

VARIÁVEIS ALEATÓRIAS

- é uma função definida no espaço de amostras.
- usada para modelar quantidades desconhecidas em análises estatísticas
- uma variável aleatória X é usada para calcular a probabilidade de que X ocorra em um conjunto S

Definição 1 (Variável Aleatória): Seja S um espaço de amostras de um experimento. Uma função real que seja definida em S é chamada de variável aleatória.

Exemplo: Considere um experimento onde são jogadas 10 moedas. Temos 2^{10} possibilidades de combinações de caras/coroas. Dentre estas possibilidades, podemos definir a variável aleatória X como sendo o número de caras em uma jogada. Portanto, em cada jogada, X assumirá qualquer valor dentre 0, 1, 2, 3, ..., 10.

Distribuição de uma Variável Aleatória

Quando se mede uma quantidade em um espaço de amostras de um experimento, podemos calcular as probabilidades dos diferentes possíveis valores que X pode assumir.

Definição 2 (Distribuição): Seja S um espaço de amostras de um experimento e X uma variável aleatória neste espaço e C um subespaço de S tal que $\{X \in C\}$. A distribuição de X é a coleção de todas as probabilidades na forma $\Pr(X \in C)$ para todos os conjuntos de números reais tal que $\{X \in C\}$ seja um evento.

Exemplo 1

Considere o caso de uma jogada de 10 moedas. Se X for a variável aleatória relacionada com o número de caras, calcule $\Pr(X = x)$, onde $x = 0, 1, \dots, 10$.

Como se supõe que a moeda seja justa, temos que X será dado pela probabilidade de termos x caras, dentro de um universo de 2^{10} possibilidades. Assim:

$$\Pr(X = x) = \binom{10}{x} \frac{1}{2^{10}}, \text{ para } x = 0, 1, \dots, 10.$$

Para os dados acima temos:

$\Pr(X = 0) = 0.0009765625$	$\Pr(X = 6) = 0.205078125$
$\Pr(X = 1) = 0.0097656250$	$\Pr(X = 7) = 0.117187500$
$\Pr(X = 2) = 0.043945312$	$\Pr(X = 8) = 0.04394531$
$\Pr(X = 3) = 0.117187500$	$\Pr(X = 9) = 0.009765625$
$\Pr(X = 4) = 0.20507812$	$\Pr(X = 10) = 0.0009765625$
$\Pr(X = 5) = 0.246093750$	

DISTRIBUIÇÕES DISCRETAS

São aquelas onde X só pode assumir valores discretos.

Definição 3 (Função Probabilidade): Se uma variável aleatória X é discreta, a função probabilidade f.p. f correspondente é definida para todo número real x , tal que:

$$f(x) = \Pr(X = x).$$

Temos a seguinte restrição no conjunto de x : $x : f(x) > 0$.

Propriedades de f

Definição 4: Seja X é uma variável aleatória discreta com f.p. f . Se x não for um dos possíveis valores de X , então $f(x) = 0$. Adicionalmente, se a sequência x_1, x_2, \dots incluir todos os possíveis valores de x , então:

$$\sum_{i=1}^{\infty} f(x_i) = 1.$$

Definição 5: Seja X é uma variável aleatória discreta com f.p. f em um espaço de amostras S . Se C for um subconjunto de S , então:

$$\Pr(X \in C) = \sum_{x_i \in C} f(x_i).$$

Exemplo 2

Um carregamento de 20 notebooks de uma loja contém 3 defeituosos. Uma escola realiza uma compra de 2 computadores destes. Encontre a f.p. que descreve o número de computadores defeituosos dentro deste conjunto de 2 computadores comprados.

Se X for a p.f. procurada, temos que $X = 0, 1, 2$. As probabilidades de cada um destes casos é:

$$\begin{aligned}\Pr(X = 0) &= \frac{\binom{17}{2}}{\binom{20}{2}} = \frac{136}{190} \\ \Pr(X = 1) &= \frac{3 \times \binom{17}{1}}{\binom{20}{2}} = \frac{51}{190} \\ \Pr(X = 2) &= \frac{\binom{3}{2}}{\binom{20}{2}} = \frac{3}{190}\end{aligned}$$

Assim, a p.f. procurada é:

x	0	1	2
$f(x)$	$\frac{136}{190}$	$\frac{51}{190}$	$\frac{3}{190}$

Definição 7 (Distribuição de Bernoulli): Uma distribuição discreta é dita distribuição de Bernoulli com p quando sua f.p. obedece:

$$f(x) = \begin{cases} 1 - p, & \text{para } x = 0 \\ p, & \text{para } x = 1 \\ 0, & \text{caso contrário.} \end{cases}$$

Uma forma de expressar a função probabilidade para o ensaio de Bernoulli, em uma única expressão, é utilizarmos:

$$f(x) = p^x (1 - p)^{1-x}$$

Distribuição Binomial

Considere o caso em que uma fábrica produz determinado produto. A probabilidade de que um produto seja defeituoso é p e, portanto, a probabilidade de que o mesmo produto não seja defeituoso é dada por $(1 - p)$. Podemos calcular a probabilidade $P(m)$ de que se pegarmos 6 produtos, escolhidos aleatoriamente, m , com $m \leq 6$ destes sejam defeituosos. Como o fato de que cada produto seja ou não defeituoso não influencia a probabilidade de que qualquer outro produto também o seja, então podemos tratar os eventuais defeitos como sendo independentes entre si. Digamos, por suposição, que m seja 2 e que os 2 primeiros produtos sejam defeituosos: a probabilidade P' disto ocorrer é igual à probabilidade dos dois produtos serem defeituosos e dos outros $6 - 2 = 4$ produtos não serem defeituosos:

$$P' = p \times p \times (1 - p) \times (1 - p) \times (1 - p) \times (1 - p) = p^2 \times (1 - p)^4.$$

Entretanto, qualquer combinação dos dois produtos defeituosos, nas 6 possíveis posições, também corresponderá ao fato de que temos 2 produtos defeituosos. Assim:

$$P = \binom{6}{2} \times p^2 \times (1 - p)^4.$$

Podemos generalizar aqui. Se n é o número de peças que pegamos, m é o número de peças defeituosas e p é a probabilidade de que uma peça seja defeituosa, temos que a probabilidade de pegarmos m peças defeituosas em n , sendo que p é a probabilidade de que uma peça individual seja defeituosa, é dada por:

$$\Pr(n; p; m) = \binom{n}{m} \times p^m \times (1 - p)^{(n-m)}.$$

Exemplo 3

Calcule a probabilidade de que no lançamento de 100 vezes de uma moeda justa, tenhamos 10, 30 e 50 caras, respectivamente.

DISTRIBUIÇÕES DISCRETAS ESPECIAIS

Algumas distribuições discretas são tão comuns que recebem nomes especiais. A seguir algumas serão descritas.

Distribuição Uniforme

Definição 6 (Distribuição Uniforme de Inteiros): Sejam a e b dois inteiros tais que $a \leq b$. Suponha que X represente uma variável aleatória tal que X seja igual para qualquer valor a, \dots, b . Neste caso se diz que X é uma distribuição uniforme dos inteiros a, \dots, b .

Exemplo: A distribuição X de que um número x seja sorteado na loteria.

Teorema 1: Seja X uma variável aleatória que represente uma f.p. uniforme, a p.f. de X é

$$f(x) = \begin{cases} \frac{1}{b - a + 1}, & \text{para } x = a, \dots, b \\ 0, & \text{caso contrário.} \end{cases}$$

Ensaio de Bernoulli

Se dá o nome de ensaio de Bernoulli o tipo de experimentação onde são possíveis apenas dois resultados. Um deles é chamado de "sucesso" e o outro de "fracasso". Os termos aqui apenas estão relacionados com a expectativa de um resultado: se o resultado for o esperado, tivemos um sucesso; caso contrário temos um fracasso. Se a probabilidade de sucesso em determinado experimento for p , então a probabilidade de fracasso será $1 - p$. Em geral o sucesso é representado pelo valor $x = 1$, enquanto que o fracasso é representado pelo número $x = 0$. Portanto $\Pr(X = 1) = p$ e $\Pr(X = 0) = 1 - p$.

Podemos identificar que $p = 0,5$, $n = 100$ e $m = \{10, 30, 50\}$. Assim:

$$\Pr(100; 0,5, 10) = \binom{100}{10} \times (0,5)^{10} \times (1 - 0,5)^{90} \\ \sim 1.365\,542 \times 10^{-17}$$

$$\Pr(100; 0,5, 30) = \binom{100}{30} \times (0,5)^{30} \times (1 - 0,5)^{70} \\ \sim 2.317\,069 \times 10^{-5}$$

$$\Pr(100; 0,5, 50) = \binom{100}{50} \times (0,5)^{50} \times (1 - 0,5)^{50} \\ \sim 0.079\,589\,237\,38$$

Definição 8 (Distribuição Binomial): Uma distribuição discreta é dita distribuição binomial com parâmetros n e p quando sua f.p. obedece:

$$f(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{(n-x)}, & \text{para } x = 0, \dots, n \\ 0, & \text{caso contrário.} \end{cases}$$

DISTRIBUIÇÕES CONTÍNUAS

Ocorre quanto f é associada com uma função contínua.

Definição 9 (Variável de Distribuição Contínua): Uma variável X tem uma distribuição contínua se existe uma função f positiva, definida no eixo real, de forma que em cada intervalo real (fechado ou não) a probabilidade de que X seja igual a um valor dentro deste intervalo seja dada pela integral de f neste intervalo.

Como consequência:

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

De maneira similar:

$$\Pr(X < a) = \int_{-\infty}^a f(x) dx,$$

e

$$\Pr(X > a) = \int_a^{\infty} f(x) dx.$$

Definição 10 (Função Densidade de Probabilidade / f.d.p.): Caso X seja uma variável aleatória X , que tem distribuição contínua, então a respectiva função f , que obedece a definição anterior, é chamada de função densidade de probabilidade f.d.p.

Propriedades

Cada f.d.p. obedece duas propriedades:

- é sempre positiva:

$$f(x) \geq 0, \text{ para qualquer } x.$$

- área total igual a 1:

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

Distribuições Contínuas Uniformes por Intervalos

Definição 11: Sejam a e b dois números reais, tal que $a < b$. Se X for uma variável aleatória tal que $a \leq X \leq b$ e para cada subintervalo de $[a, b]$, a probabilidade de que X pertença a este intervalo é proporcional ao tamanho do intervalo, então se diz que X tem uma distribuição uniforme no intervalo $[a, b]$.

Este é o caso que ocorre quando dizemos que “um ponto p foi escolhido aleatoriamente no intervalo $[a, b]$ ”.

Teorema 2 (f.d.p da Distribuição Uniforme): Caso X seja uma variável aleatória em uma distribuição uniforme no intervalo $[a, b]$, então a f.d.p de X é:

$$f(x) = \begin{cases} \frac{1}{b-a}, & \text{para } a \leq X \leq b \\ 0, & \text{caso contrário.} \end{cases}$$

Prova: Se a f.d.p. contém valores não nulos somente dentro do intervalo $[a, b]$, então:

$$\Pr(a \leq X) = \int_{-\infty}^{\infty} f(x) dx = \int_a^b f(x) dx = 1.$$

Como f é uma constante (digamos que esta constante seja k), então:

$$k \int_a^b f(x) dx = 1 \Rightarrow k x|_a^b.$$

Isolando k temos $k = \frac{1}{b-a}$.

Exemplo 4

Suponha que o erro na temperatura de uma reação em uma experiência de laboratório, em $^{\circ}\text{C}$, seja dado por uma variável aleatória X tendo a f.d.p:

$$f(x) = \begin{cases} \frac{x^2}{3}, & \text{para } -1 \leq x \leq 2, \\ 0, & \text{caso contrário.} \end{cases}$$

a) Verifique que $f(x)$ é realmente uma f.d.p. e b) Encontre $\Pr(0 < X \leq 1)$.

a) Temos:

$$\int_{-1}^2 \frac{x^2}{3} dx = \frac{1}{3} \frac{x^3}{3} = \frac{1}{9} [2^3 - (-1)^3] = 1.$$

b) Temos:

$$\Pr(0 < X \leq 1) = \int_0^1 \frac{x^2}{3} dx = \frac{1}{3} \frac{x^3}{3} = \frac{1}{9} [1^3 - (0)^3] = \frac{1}{9}$$

Exemplo 5

O número de horas totais, medidos em unidades de 100 horas, que uma família utiliza um aspirador de pó no período de um ano é dado pela variável X que tem f.d.:

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

Encontre a probabilidade, no período de um ano, de que o aspirador de pó seja utilizado a) menos de 120 horas e b) entre 50 e 100 horas.

a) Para calcularmos esta probabilidade de que sejam utilizadas menos do que 120 horas (ou seja, 1,2 em unidades de centenas de horas) utilizamos diretamente a equação correspondente:

$$\begin{aligned} \Pr(X < 1,20) &= \int_0^{1,2} f(x)dx \\ &= \int_0^1 f(x)dx + \int_1^{1,2} f(x)dx \\ &= \int_0^1 xdx + \int_1^{1,2} (2-x)dx \\ &= \left. \frac{x^2}{2} \right|_0^1 + \left[2x - \frac{x^2}{2} \right]_1^{1,2} \\ &= \frac{1}{2} + \left[2,4 - \frac{(1,2)^2}{2} - 2 + \frac{1}{2} \right] \\ &= \frac{1}{2} + \frac{9}{50} = \frac{34}{50} \\ &= 0,680 \end{aligned}$$

b) Neste caso temos, diretamente:

$$\begin{aligned} \Pr(0,5 \leq X \leq 1,0) &= \int_{0,5}^{1,0} f(x)dx \\ &= \int_{0,5}^1 f(x)dx \\ &= \int_{0,5}^1 xdx \\ &= \left. \frac{x^2}{2} \right|_{0,5}^1 \\ &= \frac{1}{2} - \left[\frac{(0,5)^2}{2} \right] \\ &= \frac{1}{2} - \frac{1}{8} = \frac{3}{8} \\ &= 0,375 \end{aligned}$$

Nota: Densidade não é probabilidade Por causa da normalização, pode ser que existam pontos em que $f(x) > 1$, se o intervalo $[a, b]$ for menor do que a unidade.

FUNÇÃO DISTRIBUIÇÃO CUMULATIVA

Definição 12 (Função Distribuição Cumulativa): Uma função distribuição cumulativa f.d.c F de uma variável aleatória X é a função:

$$F(x) = \Pr(X \leq x) \text{ para } -\infty < x < \infty.$$

Isto vale para qualquer distribuição, seja ela discreta, contínua ou mista.

Exemplo 6

Construa a função densidade de probabilidade cumulativa para a variável aleatória X , tal que x seja o número de “caras” encontrados no lançamento de duas moedas justas.

Conforme podemos ver facilmente, no lançamento de duas moedas, a probabilidade de termos x “caras” é dada por:

$$\Pr(X = 0) = 1/4$$

$$\Pr(X = 1) = 1/2$$

$$\Pr(X = 2) = 1/4$$

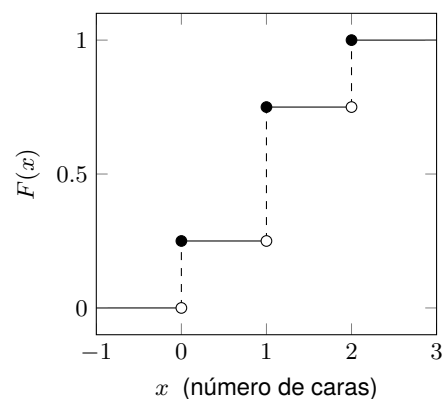
Como esta é uma distribuição discreta, temos que a função probabilidade $f(x)$ é igual ao valor de $\Pr(X = x)$. Assim a função densidade de probabilidade cumulativa fica sendo:

$$F(X = 0) = \Pr(X \leq 0) = 1/4$$

$$F(X = 1) = \Pr(X \leq 1) = 1/4 + 1/2 = 3/4$$

$$F(X = 2) = \Pr(X \leq 2) = 3/4 + 1/4 = 1$$

Podemos visualizar esta função facilmente:



Propriedades

Uma f.d.c obedece as seguintes propriedades:

1. a função $F(x)$ é crescente à medida que x aumenta, isto é, se $x_2 > x_1$, então $F(x_2) > F(x_1)$.

Prova: Se $x_1 > x_2$, então o evento $\{X \leq x_1\}$ é um subconjunto do evento $\{X \leq x_2\}$ e, portanto $\Pr(X \leq x_1) < \Pr(X \leq x_2)$.

2. nos limites $\lim_{x \rightarrow -\infty} F(x) = 0$ e $\lim_{x \rightarrow \infty} F(x) = 1$.

3. uma f.d.c é sempre contínua pela direita, isto é, $F(x) = F(x^+)$ em qualquer ponto x .

Algumas propriedades secundárias:

1. Para qualquer x :

$$\Pr(X > x) = 1 - F(x).$$

Prova: Como os eventos $\{X \leq x\}$ e $\{X > x\}$ são mutuamente excludentes, e sua união é o conjunto S , então $\Pr(\{X \leq x\}) + \Pr(\{X > x\}) = 1$, o que completa a prova.

2. Para todos os valores de x_1 e x_2 tais que $x_1 < x_2$:

$$\Pr(x_1 < X \leq x_2) = F(x_2) - F(x_1)$$

Prova: Sejam os eventos $A = \{x_1 < X \leq x_2\}$, $B = \{X \leq x_1\}$ e $C = \{X \leq x_2\}$, e vendo que os eventos A e B são mutuamente excludentes, temos que $C = A \cup B$, o que resulta em:

$$\Pr(C) = \Pr(A) + \Pr(B) \rightarrow \Pr(A) = \Pr(C) - \Pr(B).$$

3. Para cada valor x :

$$\Pr(X < x) = F(x^-)$$

O sinal x^- indica que devemos pegar o limite na função $F(x)$ pela esquerda no valor x .

4. Para cada x :

$$\Pr(X = x) = F(x) - F(x^-)$$

F.D.C de Distribuição Discreta

A função apresenta saltos nos valores que X pode assumir. Cada salto é de tamanho $f(x_i)$. Neste caso:

$$F(x) = \Pr(X \leq x) = \sum_{t \leq x} f(t), \text{ para } -\infty < x < \infty$$

Exemplo 7

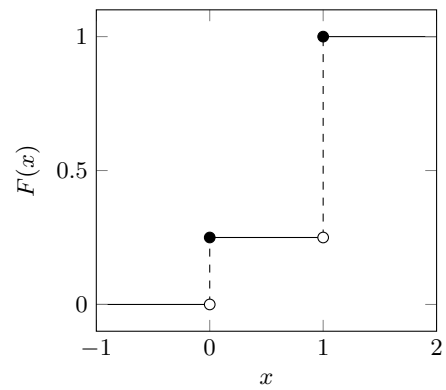
Considere a distribuição de Bernoulli. Encontre $F(x)$.

Neste caso, temos $\Pr(X = 0) = 1 - p$ e $\Pr(X = 1) = p$.

Daí segue que:

$$F(x) = \begin{cases} 0, & \text{para } x < 0, \\ 1 - p, & \text{para } 0 \leq x < 1, \\ 1, & \text{para } x \geq 1. \end{cases}$$

Graficamente ($p = 0.25$):



F.D.C de Distribuição Contínua

F é contínua. Neste caso:

$$F(x) = \Pr(X \leq x) = \int_{-\infty}^x f(t) dt, \text{ para } -\infty < x < \infty$$

Temos também, como consequência, que:

$$\frac{dF(x)}{dx} = f(x)$$

Exemplo 8

Para o exemplo anterior, do erro na temperatura de uma reação, encontre $F(x)$ e, a partir dela, calcule a probabilidade do erro na temperatura X estar entre $0 < X < 1$.

Primeiramente, temos que $F(x)$ é dada por:

$$F(x) = \int_{-\infty}^x f(x) dx = \int_{-1}^x \frac{x^2}{3} dx$$

que resulta em $F(x) = x^3/9$ para $-1 \leq x \leq 2$. A probabilidade buscada é:

$$\begin{aligned} \Pr(0 \leq x \leq 1) &= F(1) - F(0) \\ &= \frac{1}{9} - \frac{0}{9} = \frac{1}{9}, \end{aligned}$$

que é o mesmo resultado já obtido.