Aula 12 – Algumas distribuições discretas

Nesta aula estudaremos alguns modelos de variáveis aleatórias discretas. O objetivo de tais modelos é descrever situações gerais que se encaixam no contexto definido para cada um deles. Dentre os vários modelos de variáveis aleatórias discretas, estudaremos os seguintes:

- distribuição uniforme discreta
- distribuição de Bernoulli
- distribuição binomial
- distribuição hipergeométrica

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 acima, os experimentos citados têm algo em comum: em um 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 distribuição 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.

Nesse 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.

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

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

$$f_X(x_i) = \Pr(X = x_i) = \frac{1}{n} \qquad \forall i = 1, 2, \ldots, n \tag{12.1}$$

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 $\sum_x f_X(x) = 1$.

Esperança e variância

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

$$\begin{aligned} E(X) &= \frac{1}{n}x_1 + \frac{1}{n}x_2 + \cdots + \frac{1}{n}x_n \\ &= \overline{x} \end{aligned}$$

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

Com relação à variância, temos, por definição, que

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

que é a mesma fórmula vista na parte 1 do curso para a variância populacional de um conjunto de dados.

Exemplo 12.1

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

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

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

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

Atividade 12.1

1. Os defeitos em determinada máquina ocorrem aproximadamente na mesma freqüência. 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?

2. No lançamento de um dado, define-se a v.a. X como sendo o número da face obtida. Explique qual(is) é(são) a(s) hipótese(s) necessária(s) para que a fdp de X seja uma distribuição uniforme.

Distribuição de Bernoulli

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

Definição
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

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

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

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

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

(12.2)

Obviamente, as condições definidoras de uma fdp são satisfeitas, uma vez que $p > 0$, $1 - p > 0$ e $p + (1 - p) = 1$. 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:

$$F_X(x) = \begin{cases} 0 & \text{se } x < 0 \\ 1-p & \text{se } 0 \leq x < 1 \\ 1 & \text{se } x \geq 1 \end{cases} \tag{12.3}$$

Na **Figura 12.1** temos os gráficos da fdp e da fda de uma distribuição de Bernoulli.

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:

$$X \sim Bern(p) \quad \Rightarrow \quad \begin{cases} E(X) = p \\ \\ Var(X) = p(1-p) \end{cases} \tag{12.4}$$

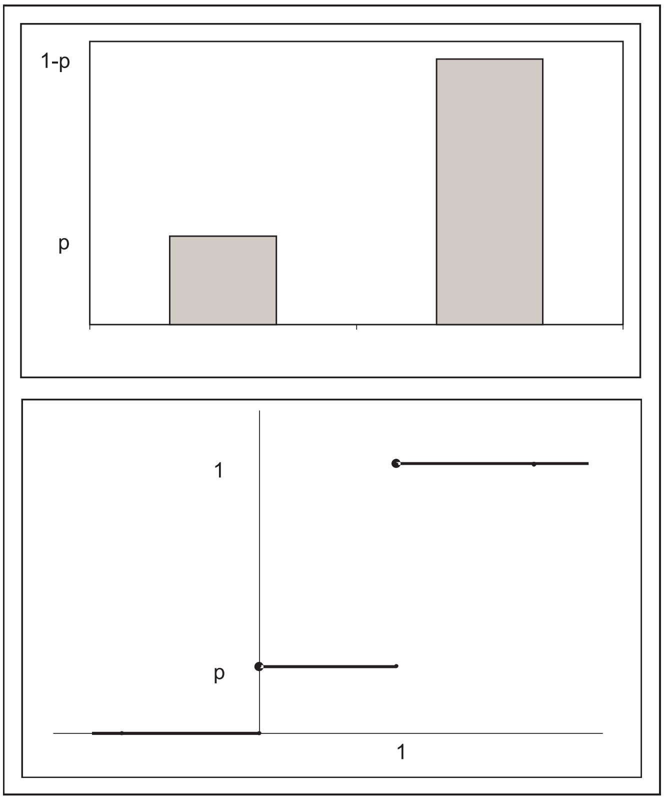

Figura 12.1: Distribuição de Bernoulli com parâmetro p.

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

Exemplo 12.2

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

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

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 12.3

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, onde 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.

Distribuição binomial

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

Exemplo 12.4

Considere o seguinte experimento: uma moeda é lançada n vezes e sabe-se que $p = \Pr(\text{cara})$. 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 distribuição de probabilidade de X, o primeiro fato a notar é que os valores possíveis de X são: 0, que equivale à ocorrência de n coroas; 1, que equivale à ocorrência de apenas 1 cara; 2, que equivale à ocorrência de 2 caras e, assim por diante, até n, que equivale à ocorrência de n caras. Assim, os possíveis valores de X são:

$$X = 0, 1, 2, \ldots, n$$

Vamos, agora, calcular a probabilidade de cada um desses valores, de modo a completar a especificação da fdp 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:

$$\{X = 0\} \equiv \{C_1 \cap C_2 \cap \cdots \cap C_n\}$$

  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

$$\begin{aligned} \Pr(C_1 \cap C_2 \cap \cdots \cap C_n) &= \Pr(C_1) \times \Pr(C_2) \times \cdots \times \Pr(C_n) \\ &= (1-p) \times (1-p) \times \cdots \times (1-p) \\ &= (1-p)^n \end{aligned}$$

- $X = 1$

  O evento $X = 1$ corresponde à ocorrência de 1 cara e, conseqüentemente, de $n-1$ coroas. Uma seqüência possível de lançamentos é $K_1 \cap C_2 \cap C_3 \cap \cdots \cap C_n$. Vamos calcular a probabilidade desse resultado. Como antes, os lançamentos são eventos independentes e, portanto,

$$\begin{aligned} \Pr(K_1 \cap C_2 \cap C_3 \cap \cdots \cap C_n) &= \Pr(K_1) \times \Pr(C_2) \times \cdots \times \Pr(C_n) \\ &= p \times (1-p) \times \cdots \times (1-p) \\ &= p(1-p)^{n-1} \end{aligned}$$

  Mas qualquer seqüência com 1 cara resulta em $X = 1$; por exemplo, o evento $C_1 \cap C_2 \cap \cdots \cap C_{n-1} \cap K_n$ também resulta em $X = 1$. Na verdade, a face cara poderia estar em qualquer posição e todas essas seqüências resultariam 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:

$$\begin{aligned} \{X = 1\} \equiv \quad & \{K_1 \cap C_2 \cap C_3 \cap \cdots \cap C_n\} \\ & \cup \{C_1 \cap K_2 \cap C_3 \cap \cdots \cap C_n\} \\ & \cup \cdots \\ & \cup \{C_1 \cap C_2 \cap \cdots \cap C_{n-1} \cap K_n\} \end{aligned}$$

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

$$
\begin{aligned}
\Pr(X=1) = \quad & \Pr(K_1 \cap C_2 \cap C_3 \cap \cdots \cap C_n) \\
& + \Pr(C_1 \cap K_2 \cap C_3 \cap \cdots \cap C_n) \\
& + \cdots \\
& + \Pr(C_1 \cap C_2 \cap \cdots \cap C_{n-1} \cap K_n) \\
= \quad & p \times (1-p) \times (1-p) \times \cdots \times (1-p) \\
& +(1-p) \times p \times (1-p) \times \cdots \times (1-p) \\
& + \cdots \\
& +(1-p) \times (1-p) \times \cdots \times (1-p) \times p \\
= \quad & p(1-p)^{n-1} + p(1-p)^{n-1} + \cdots + p(1-p)^{n-1} \\
= \quad & np(1-p)^{n-1} \\
= \quad & \binom{n}{1} p(1-p)^{n-1}
\end{aligned}
$$

- $X = 2$

O evento $X = 2$ corresponde à ocorrência de 2 caras e, conseqüentemente, de $n-2$ coroas. Uma seqüência possível de lançamentos é $K_1 \cap K_2 \cap C_3 \cap \cdots \cap C_n$ e a probabilidade de tal seqüência é $p^2(1-p)^{n-2}$. Mas as 2 caras podem ocupar quaisquer posições e existem $\binom{n}{2}$ maneiras de colocar 2 caras em uma seqüência de n lançamentos. Todas essas $\binom{n}{2}$ maneiras têm a mesma probabilidade e correspondem a eventos mutuamente exclusivos. Temos a seguinte equivalência:

$$
\begin{aligned}
\{X=2\} \equiv \quad & \{K_1 \cap K_2 \cap C_3 \cap \cdots \cap C_n\} \\
& \cup \{K_1 \cap C_2 \cap K_3 \cap \cdots \cap C_n\} \\
& \cup \cdots \\
& \cup \{C_1 \cap C_2 \cap \cdots \cap C_{n-2} \cap K_{n-1} \cap K_n\}
\end{aligned}
$$

e, portanto

$$
\begin{aligned}
\Pr(X=2) = \quad & \Pr(K_1 \cap K_2 \cap C_3 \cap \cdots \cap C_n) \\
& + \Pr(K_1 \cap C_2 \cap K_3 \cap \cdots \cap C_n) \\
& + \cdots \\
& + \Pr(C_1 \cap C_2 \cap \cdots \cap C_{n-2} \cap K_{n-1} \cap K_n) \\
= \quad & p \times p \times (1-p) \times \cdots \times (1-p) \\
& + p \times (1-p) \times p \times \cdots \times (1-p) \\
& + \cdots \\
& + (1-p) \times \cdots \times (1-p) \times p \times p \\
= \quad & p^2(1-p)^{n-2} + p^2(1-p)^{n-2} + \cdots + p^2(1-p)^{n-2} \\
= \quad & \binom{n}{2} p^2(1-p)^{n-2}
\end{aligned}
$$

Aqui vale a pena fazer uma observação sobre o número combinatório. Por que combinação e não arranjo? Vamos considerar a primeira seqüência como exemplo: $K_1 \cap K_2 \cap C_3 \cap \cdots \cap C_n$. Essa seqüência, para o nosso problema, significa "cara nos 2 primeiros lançamentos e coroa nos $n-2$ últimos lançamentos". Não existe diferença entre as seqüências $K_1 \cap K_2 \cap C_3 \cap \cdots \cap C_n$ e $K_2 \cap K_1 \cap C_3 \cap \cdots \cap C_n$. Assim, a "ordem não importa" e temos que usar combinação.

- $X = x, \qquad x = 0, 1, 2, \ldots, n$

  O raciocínio visto para os casos $X = 0, X = 1$ e $X = 2$ se generaliza facilmente, o que nos leva ao seguinte resultado geral:

$$
\Pr(X = x) = \binom{n}{x} p^x (1-p)^{n-x} \qquad x = 0, 1, 2, \ldots, n
$$

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 seqüência 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 12.5

Uma urna contém 4 bolas brancas e 6 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:

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

Cada extração equivale a um experimento de Bernoulli e como o interesse está nas bolas brancas, vamos considerar sucesso = bola branca.

Os valores possíveis de X são $0, 1, 2, 3$. 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, podemos considerar as extrações como independentes e, assim, temos uma situação análoga à do exemplo anterior: temos 4 repetições de um experimento (sorteio de uma bola), essas repetições são independentes e em cada uma delas há dois resultados possíveis: bola branca ou bola verde.

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,

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

Lembre-se que, por definição, $0! = 1$

- $X = 1$

  Esse resultado equivale à extração de uma bola branca e, por conseqüência, 2 bolas verdes. A bola branca pode sair em qualquer extração e, definida a posição da bola branca, as posições das bolas verdes ficam totalmente estabelecidas. Logo,

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

- $X = 2$

  Esse resultado equivale à extração de duas bolas brancas e, por conseqüência, 1 bola verde. As bolas brancas podem sair em quaisquer duas extrações e, definidas as posições das bolas brancas, a posição da bola verde fica totalmente estabelecida. Logo,

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

- $X = 3$

  Esse resultado equivale à extração de três bolas brancas; logo,

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

A distribuição binomial

Nos dois exemplos anteriores, tínhamos repetições de um experimento que podiam ser consideradas independentes e em cada repetição havia apenas dois resultados possíveis. Essas são as condições definidoras do *contexto binomial*.

Definição
Um **experimento binomial** consiste em repetições *independentes* de um experimento de Bernoulli.

Definição
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 distribuição de probabilidade é dada por

$$f_X(x) = \Pr(X = x) = \binom{n}{x} p^x (1-p)^{n-x} \qquad x = 0, 1, 2, \ldots, n \qquad (12.5)$$

É imediato ver, da equação (12.5), que $f_X(x) \geq 0$. O fato de que $\sum\limits_{x=0}^{n} f_X(x) = 1$ segue diretamente do Teorema do Binômio de Newton que diz que, se x e y são números reais e n é um inteiro positivo, então

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}. \qquad (12.6)$$

Fazendo $x = p$ e $y = 1 - p$ em (12.6), obtém-se:

$$[p + (1-p)]^n = 1^n = 1 = \sum_{x=0}^{n} \binom{n}{x} p^x (1-p)^{n-x} = \sum_{x=0}^{n} f_X(x)$$

o que prova que $\sum\limits_{x=0}^{n} f_X(x) = 1$. Sendo assim a equação (12.5) realmente define uma função de distribuição de probabilidade. Vamos denotar por $X \sim bin(n,p)$ o fato de que a v.a. X tem distribuição binomial com parâmetros n e p.

Esperança e variância

Pode-se mostrar que

$$X \sim bin(n,p) \quad \Rightarrow \quad \begin{cases} E\left(X\right) = np \\ \\ Var\left(X\right) = np\left(1-p\right) \end{cases} \qquad (12.7)$$

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 12.6

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 1 vez?

Solução:

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

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

Exemplo 12.7

Dois adversários A e B disputam uma série de 8 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 $\Pr(X \geq 5)$, isto é, probabilidade de A ganhar mais partidas que B.

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

Atividade 12.2

1. Na manufatura de certo artigo, é sabido que 1 entre 10 artigos é defeituoso. Uma amostra de tamanho 4 é retirada com reposição, de um lote da produção. Qual a probabilidade de que a amostra contenha

   (a) nenhum defeituoso?

   (b) pelo menos 1 defeituoso?

   (c) exatamente 1 defeituoso?

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

2. 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.

Distribuição hipergeométrica

A distribuição hipergeométrica, que estudaremos a seguir, tem estreita ligação com a distribuição binomial. Para salientar as semelhanças e as diferenças entre as duas distribuições, vamos retomar a situação do Exemplo 12.5, em que consideramos extrações de uma urna composta por 4 bolas brancas e 6 bolas verdes.

Exemplo 12.8

De uma urna composta por 4 bolas brancas e 6 bolas verdes extraem-se 3 bolas *sem* reposição. Vamos definir a seguinte variável aleatória associada a esse experimento:

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

Assim como no Exemplo 12.5, temos repetições de um experimento de Bernoulli: em cada extração, podemos tirar uma bola branca ou uma bola verde. A diferença fundamental é que essas repetições *não* são independentes, ou seja, o resultado de uma repetição afeta o resultado da próxima repetição. Vamos calcular a função de distribuição de probabilidade de X. Como antes, os valores possíveis de X são

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

Para calcular a probabilidade de cada um destes valores, vamos usar a definição clássica de probabilidade, que estabelece que a probabilidade de um evento A é

$$\Pr(A) = \frac{\#A}{\#\Omega}$$

O espaço amostral Ω deste experimento é formado por todas as triplas de bolas brancas e verdes retiradas dessa urna. O número total de elementos de Ω é

$$\#\Omega = \binom{10}{3}$$

(Como antes, a ordem não importa, ou seja, $B_1B_2V_3 \equiv B_2B_1V_3 \equiv$ bola branca nas 2 primeiras extrações e bola verde na terceira extração.) Vamos calcular a probabilidade de cada valor de X.

- $X = 0$

  Esse resultado equivale a retirar apenas bolas verdes. Como há 6 bolas verdes, o número de possibilidades é $\binom{6}{3}$ e, portanto,

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

- $X = 1$

  Esse resultado equivale a tirar 1 bola branca e 2 bolas verdes. O número de possibilidades para a bola branca é $\binom{4}{1}$ e para cada uma dessas possibilidades, existem $\binom{6}{2}$ maneiras de tirar as bolas verdes. Pelo princípio fundamental da multiplicação, o número total de maneiras de tirar 1 bola branca e 2 verdes é $\binom{4}{1} \times \binom{6}{2}$ e, portanto

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

- $X = 2$

  Com raciocínio análogo, conclui-se que

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

e

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

Exemplo 12.9

De uma urna com 4 bolas brancas e 8 bolas verdes, extraem-se 6 bolas sem reposição. Mais uma vez, vamos definir a seguinte variável aleatória associada a esse experimento:

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

Note que não temos bolas brancas suficientes para tirar uma amostra só de bolas brancas, por exemplo. Mais precisamente, os valores possíveis de X são $0, 1, 2, 3, 4$. Utilizando raciocínio análogo ao do exemplo anterior, podemos ver que

$$\Pr(X = x) = \frac{\binom{4}{x}\binom{8}{6-x}}{\binom{12}{6}} \qquad x = 0, 1, 2, 3, 4$$

Se estabelecermos a notação de que $\binom{m}{j} = 0$ sempre que $j > m$, podemos definir a fdp de X como

$$\Pr(X = x) = \frac{\binom{4}{x}\binom{8}{6-x}}{\binom{12}{6}} \qquad x = 0, \dots, 6$$

e com isso estamos atribuindo probabilidade nula aos valores impossíveis 5 e 6.

Exemplo 12.10

De uma urna com 8 bolas brancas e 4 bolas verdes, extraem-se 6 bolas sem reposição. Mais uma vez, vamos definir a seguinte variável aleatória associada a esse experimento:

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

Note que não temos bolas verdes suficientes para tirar uma amostra só de bolas verdes, por exemplo. Mais precisamente, os valores possíveis de X são $2, 3, 4, 5, 6$ e

$$\Pr(X = x) = \frac{\binom{8}{x}\binom{4}{6-x}}{\binom{12}{6}} \qquad x = 2, 3, 4, 5, 6$$

Como antes, se estabelecermos a notação de que $\binom{m}{j} = 0$ sempre que $j > m$, podemos definir a fdp de X como

$$\Pr(X = x) = \frac{\dbinom{8}{x}\dbinom{4}{6-x}}{\dbinom{12}{6}} \qquad x = 0, \ldots, 6$$

e com isso estamos atribuindo probabilidade nula aos valores impossíveis 0 e 1.

Nos três exemplos anteriores, temos a seguinte situação geral: do espaço amostral, que está dividido em duas categorias (branca ou verde), retira-se, sem reposição, uma amostra ou subconjunto. O interesse está no número de elementos, nesse subconjunto, de determinada categoria. Como no experimento de Bernoulli, a categoria de interesse será identificada por "sucesso" e a outra, por "fracasso".

Definição

Considere uma população de tamanho N dividida em 2 classes, uma composta por r "sucessos" e a outra composta por $N - r$ "fracassos". Dessa população, extrai-se uma amostra de tamanho n sem reposição (ver **Figura 12.2**). Então, a variável aleatória

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

tem **distribuição hipergeométrica** com parâmetros N, r, n cuja função de distribuição de probabilidade é

$$\Pr(X = x) = \frac{\dbinom{r}{x}\dbinom{N-r}{n-x}}{\dbinom{N}{x}} \qquad x = 0, 1, 2, \ldots, n \tag{12.8}$$

Por convenção, $\binom{j}{m} = 0$ se $j > m$.

Pode-se provar que a equação (12.8) realmente define uma função de distribuição de probabilidade; isto é, $\Pr(X = k) \geq 0$ e $\sum_k \Pr(X = k) = 1$.

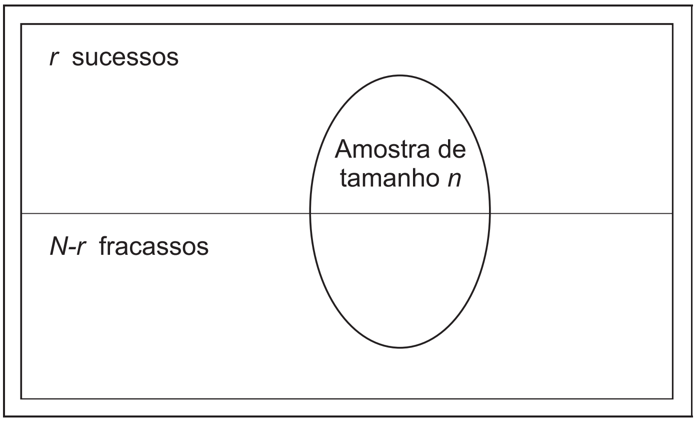

Figura 12.2: Ilustração do espaço amostral de uma v.a. hipergeométrica.

Esperança e variância

Temos os seguintes resultados:

$$X \sim hiper(N, r, n) \quad \Rightarrow \quad \begin{cases} E(X) = n\dfrac{r}{N} \\ \\ Var(X) = n\dfrac{r}{N}\dfrac{N-r}{N}\dfrac{N-n}{N-1} \end{cases} \tag{12.9}$$

Atividade 12.3

1. Uma comissão de 5 membros deve ser escolhida de um grupo formado por 12 mulheres e 18 homens. Se a comissão escolhida é formada por 5 homens, existe alguma razão para se suspeitar da lisura do processo de escolha? Suponha que seja estabelecida a seguinte regra: se a probabilidade de se obter uma comissão formada apenas por homens for muito pequena, menor que 0,01, o processo será considerado fraudulento e uma nova comissão deverá ser escolhida. Qual é a conclusão nesse caso?

2. Um caçador, após um dia de caça, verificou que matou 5 andorinhas e 2 aves de uma espécie rara, proibida de ser caçada. Como todos os espécimes tinham aproximadamente o mesmo tamanho, ele os colocou na mesma bolsa, pensando em dificultar o trabalho dos fiscais. No posto de fiscalização há dois fiscais, Manoel e Pedro, que adotam diferentes métodos de inspeção. Manoel retira três espécimes de cada bolsa

dos caçadores sem reposição. Pedro retira um espécime, classifica-o e o repõe na bolsa, retirando em seguida um segundo espécime. Em qualquer caso, o caçador é multado se é encontrado pelo menos um espécime proibido. Qual dos dois fiscais é mais favorável para o caçador em questão?

Binomial *versus* hipergeométrica

Vamos fazer agora algumas comparações entre as distribuições binomial e hipergeométrica. Colocando ambas em termos de extrações de bolas verdes de uma urna com bolas verdes e brancas, a binomial equivale a extrações independentes *com* reposição. Note que, repondo as bolas, a probabilidade de sucesso (isto é, bola branca) permanece constante ao longo das extrações. Já a hipergeométrica corresponde a extrações *sem* reposição.

A esperança da binomial é igual ao produto do tamanho da amostra pela probabilidade de sucesso; em termos da urna, a probabilidade de sucesso é $\dfrac{r}{N}$ e, portanto, a esperança é $n\dfrac{r}{N}$. 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. Em termos de urna, essas probabilidades são $\dfrac{r}{N}$ e $\dfrac{N-r}{N}$. 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 estatísticas por amostragem, normalmente lidamos com amostragem sem reposição (já imaginou visitar e entrevistar um mesmo morador duas vezes?). No entanto, os resultados teóricos sobre amostragem com reposição são bem mais simples (como você verá mais adiante nesse curso, isso equivale a lidar com variáveis independentes); assim, costuma-se usar uma aproximação, sempre que possível. Ou seja, quando o tamanho N da população 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. Isso vem dos seguintes resultados:

- na amostragem com reposição, a probabilidade de seleção de cada elemento em sorteios consecutivos é sempre $\frac{1}{N}$.
- Na amostragem sem reposição, as probabilidades em extrações suces-

sivas são $\frac{1}{N}$, $\frac{1}{N-1}$, ..., $\frac{1}{N-n}$. 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.

Resumo da Aula

- **Distribuição uniforme discreta**: X assume valores $x_1, x_2, \ldots, x_n$ tais que

$$f_X(x_i) = \Pr(X = x_i) = \frac{1}{n} \qquad \forall i = 1, 2, \ldots, n$$

$$E(X) = \overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$Var(X) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2 = \sigma_X^2$$

- **Distribuição de Bernoulli:**

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

$$\begin{aligned} E(X) &= p \\ Var(X) &= p(1-p) \end{aligned}$$

- **Experimento binomial:** repetições *independentes* de um experimento de Bernoulli.

- **Distribuição binomial:** X = número de sucessos em n repetições independentes de um experimento binomial

$$f_X(x) = \Pr(X = x) = \binom{n}{x} p^x (1-p)^{n-x} \qquad x = 0, 1, 2, \ldots, n$$

$$\begin{aligned} E(X) &= np \\ Var(X) &= np(1-p) \end{aligned}$$

- **Distribuição hipergeométrica:** X = número de sucessos em uma amostra de tamanho n, retirada sem reposição de uma população dividida em 2 classes, uma consistindo em r "sucessos" e outra consistindo em $N - r$ "fracassos"

$$\Pr(X = x) = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{x}} \qquad x = 0, 1, 2, \ldots, n$$

Por convenção, $\binom{j}{m} = 0$ se $j > m$.

Exercícios

1. Joga-se uma moeda não viciada. Qual é a probabilidade de serem obtidas 5 caras antes de 3 coroas?

2. Entre os 16 programadores de uma empresa, 12 são do sexo masculino. A empresa decide sortear 5 programadores para fazer um curso avançado de programação. Qual é a probabilidade dos 5 sorteados serem do sexo masculino?

3. **Distribuição geométrica**

   Suponha que uma moeda perfeita seja lançada até que apareça cara pela primeira vez. Obtida a primeira cara, o experimento é interrompido e conta-se o número de lançamentos feitos. Seja X o número de lançamentos. Obtenha a função de distribuição de probabilidade de X. Repita o exercício supondo que a probabilidade de cara seja p, $p \neq \frac{1}{2}$. A distribuição da v.a. X é chamada *distribuição geométrica com parâmetro p*. A definição geral da distribuição geométrica é a seguinte: Em repetições independentes de um experimento de Bernoulli com parâmetro p, a v.a. X = "número de repetições até o primeiro sucesso" tem distribuição geométrica com parâmetro p.

4. 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 1 vez?

5. A probabilidade de uma máquina produzir uma peça defeituosa em um dia é 0,1.

   (a) Qual é a probabilidade de que, em 20 peças produzidas em um dia, exatamente 1 seja defeituosa?

   (b) Qual é a probabilidade de que a $20^{\underline{a}}$ peça produzida em um dia seja a primeira defeituosa?

6. Um supermercado faz a seguinte promoção: o cliente, ao passar pelo caixa, lança um dado. Se sair face 6 tem um desconto de 30% sobre o total de sua conta. Se sair face 5 o desconto é de 20%. Se sair face 4 o desconto é de 10% e se ocorrerem 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 5 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.

7. As probabilidades de que haja 1, 2, 3, 4 ou 5 pessoas nos carros que passam por um pedágio são, respectivamente, 0,05; 0,20; 0,40; 0,25 e 0,10. Seja X = número de passageiros por veículo.

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

   (b) Calcule o número médio de passageiros por veículo.

   (c) Calcule a probabilidade de que, num grupo de 5 carros, pelo menos um tenha mais que 3 pessoas.

   (d) Calcule a probabilidade de que o quarto carro seja o primeiro a ter 5 passageiros.

8. Um fabricante de peças de automóveis garante que uma caixa de suas peças conterá, no máximo, 2 defeituosas. Se a caixa contém 18 peças e a experiência mostra que esse processo de fabricação produz 5% de peças defeituosas, qual é a probabilidade de que uma caixa satisfaça a garantia?

9. Certo curso de treinamento aumenta a produtividade de uma certa população de funcionários em 80% dos casos. Se 10 funcionários quaisquer participam deste curso, encontre a probabilidade de:

   (a) exatamente 7 funcionários aumentarem a produtividade;

   (b) pelo menos 3 funcionários não aumentarem a produtividade;

   (c) não mais que 8 funcionários aumentarem a produtividade.

10. Determinado tipo de parafuso é vendido em caixas com 1.000 peças. É uma característica da fabricação produzir 10% de defeituosos. Normalmente, cada caixa é vendida por 13,50 u.m.. Um comprador faz a seguinte proposta para o produtor: de cada caixa, ele escolhe uma amostra de 20 peças; se ele encontrar 0 defeituosa, ele paga 20,00 u.m. pela caixa; 1 ou 2 defeituosas, ele paga 10,00 u.m.; 3 ou mais defeituosas, ele paga 8,00 u.m.. Qual é a alternativa mais vantajosa para o fabricante?

11. Um industrial fabrica peças, das quais 20% são defeituosas. Dois compradores, A e B, classificam as partidas adquiridas em categorias I e II, pagando 1,20 u.m. e 0,80 u.m. respectivamente, do seguinte modo:

   - Comprador A: retira uma amostra de 5 peças; se encontrar mais de uma defeituosa, classifica como II;
   - Comprador B: retira uma amostra de 10 peças; se encontrar mais de 2 defeituosas, classifica como II.

   Em média, qual comprador oferece maior lucro para o fabricante?

Solução das Atividades

Atividade 12.1

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

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

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

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

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

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

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

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

2. O dado tem que ser honesto.

Atividade 12.2

1. Como a amostra é retirada com reposição, as extrações são repetições independentes de um experimento de Bernoulli com parâmetro 0,1. Seja X = "número de artigos defeituosos na amostra".

   (a) $\Pr(X = 0) = \binom{4}{0}(0,1)^0(0,9)^4 = 0,6561$

   (b) $\Pr(X \geq 1) = 1 - \Pr(X < 1) = 1 - \Pr(X = 0) = 0,3439$

   (c) $\Pr(X = 1) = \binom{4}{1}(0,1)^1(0,9)^3 = 0,2916$

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

Atividade 12.3

1. Vamos definir a seguinte v.a associada a este experimento:

$$X = \text{“número de homens na comissão”}$$

Queremos calcular $\Pr(X = 5)$. O número total de comissões possíveis é $\#\Omega = \binom{30}{5}$ e

$$\begin{aligned}\Pr(X &= 5) = \frac{\binom{18}{5}}{\binom{30}{5}} = \frac{\frac{18!}{5!13!}}{\frac{30!}{5!25!}} \\ &= \frac{18 \times 17 \times 16 \times 15 \times 14}{30 \times 29 \times 28 \times 27 \times 26} \\ &= 0,060124\end{aligned}$$

Como a probabilidade é maior que 0,01, não há razão para se sortear outra comissão.

2. Seja X = número de aves proibidas (sucessos) encontradas por um fiscal. No caso de Manoel, temos que $X \sim hiper(7; 2; 3)$ e no caso do fiscal Pedro, $X \sim bin\left(2; \frac{2}{7}\right)$. Queremos calcular $\Pr(\text{multa}) = \Pr(X \geq 1) = 1 - \Pr(X = 0)$.

Manoel:

$$\Pr(\text{multa}) = 1 - \Pr(X = 0) = 1 - \frac{\binom{2}{0}\binom{5}{3}}{\binom{7}{3}} = 1 - \frac{2}{7} = \frac{5}{7} = \frac{35}{49}$$

Pedro:

$$\Pr(\text{multa}) = 1 - \Pr(X = 0) = 1 - \binom{2}{0}\left(\frac{5}{7}\right)^2 = 1 - \frac{25}{49} = \frac{24}{49}$$

Logo, a probabilidade de multa é maior no caso do fiscal Manoel, e, portanto, Pedro é o fiscal mais favorável para o caçador.

Solução dos Exercícios

1. Vamos considerar a seguinte v.a. de Bernoulli

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

Então, $\Pr(X = 0) = \Pr(X = 1) = 0,5$ e temos repetições independentes de um experimento de Bernoulli. A ocorrência de 5 caras antes de 3 coroas só é possível se, nas 7 primeiras repetições, tivermos pelo menos 5 caras. Seja, então, $Y =$ "número de caras em 7 repetições". Logo, $Y \sim bin(7; 0,5)$ e o problema pede $\Pr(Y \geq 5)$.

$$\begin{aligned}
\Pr(Y \geq 5) &= \Pr(Y = 5) + \Pr(Y = 6) + \Pr(Y = 7) \\
&= \tbinom{7}{5}(0,5)^5(0,5)^2 + \tbinom{7}{6}(0,5)^6(0,5) + \tbinom{7}{7}(0,5)^7(0,5)^0 \\
&= \tbinom{7}{5}(0,5)^7 + \tbinom{7}{6}(0,5)^7 + \tbinom{7}{7}(0,5)^7 \\
&= 0,2265625
\end{aligned}$$

2. Se $X =$ número de homens sorteados, então $X \sim hiper(16; 12; 5)$ e o problema pede

$$\Pr(X = 5) = \frac{\binom{12}{5}}{\binom{16}{5}} = \frac{12 \times 11 \times 10 \times 9 \times 8}{16 \times 15 \times 14 \times 13 \times 12} = \frac{33}{14 \times 13} = 0,181319$$

3. A primeira observação diz respeito aos valores possíveis de X. Podemos ter muita sorte e obter cara no primeiro lançamento; nesse caso, $X = 1$. Nossa "sorte" pode começar a diminuir de modo que obtemos cara no segundo lançamento; nesse caso, $X = 2$. Continuando, podemos ser bastante infelizes e ter que ficar jogando a moeda "infinitas" vezes até obter a primeira cara. Esse é um exemplo de v.a. discreta em que o espaço amostral é enumerável mas infinito: os valores possíveis de X são $1, 2, 3, \ldots$. Cada resultado desses significa que os primeiros lançamentos foram coroa (C) e o último, cara (K). Como os lançamentos podem ser considerados independentes, resulta que:

$$\Pr(X = 1) = \Pr(K) = \frac{1}{2}$$

$$\begin{aligned}
\Pr(X = 2) &= \Pr(C_1 \cap K_2) \\
&= \frac{1}{2} \times \frac{1}{2} = \left(\frac{1}{2}\right)^2
\end{aligned}$$

$$\begin{aligned}
\Pr(X = 3) &= \Pr(C_1 \cap C_2 \cap K_3) \\
&= \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{4} \times \frac{1}{2} = \frac{1}{8} = \left(\frac{1}{2}\right)^3
\end{aligned}$$

$$\begin{aligned}
\Pr(X = 4) &= \Pr(C_1 \cap C_2 \cap C_3 \cap K_4) \\
&= \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8} \times \frac{1}{2} \\
&= \frac{1}{16} = \left(\frac{1}{2}\right)^4
\end{aligned}$$

En geral,

$$\Pr(X = x) = \left(\frac{1}{2}\right)^{x-1}\left(\frac{1}{2}\right) = \left(\frac{1}{2}\right)^{x} \qquad x = 1, 2, 3, \ldots$$

Se a probabilidade de cara é p, então a única diferença com relação ao visto anteriormente é que $\Pr(K) = p$ e $\Pr(C) = 1 - p$. Então,

$$\Pr(X = x) = (1-p)^{x-1} p \qquad x = 1, 2, 3, \ldots$$

É interessante notar que, tanto na distribuição binomial quanto na geométrica, temos repetições independentes de um experimento de Bernoulli. Na binomial, o número de repetições é fixo e estamos interessados no número de sucessos. Na geométrica, o número de sucessos é fixo (igual a 1) e estamos interessados no número de repetições. A distribuição binomial negativa generaliza a distribuição geométrica, no seguinte sentido: a v.a. de interesse é $X =$ "número de sucessos até o $r-$ésimo sucesso, $r \geq 1$".

4. Nossa variável aleatória de Bernoulli é a seguinte:

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

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

(a) Seja $Z =$ "número de tiros até primeiro acerto no alvo"; então, $Z \sim geom(0,2)$ e

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

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

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

5. Nossa variável aleatória de Bernoulli é a seguinte:

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

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

(a) Seja Y = "número de peças defeituosas na amostra de tamanho 20". Então $Y \sim bin(20; 0,1)$ e

$$\Pr(Y = 1) = \binom{20}{1}(0,10)(0,9)^{19} = 0,27017$$

(b) Seja Z = "número de repetições até primeira peça defeituosa"; então, $Z \sim geom(0,1)$ e

$$\Pr(Z = 20) = (0,9)^{19}(0,10) = 0,013509$$

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

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

(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, num grupo de 5, que recebem desconto maior que 10%. Então, $Y \sim bin\left(5; \frac{2}{6}\right)$. Logo,

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

(d) Seja Z = número de clientes que passam pelo caixa até primeiro desconto de 30% (probabilidade $\frac{1}{6}$). Então $Z \sim geom\left(\frac{1}{6}\right)$ e, portanto,

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

7. X = "número de pessoas em cada carro"

(a) A fdp de X é

x	1	2	3	4	5
$f_X(x) = \Pr(X = x)$	$0,05$	$0,20$	$0,40$	$0,25$	$0,10$

(b) $E(X) = 0,05 + 0,40 + 1,20 + 1,0 + 0,5 = 3,15$ pessoas por carro

(c) A probabilidade de haver mais de 3 pessoas em um carro é $0,35 = \Pr(X=4)+\Pr(X=5) = 0,25+0,10$. Seja $Y =$ número de carros, num grupo de 5, com mais de 3 pessoas. Então, $Y \sim bin\,(5;0,35)$. Logo

$$\Pr(Y \geq 1) = 1-\Pr(Y=0) = 1-\binom{5}{0}(0,35)^0(0,65)^5 = 0,883971$$

(d) Seja $Z =$ número de carros até primeiro carro com 5 passageiros. Então, $Z \sim geom\,(0,10)$ e, assim

$$\Pr(Z=4) = (0,90)^3(0,10) = 0,0729$$

8. Se $X =$ "número de peças defeituosas em uma caixa", resulta que $X \sim bin\,(18;0,05)$.

A caixa satisfaz a garantia se $X \leq 2$. Logo, a probabilidade de uma caixa satisfazer a garantia é

$$\begin{aligned}
\Pr(X \leq 2) &= \Pr(X=0)+\Pr(X=1)+\Pr(X=2) = \\
&= \binom{18}{0}(0,05)^0(0,95)^{18} + \binom{18}{1}(0,05)^1(0,95)^{17} \\
&\quad + \binom{18}{2}(0,05)^2(0,95)^{16} = \\
&= 0,397214+0,376308+0,168348 \\
&= 0,941871
\end{aligned}$$

9. Podemos pensar nos funcionários selecionados para o curso como experimentos de Bernoulli (aumenta ou não a produtividade) independentes. Seja $X =$ número de funcionários, dentre os 10, que aumentam produtividade.

(a)
$$\Pr(X=7) = \binom{10}{7}(0,80)^7(0,20)^3 = 0,201327$$

(b) Pelo menos 3 não aumentarem a produtividade é equivalente a no máximo 7 dos 10 aumentarem a produtividade. Logo, a probabilidade pedida é

$$\begin{aligned}
\Pr(X \leq 7) &= 1-\Pr(X>7) = 1-\Pr(X=8)-\Pr(X=9) \\
&\quad -\Pr(X=10) \\
&= 1-\binom{10}{8}(0,80)^8(0,20)^2-\binom{10}{9}(0,80)^9(0,20)^1 \\
&\quad -\binom{10}{10}(0,80)^{10}(0,20)^0 \\
&= 0,32220
\end{aligned}$$

(c)

$$\begin{aligned}
\Pr(X \leq 8) &= \Pr(X \leq 7) + \Pr(X = 8) \\
&= 0,322200 + \binom{10}{8} (0,80)^8 (0,20)^2 = \\
&= 0,62419
\end{aligned}$$

10. Numa população de 1.000, retirar uma amostra de 20 pode ser vista como repetições de experimentos independentes de Bernoulli.

    Seja X = número de defeituosos na amostra de 20. Então, $X \sim bin(20; 0,10)$

    Seja V = valor de compra proposto pelo cliente. Então, V pode assumir os valores 20, 10 ou 8 u.m. e, pela regra dada,

$$\Pr(V = 20) = \Pr(X = 0) = \binom{20}{0} (0,10)^0 (0,90)^{20} = 0,1216$$

$$\begin{aligned}
\Pr(V = 10) &= \Pr(X = 1) + \Pr(X = 2) = \\
&= \binom{20}{1} (0,10)^1 (0,90)^{19} + \binom{20}{2} (0,10)^2 (0,90)^{18} = 0,5553
\end{aligned}$$

$$\Pr(V = 8) = \Pr(X \geq 3) = 1 - \Pr(X = 0) - \Pr(X = 1) - \Pr(X = 2) = 0,3231$$

v	8	10	20
$f_V(v)$	0,3231	0,5553	0,1216

$$E(V) = 8 \times 0,3231 + 10 \times 0,5553 + 20 \times 0,1216 = 10,5698$$

    A proposta do cliente é mais desvantajosa para o fabricante, já que, em média, ele paga menos do que o preço normal de 13,50.

11. Sejam os seguintes eventos: A = comprador A classifica partida como tipo II e B = comprador B classifica partida como tipo II. Sejam X_A número de peças defeituosas na amostra do comprador A e X_B o número de peças defeituosas na amostra do comprador B. Então, $X_A \sim bin(5; 0,20)$ e $X_B \sim bin(10; 0,20)$

$$\begin{aligned}
\Pr(A) &= \Pr(X_A > 1) = 1 - \Pr(X_A \leq 1) \\
&= 1 - \binom{5}{0} (0,2)^0 (0,8)^5 - \binom{5}{1} (0,2)^1 (0,8)^4 = \\
&= 0,2627
\end{aligned}$$

$$\begin{aligned}
\Pr(B) &= \Pr(X_B > 2) = 1 - \Pr(X_A \leq 2) = \\
&= 1 - \binom{10}{0}(0,2)^0(0,8)^{10} - \binom{10}{1}(0,2)^1(0,8)^9 - \\
&\quad \binom{10}{2}(0,2)^2(0,8)^8 = \\
&= 0,3222
\end{aligned}$$

Sejam P_A e P_B os preços pagos pelos compradores A e B respectivamente. Então, as distribuições de probabilidade dessas variáveis são:

P_A	0,8	1,2
Probabilidade	0,2627	0,7373

$E(P_A) = 1,095$

P_B	0,8	1,2
Probabilidade	0,3222	0,6778

$E(P_B) = 1,071$

A proposta do comprador A é mais vantajosa.