Introdução à teoria das probabilidades

Variáveis aleatórias

Francisco A. Rodrigues

Universidade de São Paulo

Variáveis aleatórias

- Para estudarmos um sistema físico, precisamos definir as quantidades de interesse: posição, velocidade, temperatura, voltagem, etc.

- Para estudarmos um sistema aleatório, precisamos definir algo similar: número de caras, número de sucessos em n experimentos, nota de um aluno, valor de imóvel, etc.

- Para isso, vamos definir o conceito de variável aleatória.

Variáveis aleatórias

- Suponha que lançamos três moedas.
- Definirmos o evento C para representar uma cara observada e R para uma coroa.

Variáveis aleatórias

Suponha que lançamos três moedas. Se definirmos o evento C para representar uma cara observada e R para uma coroa, teremos os seguintes resultados possíveis:

$$\begin{aligned}\Omega =& \{(C, C, C), (C, C, R), (C, R, C), (R, C, C), \\ & (R, C, R), (R, R, C), (C, R, R), (R, R, R)\}.\end{aligned}$$

A partir do espaço amostral e dos eventos, podemos calcular diversas quantidades como o número de caras, o número de saídas em que as faces são diferentes, ou mesmo o lucro médio se cada cara resulta em um ganho de 10 reais. Essas quantidades podem ser calculadas a partir de variáveis aleatórias.

Variáveis aleatórias

Se definirmos uma função que retorna o número de caras e assumirmos que cada cara sai com probabilidade p, podemos calcular essas quantidades usando a tabela.

Saída	Número de caras	Probabilidade
(C,C,C)	3	p^3
(C,C,R)	2	$p^2(1-p)$
(C,R,C)	2	$p^2(1-p)$
(R,C,C)	2	$p^2(1-p)$
(R,C,R)	1	$p(1-p)^2$
(R,R,C)	1	$p(1-p)^2$
(C,R,R)	1	$p(1-p)^2$
(R,R,R)	0	$(1-p)^3$

Exemplo

Definição: Uma variável aleatória é uma função que mapeia o espaço amostral na reta real (X: $\Omega \longrightarrow \mathbb{R}$), onde que cada elemento ω do espaço amostral é mapeado em um valor real, isto é, X: $\omega \in \Omega \longrightarrow X(\omega) = x \in \mathbb{R}$.

Variáveis aleatórias

Exemplo: Suponha que lancemos duas moedas. O espaço amostral associado ao experimento, sendo os eventos C : "sai uma cara" e R : "sai uma coroa", é dado por:

$$\Omega = \{CC, CR, RC, RR\}.$$

Uma possível variável aleatória associada ao experimento é definida por:

$$X : \text{"número de caras obtidas no experimento"}.$$

Variáveis aleatórias

Exemplo: Suponha que lancemos duas moedas. O espaço amostral associado ao experimento, sendo os eventos C : "sai uma cara" e R : "sai uma coroa", é dado por:

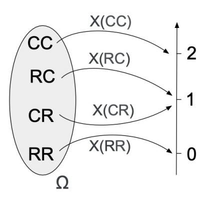
$$\Omega = \{CC, CR, RC, RR\}.$$

Uma possível variável aleatória associada ao experimento é definida por:

$$X : \text{"número de caras obtidas no experimento"}.$$

Então, a partir da definição, temos:

$$X(CC) = 2,\ X(CR) = 1,\ X(RC) = 1 \text{ e } X(RR) = 0.$$



Variáveis aleatórias

Representamos variáveis aleatórias por letras maiúsculas, enquanto que usamos letras minúsculas para indicar os valores das variáveis aleatórias. Usamos X, Y, Z, W para variáveis aleatórias, enquanto que x, y, z, w representam valores observados das variáveis aleatórias. Por exemplo, se medirmos a altura dos alunos de uma universidade, teremos:

- X : variável aleatória representando as alturas,
- x : valor da altura de um aluno selecionado.

Variáveis aleatórias discretas

Definição: Seja X uma variável aleatória. Se o número de valores possíveis de X for finito ou infinito enumerável, denominamos X como uma *variável aleatória discreta*.

Variáveis aleatórias discretas

Definição: A função que atribui a cada valor da variável aleatória a sua respectiva probabilidade é denominada função discreta de probabilidade, ou simplesmente, distribuição de probabilidade. A notação a ser utilizada é:

$$P(X = x_i) = p(x_i) = p_i, \quad i = 1, 2, 3, \ldots, \tag{1}$$

onde x_i, $i = 1, 2, \ldots$, são os valores possíveis da variável aleatória X.
A distribuição de probabilidade também é chamada função massa de probabilidade.

Variáveis aleatórias discretas

Definição: A função que atribui a cada valor da variável aleatória a sua respectiva probabilidade é denominada função discreta de probabilidade, ou simplesmente, distribuição de probabilidade. A notação a ser utilizada é:

$$P(X = x_i) = p(x_i) = p_i, \quad i = 1, 2, 3, \ldots, \tag{2}$$

onde x_i, $i = 1, 2, \ldots$, são os valores possíveis da variável aleatória X.
A distribuição de probabilidade também é chamada função massa de probabilidade. Podemos representar essa distribuição através de uma tabela:

X	x_1	x_2	x_3	$\ldots$
$P(X = x_i)$	p_1	p_2	p_3	$\ldots$

Variáveis aleatórias discretas

Exemplo: Duas bolas são retiradas sucessivamente, sem reposição, de uma caixa que contém 10 bolas vermelhas e 5 azuis. Seja a variável aleatória:

$$X : \text{“número de bolas vermelhas retiradas no experimento”}.$$

Determine a distribuição de probabilidade de X.

Variáveis aleatórias discretas

Solução: Sejam os eventos:

$$\begin{cases} V_i\text{: “uma bola vermelha retirada na tentantiva i”} \\ A_i\text{: “uma bola azul é retirada na tentativa i”} \end{cases}$$

A probabilidade associada a cada saída possível do experimento é calculada a seguir. Para duas bolas azuis retiradas:

$$P(X = 0) = P((A_1, A_2)) = \frac{5}{15} \times \frac{4}{14} = \frac{20}{210} = \frac{2}{21}.$$

Portanto, temos a seguinte tabela:

Saída	$X = x$	$P(X = x)$
(V_1, A_2)	1	$10/15 \times 5/14 = 5/21$
(A_1, V_2)	1	$5/15 \times 10/14 = 5/21$
(V_1, V_2)	2	$10/15 \times 9/14 = 9/21$
(A_1, A_2)	0	$5/15 \times 4/14 = 2/21$

Variáveis aleatórias discretas

Com isso, construímos a distribuição de probabilidade de X, que é dada por:

X	0	1	2
$P(X = x_i)$	$2/21$	$10/21$	$9/21$

Notem que $\sum_{k=0}^{2} P(X = k) = 1$.

Variáveis aleatórias discretas

Exemplo: Em um cassino da cidade, um novo jogo de dados é apresentado. Nesse jogo, um jogador paga 10 reais para participar. O jogador e a banca lançam cada um o seu dado e a seguinte regra de premiação é estabelecida:

- Se o ponto do jogador é maior, ele ganha 3 vezes a diferença entre o seu ponto e o obtido pelo oponente.
- Se o ponto do jogador é menor ou igual ao do oponente, ele não ganha nada.

Esse jogo é favorável ao jogador?

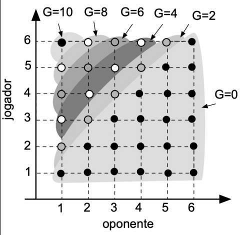
Variáveis aleatórias discretas

Solução: Vamos definir as variáveis aleatórias,

- G : "ganho bruto do jogador",
- X : "valor obtido pelo oponente"
- Y : "valor obtido pelo jogador"

Considerando a regra de premiação, podemos escrever a variável aleatória G em função de X e Y:

$$G = \begin{cases} 3(Y - X), & \text{se } Y > X \\ 0 & \text{se } Y \leq X. \end{cases} \tag{3}$$



Variáveis aleatórias discretas

A distribuição de probabilidade de G é dada por:

G	0	3	6	9	12	15
$P(G = g_i)$	$21/36$	$5/36$	$4/36$	$3/36$	$2/36$	$1/36$

Portanto, para ter lucro é preciso que $P(G \geq 10)$, pois o jogador apostou 10 reais,

$$\begin{aligned} P(G \geq 10) &= P(G = 12) + P(G = 15) \\ &= \frac{2}{36} + \frac{1}{36} = \frac{3}{36} = \frac{1}{12} = 0,08. \end{aligned}$$

Assim, a chance de lucro é muito pequena. Logo, não vale a pena jogar esse jogo.

Variáveis aleatórias discretas

Simulando o problema:

```
from random import randint
aposta = 10
Ns = 100 # número de simulações (jogos)
p = 0 # frequência que obtém lucro
for n in range(1,Ns):
    G = 0 # ganho do jogador
    j = randint(1, 6) # valor do jogador
    b = randint(1, 6) # valor do oponente
    if(j > b): # se o valor do jogador é maior
        G = 3*(j-b)
    if(G > aposta): # se houver lucro
        p = p + 1
p = p/Ns # probabilidade de ter lucro
print('Aposta:', aposta, 'reais.')
print("Probabilidade de obter lucro:", p)
```

Variáveis aleatórias contínuas

Definição: Dizemos que X é uma variável aleatória contínua se existir uma função f, denominada função de densidade de probabilidade de X, que satisfaça:

$$f(x) \geq 0, \quad \forall x \in \mathbb{R} \tag{4}$$

$$\int_{-\infty}^{\infty} f(x)dx = 1 \tag{5}$$

$$P(a \leq X \leq b) = \int_{a}^{b} f(x)dx, \quad -\infty < a < b < \infty. \tag{6}$$

Notem que $f(x)$ é uma função com valores estritamente positivos e área unitária.

Variáveis aleatórias contínuas

Exemplo: Seja X uma variável aleatória contínua com função de densidade de probabilidade:

$$f(x) = \begin{cases} 2x, & 0 < x < 1, \\ 0, & \text{caso contrário} \end{cases}$$

Verifique se $f(x)$ é uma função de densidade de probabilidade e calcule a probabilidade $P(X \geq 1/2)$.

Variáveis aleatórias contínuas

Solução: Para que $f(x)$ seja uma função de densidade de probabilidade, temos que:

$$f(x) \geq 0, \quad \forall x \in \mathbb{R}.$$

Além disso,

$$\begin{aligned}
\int_{-\infty}^{\infty} f(x)dx &= \int_{-\infty}^{0} f(x)dx + \int_{0}^{1} f(x)dx + \int_{1}^{\infty} f(x)dx \\
&= \int_{0}^{1} 2xdx \\
&= \frac{2x^2}{2}\Big|_0^1 \\
&= 1^2 - 0^2 \\
&= 1,
\end{aligned}$$

Vamos calcular $P(X \geq 1/2)$:

$$P(X \geq 1/2) = \int_{1/2}^{\infty} f(x)dx = \int_{1/2}^{1} 2xdx = \frac{2x^2}{2}\Big|_{1/2}^{1} = 1 - \frac{1}{4} = \frac{3}{4}.$$

Variáveis aleatórias contínuas

Exemplo: Seja X uma variável aleatória contínua com função de densidade de probabilidade:

$$f(x) = \begin{cases} C(4x - 2x^2), & 0 \leq x \leq 2, \\ 0 & \text{caso contrário.} \end{cases}$$

Determine o valor da constante C e calcule $P(X > 1)$.

Variáveis aleatórias contínuas

Solução: Como $f(x)$ é uma função de densidade de probabilidade, temos:

$$\int_{-\infty}^{\infty} f(x)dx = 1 \implies \int_0^1 C(4x - 2x^2)dx = 1$$

$$C\left[\frac{4x^2}{2} - \frac{2x^3}{3}\right]\Big|_{x=0}^{2} = 1 \implies C = \frac{3}{8}.$$

Podemos calcular $P(X > 1)$:

$$P(X > 1) = \int_1^2 \frac{3}{8}(4x - 2x^2)dx = \frac{3}{8}\left(\frac{4x^2}{2} - \frac{2x^3}{3}\right)\Big|_1^2 = \frac{1}{2}.$$

Probabilidade condicional

Definição: Seja X uma variável aleatória discreta ou contínua. A probabilidade condicional de que $X \in S$ dado que $X \in V$ é dada por:

$$P(X \in S | X \in V) = \frac{P(X \in S \cap X \in V)}{P(X \in V)}, \tag{7}$$

onde S e V são regiões da reta real.

Variáveis aleatórias contínuas

Exemplo: Seja X uma variável aleatória contínua com função de densidade de probabilidade:

$$f(x) = \begin{cases} 2x, & 0 < x < 1, \\ 0, & \text{caso contrário} \end{cases}$$

Calcule $P(X > 1/2 | X > 1/4)$.

Variáveis aleatórias contínuas

Solução: Usando a definição de probabilidade condicional, temos:

$$P(X > 1/2 | X > 1/4) = \frac{P(X > 1/2, X > 1/4)}{P(X > 1/4)}.$$

O termo $P(X > 1/2, X > 1/4)$ indica a probabilidade de que $X > 1/2$ e $X > 1/4$, ou seja, é a probabilidade da interseção desses dois intervalos:

$$\{X > 1/2 \cap X > 1/4\} \equiv \{X > 1/2\}.$$

Assim,

$$P(X > 1/2 | X > 1/4) = \frac{P(X > 1/2)}{P(X > 1/4)} = \frac{\int_{1/2}^{1} 2x dx}{\int_{1/4}^{1} 2x dx} = \frac{3/4}{15/16} = \frac{4}{5} = 0,8.$$

Sumário

- Variáveis aleatórias discretas.
- Variáveis aleatórias contínuas.
- Probabilidade condicional.