabilidade de sucesso ("peça defeituosa") permanece constante. Logo, $Y \sim bin(4; 0,1)$ e

(a) $P(Y = 0) = \binom{4}{0}(0,10)^0(0,9)^4 = 0,6501$

(b) $P(Y \geq 1) = 1 - P(Y = 0) = 1 - 0,6501 = 0,3439$

(c) $P(Y = 1) = \binom{4}{1}(0,10)(0,9)^3 = 0,2916$

11. (a) Supondo que o dado seja honesto, a fdp de X é

Valor do desconto x	0,30	0,20	0,10	0,05
$P(X = x)$	1/6	1/6	1/6	3/6

(b) Temos que

$$E(X) = \frac{0,30 + 0,20 + 0,10 + 3 \times 0,05}{6} = 0,125$$

ou um desconto médio de 12,5%.

(c) A probabilidade de se ter um desconto maior que 10% (20% ou 30%) é de $\frac{2}{6}$. Seja Y = número de clientes, em um grupo de cinco, que recebem desconto maior que 10%. Então, $Y \sim bin\left(5; \frac{2}{6}\right)$. Logo,

$$\begin{aligned} P(Y \geq 1) &= 1 - P(Y < 1) \\ &= 1 - P(Y = 0) \\ &= 1 - \binom{5}{0}\left(\frac{2}{6}\right)^0\left(\frac{4}{6}\right)^5 = 0,868313 \end{aligned}$$

(d) Seja Z = número de clientes que passam pelo caixa até primeiro desconto de 30%. O evento $\{Z = 4\}$ corresponde a 3 clientes que não têm desconto de 30% seguidos do primeiro que tem desconto de 30%. Logo,

$$P(Z = 4) = \left(\frac{5}{6}\right)^3\left(\frac{1}{6}\right) = 0,09645$$

12. Temos uma variável aleatória de Bernoulli, a saber:

$$X = \begin{cases} 1 \text{ se acerta no alvo} \\ 0 \text{ se não acerta no alvo} \end{cases}$$

e $P(X = 1) = 0,20$, o que implica que $P(X = 0) = 0,8$.

(a) Seja Z = "número de tiros até primeiro acerto no alvo". Então, $\{Z = 10\}$ significa que o atirador erra os 9 primeiros e acerta o décimo tiro:

$$P(Z = 10) = (0,8)^9(0,20) = 0,026844$$

(b) Seja Y = "número de acertos em 10 tiros". Então, $Y \sim bin(10; 0,2)$ e

$$P(Y = 1) = \binom{10}{1}(0,20)(0,8)^9 = 0,26844$$

Apêndice A

Demonstrações de propriedades de variáveis aleatórias discretas

A.1 Distribuição Binomial

A.1.1 Esperança

$$
\begin{aligned}
\mathrm{E}(X) &= \sum_{k=0}^{n} k \Pr(X = k) = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k} = \\
&= \sum_{k=0}^{n} k \frac{n!}{k!\,(n-k)!} p^k (1-p)^{n-k}
\end{aligned}
$$

Quando $k = 0$, a parcela correspondente no somatório é nula. Logo, podemos escrever (note o índice do somatório!):

$$
\mathrm{E}(X) = \sum_{k=1}^{n} k \frac{n!}{k!\,(n-k)!} p^k (1-p)^{n-k} = \sum_{k=1}^{n} k \frac{n!}{k\,(k-1)!\,(n-k)!} p^k (1-p)^{n-k}
$$

e como $k \neq 0$, podemos fazer a divisão, o que resulta na simplificação

$$
\begin{aligned}
\mathrm{E}(X) &= \sum_{k=1}^{n} \frac{n!}{(k-1)!\,(n-k)!} p^k (1-p)^{n-k} = \sum_{k=1}^{n} \frac{n\,(n-1)!}{(k-1)!\,(n-k)!} \left(p \times p^{k-1}\right) (1-p)^{n-k} = \\
&= np \sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!\,(n-k)!} p^{k-1} (1-p)^{n-k} = np \sum_{k=1}^{n} \binom{n-1}{k-1} p^{k-1} (1-p)^{n-k}
\end{aligned}
$$

Fazendo $j = k - 1$, temos que $k = j + 1$, $k = 1 \Rightarrow j = 0$ e $k = n \Rightarrow j = n - 1$. Logo,

$$
\mathrm{E}(X) = np \sum_{j=0}^{n-1} \binom{n-1}{j} p^j (1-p)^{n-1-j}
$$

Mas note que esse somatório é a expressão do binômio de Newton para $(x+y)^{n-1}$ com $x=p$ e $y=1-p$) e, portanto, é igual a $1^{n-1}=1$. (Esse somatório é, também, a soma das probabilidades de uma variável binomial com parâmetros $n-1$ e p.)

$$X \sim bin(n,p) \quad \Rightarrow \quad E(X)=np \tag{A.1}$$

A.1.2 Variância

Vamos calcular $E\left(X^2\right)$. Usando raciocínio análogo ao usado no cálculo da esperança, temos que:

$$\begin{aligned}
E\left(X^2\right) &= \sum_{k=0}^{n} k^2 \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=1}^{n} k^2 \frac{n!}{k!\,(n-k)!} p^k (1-p)^{n-k} = \\
&= \sum_{k=1}^{n} k^2 \frac{n!}{k\,(k-1)!\,(n-k)!} p^k (1-p)^{n-k} = \\
&= \sum_{k=1}^{n} k \frac{n!}{(k-1)!\,(n-k)!} p^k (1-p)^{n-k} = \\
&= \sum_{k=1}^{n} k \frac{n\,(n-1)!}{(k-1)!\,(n-k)!} \left(p \times p^{k-1}\right) (1-p)^{n-k} = \\
&= np \sum_{k=1}^{n} k \frac{(n-1)!}{(k-1)!\,(n-k)!} p^{k-1} (1-p)^{n-k} = \\
&= np \sum_{j=0}^{n-1} (j+1) \frac{(n-1)!}{j!\,(n-j-1)!} p^j (1-p)^{n-j-1} = \\
&= np \sum_{j=0}^{n-1} j \frac{(n-1)!}{j!\,(n-j-1)!} p^j (1-p)^{n-j-1} + np \sum_{j=0}^{n-1} \frac{(n-1)!}{j!\,(n-j-1)!} p^j (1-p)^{n-j-1} =
\end{aligned}$$

$$\begin{aligned}
&= np \sum_{j=0}^{n-1} j \frac{(n-1)!}{j!\,(n-1-j)!} p^j (1-p)^{n-1-j} + np \sum_{j=0}^{n-1} \frac{(n-1)!}{j!\,(n-1-j)!} p^j (1-p)^{n-1-j} = \\
&= np \sum_{j=0}^{n-1} j \binom{n-1}{j} p^j (1-p)^{n-1-j} + np \sum_{j=0}^{n-1} \binom{n-1}{j} p^j (1-p)^{n-1-j}
\end{aligned}$$

Mas o primeiro somatório é a esperança de uma binomial com parâmetros $(n-1)$ e p; portanto, pelo resultado (A.1), é igual a $(n-1)\,p$. Já o segundo somatório é a soma das probabilidades dos valores de uma binomial com esses mesmos parâmetros (ou binômio de Newton); logo, é igual a 1. Segue, então, que

$$E\left(X^2\right) = np\left[(n-1)\,p+1\right] = n^2p^2 - np^2 + np$$

e, portanto,

$$Var\,(X) = n^2p^2 - np^2 + np - (np)^2 = np - np^2$$

ou seja,

$$X \sim bin(n,p) \quad \Rightarrow \quad Var\,(X) = np\,(1-p) \tag{A.2}$$

A.2 Distribuição geométrica

A.2.1 Esperança

Por definição, temos que

$$E(X) = \sum_{k=1}^{\infty} k\,\mathrm{P}(X=k) = \sum_{k=1}^{\infty} kp(1-p)^{k-1} = p\sum_{k=1}^{\infty} k(1-p)^{k-1}$$

Fazendo a mudança de variável $k-1=j$, resulta que $k=j+1$, $k=1 \Rightarrow j=0$ e $k=\infty \Rightarrow j=\infty$. Logo,

$$E(X) = p\sum_{j=0}^{\infty}(j+1)(1-p)^j$$

Usando o resultado (3.27) da seção 3.9 com $r=1-p$ e $k=1$, obtemos que:

$$E(X) = p \times \frac{1!}{[1-(1-p)]^{1+1}} = \frac{p}{p^2}$$

Logo,

$$X \sim Geom(p) \Rightarrow E(X) = \frac{1}{p} \tag{A.3}$$

A.2.2 Variância

Para calcular a variância, temos que calcular $E(X^2)$. Por definição,

$$\begin{aligned} E(X^2) &= \sum_{k=1}^{\infty} k^2p(1-p)^{k-1} = p\sum_{k=1}^{\infty}\left(k^2-k+k\right)(1-p)^{k-1} = \\ &= p\sum_{k=1}^{\infty}[(k^2-k)(1-p)^{k-1} + k(1-p)^{k-1}] \end{aligned}$$

Como ambos os termos convergem, podemos escrever

$$\begin{aligned} E(X^2) &= p\sum_{k=1}^{\infty}\left(k^2-k\right)(1-p)^{k-1} + p\sum_{k=1}^{\infty}k(1-p)^{k-1} \\ &= p\sum_{k=1}^{\infty}k(k-1)(1-p)^{k-1} + p\sum_{k=1}^{\infty}k(1-p)^{k-1} \end{aligned}$$

No primeiro somatório, a parcela correspondente a $k = 1$ é nula, logo, podemos escrever (note o índice do somatório!):

$$\begin{aligned} E(X^2) &= p\sum_{k=2}^{\infty} k(k-1)(1-p)^{k-2}(1-p) + \sum_{k=1}^{\infty} kp(1-p)^{k-1} = \\ &= p(1-p)\sum_{k=2}^{\infty} k(k-1)(1-p)^{k-2} + \sum_{k=1}^{\infty} kp(1-p)^{k-1} \end{aligned}$$

O segundo somatório é a esperança da distribuição geométrica com parâmetro p; logo, ele é igual a $\frac{1}{p}$. Fazendo a mudança de variável $k - 2 = j$ no primeiro somatório, resulta que

$$E(X^2) = p(1-p)\sum_{j=0}^{\infty}(j+2)(j+1)(1-p)^j + \frac{1}{p}$$

Usando o resultado (3.27) da seção 3.9 com $r = 1 - p$ e $k = 2$, obtemos que:

$$\begin{aligned} E\left(X^2\right) &= p(1-p) \times \frac{2!}{[1-(1-p)]^{2+1}} + \frac{1}{p} = \\ &= \frac{2(1-p)}{p^2} + \frac{1}{p} = \frac{2-2p+p}{p^2} = \frac{2-p}{p^2} \end{aligned}$$

Segue que:

$$Var(X) = E(X^2) - [E(X)]^2 = \frac{2-p}{p^2} - \frac{1}{p^2}$$

Logo,

$$X \sim Geom(p) \Rightarrow Var(X) = \frac{1-p}{p^2} \tag{A.4}$$

A.3 Distribuição Hipergeométrica

A.3.1 Condições definidoras de uma função de probabilidade

Vimos que, se $X \sim hiper(n, r, n)$, então

$$\mathrm{P}\left(X = k\right) = \frac{\binom{r}{k}\binom{N-r}{n-k}}{\binom{N}{n}} \qquad k = 0, \ldots, n \tag{A.5}$$

Obviamente, $\mathrm{P}(X = k \geq 0$. Provar que $\sum_k \Pr\left(X = k\right) = 1$, equivale a provar que

$$\sum_{k=0}^{n}\binom{r}{k}\binom{N-r}{n-k} = \binom{N}{n} \tag{A.6}$$

Do teorema do binômio de Newton sabemos que

$$(x+y)^n = \sum_{j=0}^{n} \binom{n}{j} x^j y^{n-j} \tag{A.7}$$

e temos também a igualdade

$$(1+x)^r (1+x)^{N-r} = (1+x)^N \tag{A.8}$$

Para provar o resultado (A.6), vamos calcular os coeficientes de x^n em ambos os termos da igualdade (A.8). Esses coeficientes têm que ser iguais!

Por (A.7), a expressão do lado direito de (A.8) é

$$(1+x)^N = \sum_{j=0}^{N} \binom{N}{j} x^{N-j}$$

e, portanto, o coeficiente de x^n é obtido fazendo $N-j=n \Rightarrow j=N-n$, ou seja, o coeficiente de x^n é (:

$$\binom{N}{N-n} = \binom{N}{n} \tag{A.9}$$

No lado esquerdo de (A.8), a potência x^n decorre da multiplicação de x^k, vindo do primeiro termo $(1+x)^r$, por x^{n-k}, vindo do segundo termo $(1+x)^{N-r}$.

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

Logo, o coeficiente de x^k é obtido fazendo $r-j=k \Rightarrow j=r-k$, ou seja, o coeficiente é

$$\binom{r}{r-k} = \binom{r}{k}$$

Analogamente, o coeficiente de x^{n-k} em $(1+x)^{N-r}$ é obtido fazendo $n-k=N-r-j \Rightarrow j=(N-r)-(n-k)$, ou seja, o coeficiente é

$$\binom{N-r}{N-r-(n-k)} = \binom{N-r}{n-k}$$

Sendo assim, os coeficientes de $x^k x^{n-k}$ são $\binom{r}{k}\binom{N-r}{n-k}$, $\quad k=0,\ldots,n$, o que implica que o coeficiente de x^n no lado esquerdo de (A.8) é

$$\sum_{k=0}^{n} \binom{r}{k}\binom{N-r}{n-k} \tag{A.10}$$

Por (A.8), os coeficientes dados em (A.9) e (A.10) têm que ser iguais. Igualando-os, obtemos o resultado desejado dado em (A.6).

A.3.2 Esperança

$$
\begin{aligned}
E(X) &= \sum_{k=0}^{n} k \frac{\binom{r}{k}\binom{N-r}{n-k}}{\binom{N}{n}} = \frac{1}{\binom{N}{n}} \sum_{k=1}^{n} k \binom{r}{k}\binom{N-r}{n-k} = && \text{(note o índice)} \\
&= \frac{1}{\binom{N}{n}} \sum_{k=1}^{n} k \frac{r(r-1)!}{k(k-1)!\,(r-k)!}\binom{N-r}{n-k} = && \text{(porque } k \neq 0\text{)} \\
&= \frac{r}{\binom{N}{n}} \sum_{k=1}^{n} \frac{(r-1)!}{(k-1)!\,(r-k)!}\binom{N-r}{n-k} = && (j = k-1) \\
&= \frac{r}{\binom{N}{n}} \sum_{j=0}^{n-1} \frac{(r-1)!}{j!\,(r-1-j)!}\binom{N-r}{n-1-j} = \\
&= \frac{r}{\binom{N}{n}} \sum_{j=0}^{n-1} \binom{r-1}{j}\binom{N-1-(r-1)}{n-1-j} = && \text{(por (A.6))} \\
&= \frac{r}{\binom{N}{n}} \binom{N-1}{n-1} = \frac{rn!\,(N-n)!}{N!}\frac{(N-1)!}{(n-1)!\,(N-n)!}
\end{aligned}
$$

Logo,

$$
X \sim hiper(N, r, n) \quad \Rightarrow \quad E(X) = n\frac{r}{N} \tag{A.11}
$$

A.3.3 Variância

Vamos calcular $E(X^2)$.

$$
\begin{aligned}
E\left(X^2\right) &= \sum_{k=0}^{n} k^2 \frac{\binom{r}{k}\binom{N-r}{n-k}}{\binom{N}{n}} = \frac{1}{\binom{N}{n}} \sum_{k=1}^{n} k^2 \binom{r}{k}\binom{N-r}{n-k} = && \text{(note o índice)} \\
&= \frac{1}{\binom{N}{n}} \sum_{k=1}^{n} k^2 \frac{r(r-1)!}{k(k-1)!\,(r-k)!}\binom{N-r}{n-k} = \\
&= \frac{r}{\binom{N}{n}} \sum_{k=1}^{n} k \frac{(r-1)!}{(k-1)!\,(r-k)!}\binom{N-r}{n-k} = && \text{(porque } k \neq 0\text{)}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{r}{\binom{N}{n}} \sum_{j=0}^{n-1} (j+1) \frac{(r-1)!}{j!\,(r-1-j)!} \binom{N-r}{n-1-j} = \qquad\qquad (j = k-1) \\
&= \frac{r}{\binom{N}{n}} \sum_{j=0}^{n-1} (j+1) \binom{r-1}{j} \binom{N-r}{n-1-j} = \\
&= \frac{r}{\binom{N}{n}} \left[\sum_{j=0}^{n-1} j \binom{r-1}{j} \binom{N-r}{n-1-j} + \sum_{j=0}^{n-1} \binom{r-1}{j} \binom{N-r}{n-1-j} \right] = \\
&= \frac{r\binom{N-1}{n-1}}{\binom{N}{n}} \left[\sum_{j=0}^{n-1} j \frac{\binom{r-1}{j}\binom{N-1-(r-1)}{n-1-j}}{\binom{N-1}{n-1}} + \sum_{j=0}^{n-1} \frac{\binom{r-1}{j}\binom{N-1-(r-1)}{n-1-j}}{\binom{N-1}{n-1}} \right]
\end{aligned}
$$

Mas o primeiro somatório é a esperança de uma hipergeométrica com parâmetros $N-1, n-1$ e $r-1$ e o segundo somatório é a soma das probabilidades no espaço amostral de uma hipergeométrica com os mesmos parâmetros. Segue, então, que

$$
\begin{aligned}
E\left(X^2\right) &= \frac{r\binom{N-1}{n-1}}{\binom{N}{n}} \left[(n-1)\frac{r-1}{N-1} + 1 \right] = \\
&= \frac{rn}{N} \times \frac{(n-1)(r-1) + N - 1}{N-1}
\end{aligned}
$$

e, portanto,

$$
\begin{aligned}
\mathrm{Var}(X) &= \frac{rn}{N} \times \frac{(n-1)(r-1) + N - 1}{N-1} - \frac{n^2r^2}{N^2} = \\
&= \frac{rn}{N} \times \left[\frac{Nnr - nN - Nr + N + N^2 - N - Nnr + nr}{N(N-1)} \right] = \\
&= n\frac{r}{N} \frac{N(N-n) - r(N-n)}{N(N-1)}
\end{aligned}
$$

ou seja,

$$
X \sim hiper(N, r, n) \quad \Rightarrow \quad Var(X) = n\frac{r}{N}\frac{N-r}{N}\frac{N-n}{N-1} \tag{A.12}
$$