

Variável Aleatória

Prof. Leandro Chaves Rêgo

Departamento de Estatística e Matemática Aplicada - UFC

Fortaleza, 17 de novembro de 2021

Variável Aleatória

Suponha que um experimento aleatório consista em selecionar ao acaso um perfil em uma rede social. Quantos amigos na rede social possui o dono deste perfil? Quantidades desse tipo é o que tradicionalmente têm sido chamadas de *variáveis aleatórias*. Intuitivamente, são variáveis aleatórias porque seus valores variam dependendo do perfil escolhido. O adjetivo “aleatória” é usado para enfatizar que o seu valor é de certo modo incerto, pois a escolha do perfil envolve incerteza.

Formalmente, contudo, uma variável aleatória não é nem “aleatória” nem é uma variável. Na verdade, variável aleatória é uma função real do espaço amostral. Quando o interesse é trabalhar com o espaço amostral sendo o conjunto dos números reais, em geral, o interesse é saber sobre as probabilidades dos intervalos.

Definição

Seja (Ω, \mathcal{A}, P) um espaço de probabilidade. Uma função $X : \Omega \rightarrow \mathbb{R}$ é chamada de variável aleatória se para todo número real λ , $\{\omega \in \Omega : X(\omega) \leq \lambda\} \in \mathcal{A}$.

Esquema Funcional

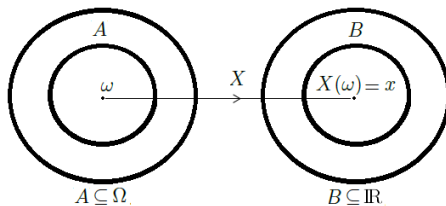


Figura: Esquema funcional de uma variável aleatória.

Exemplo

Exemplo

Considere três lançamentos de uma moeda honesta. O espaço amostral para este experimento aleatório consiste de todas as possíveis sequências de tamanho 3 de caras e coroas, isto é:

$$\Omega = \{(cara, cara, cara), (cara, cara, coroa), (cara, coroa, cara), \\ (cara, coroa, coroa), (coroa, cara, cara), (coroa, cara, coroa), \\ (coroa, coroa, cara), (coroa, coroa, coroa)\}.$$

Seja \mathcal{A} o conjunto de todos os subconjuntos de Ω . Neste caso qualquer função real de Ω é uma variável aleatória. Por exemplo, seja X a diferença entre o número de caras e o número de coroas obtidos nos três lançamentos. Então, X pode assumir quatro valores, 3, 1, -1, ou -3. Nosso objetivo é estudar a probabilidade de X assumir cada um desses possíveis valores. Como a moeda é honesta cada um dos possíveis resultados em Ω tem a mesma probabilidade $1/8$. Então, por exemplo, como poderemos obter a probabilidade de X ser negativo?

Notações

Notações comumente encontradas, com os respectivos significados:

$$[X = x] = \{\omega \in \Omega \mid X(\omega) = x\}, \quad B = \{x\},$$

$$[X \leq x] = \{\omega \in \Omega \mid X(\omega) \leq x\}, \quad B = (-\infty, x],$$

$$[x \leq X \leq y] = \{\omega \in \Omega \mid x \leq X(\omega) \leq y\}, \quad B = [x, y].$$

Probabilidade Induzida

A menor σ -álgebra que contém todos os intervalos reais é chamada de σ -álgebra de Borel, \mathcal{B} , e seus elementos, que são subconjuntos de números reais, são chamados de borelianos. Pelas propriedades de σ -álgebra quaisquer uniões e interseções enumeráveis de intervalos reais são também borelianos. Pode-se provar que se X é uma variável aleatória em (Ω, \mathcal{A}, P) , então para todo evento $B \in \mathcal{B}$, chamado evento Boreliano, temos que $X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{A}$.

Deste modo, dada uma variável aleatória X em (Ω, \mathcal{A}, P) , pode-se definir uma probabilidade induzida P_X no espaço mensurável (R, \mathcal{B}) da seguinte maneira: para todo $A \in \mathcal{B}$, definimos $P_X(A) = P(X^{-1}(A))$. Como X é uma variável aleatória, tem-se que $X^{-1}(A) \in \mathcal{A}$, então P_X está bem definida. Resta provar que P_X satisfaz os axiomas K2, K3, e K5' de probabilidade:

Verificando os Axiomas

K2. $P_X(A) = P(X^{-1}(A)) \geq 0.$

K3. $P_X(R) = P(X^{-1}(R)) = P(\Omega) = 1.$

K5'. Suponha que A_1, A_2, \dots são eventos Borelianos disjuntos. Então,

$$P_X(\cup_i A_i) = P(X^{-1}(\cup_i A_i)) = P(\cup_i X^{-1}(A_i)) = \sum_i P(X^{-1}(A_i)) = \sum_i P_X(A_i).$$

Continuando Exemplo

Exemplo

No exemplo anterior, temos que se o evento de interesse A são todos os reais negativos, então $X^{-1}(A)$ são todos os resultados do experimento que nos dão valores negativos para X , ou seja, são os resultados que contêm menos caras que coroas: $(cara, coroa, coroa)$, $(coroa, cara, coroa)$, $(coroa, coroa, cara)$ e $(coroa, coroa, coroa)$. Portanto, $P_X(A) = 4 \times 1/8 = 1/2$.

Observações

Vale a pena salientar que em muitos problemas, já teremos a informação sobre a distribuição induzida P_X definida em (R, \mathcal{B}) . Nestes casos, estaremos “esquecendo” a natureza funcional de X e nos preocupando apenas com os valores assumidos por X . Estes casos podem ser pensados como se o experimento aleatório fosse descrito por (R, \mathcal{B}, P_X) e $X(w) = w, \forall w \in R$, ou seja, os resultados do experimento aleatório já são numéricos e descrevem a característica de interesse que queremos analisar. É isto que acontecerá quando estivermos apresentando as principais distribuições de variáveis aleatórias.

É importante enfatizar que é usual se referir a variáveis aleatórias por letras maiúsculas X, Y, Z, \dots e aos valores que tais variáveis podem assumir por letras minúsculas x, y, z, \dots . Muitas vezes escreve-se $P(X \in A)$ para representar $P(\{w \in \Omega : X(w) \in A\})$. Por exemplo, $P(X \leq 5) = P(\{w \in \Omega : X(w) \leq 5\})$.

Exercício

Exemplo

Considere que lançamos 3 vezes uma moeda que tem probabilidade de cair cara igual $2/3$. Seja X o número de coroas obtido. Determine:

- (a) $P(X < 3)$.
- (b) $P(1 < X < 3)$.
- (c) $P(X > 1 | X < 3)$.



Função de Distribuição Acumulada

Para uma variável aleatória X , uma maneira simples e básica de descrever a probabilidade induzida P_X é utilizando sua *função de distribuição acumulada*.

Definição

A função de distribuição acumulada de uma variável aleatória X , representada por F_X , é definida por

$$F_X(x) = P_X((-\infty, x]), \forall x \in \mathbb{R}.$$

Propriedades

A função de distribuição acumulada F_X satisfaz as seguintes propriedades:

F1. Se $x \leq y$, então $F_X(x) \leq F_X(y)$.

$$x \leq y \Rightarrow (-\infty, x] \subseteq (-\infty, y] \Rightarrow P_X((-\infty, x]) \leq P_X((-\infty, y]) \Rightarrow F_X(x) \leq F_X(y)$$

F2. Se $x_n \downarrow x$, então $F_X(x_n) \downarrow F_X(x)$.

Se $x_n \downarrow x$, então os eventos $(-\infty, x_n]$ são decrescentes e $\cap_n (-\infty, x_n] = (-\infty, x]$. Logo, pela continuidade da medida de probabilidade, tem-se que $P_X((-\infty, x_n]) \downarrow P((-\infty, x])$, ou seja, $F_X(x_n) \downarrow F_X(x)$.

Propriedades

F3. Se $x_n \downarrow -\infty$, então $F_X(x_n) \downarrow 0$, e se $x_n \uparrow \infty$, então $F_X(x_n) \uparrow 1$.

Se $x_n \downarrow -\infty$, então os eventos $(-\infty, x_n]$ são decrescentes e $\cap_n (-\infty, x_n] = \emptyset$. Logo, pela continuidade da medida de probabilidade, tem-se que $P_X((-\infty, x_n]) \downarrow P(\emptyset)$, ou seja, $F_X(x_n) \downarrow 0$. Similarmente, se $x_n \uparrow \infty$, então os eventos $(-\infty, x_n]$ são crescentes e $\cup_n (-\infty, x_n] = \mathbf{R}$. Logo, pela continuidade da medida de probabilidade, tem-se que $P_X((-\infty, x_n]) \uparrow P(\Omega)$, ou seja, $F_X(x_n) \uparrow 1$.

Teorema

Uma função real G satisfaz F1–F3 se e somente se G é uma distribuição de probabilidade acumulada.

Prova: A prova de que se G for uma distribuição de probabilidade acumulada, então G satisfaz F1-F3 foi dada acima. A prova de que toda função real que satisfaz F1-F3 é uma função de probabilidade acumulada é complexa envolvendo o Teorema da Extensão de Carathéodory, e está fora do escopo deste curso. ■

As discontinuidades

Condição F2 significa que toda função distribuição de probabilidade acumulada F_X é contínua à direita. Ainda mais, como F_X é não-decrescente e possui valores entre 0 e 1, pode-se provar que ela tem um número enumerável de discontinuidades do tipo salto. Pela continuidade à direita, o salto no ponto x é igual a

$$\begin{aligned} F_X(x) - F_X(x^-) &= F_X(x) - \lim_{n \rightarrow \infty} F(x - \frac{1}{n}) \\ &= P_X((-\infty, x]) - \lim_{n \rightarrow \infty} P_X((-\infty, x - \frac{1}{n}]) \\ &= \lim_{n \rightarrow \infty} P_X((x - \frac{1}{n}, x]). \end{aligned}$$

As descontinuidades

Como a sequência de eventos $(x - \frac{1}{n}, x]$ é decrescente e $\cap_n (x - \frac{1}{n}, x] = \{x\}$.
Temos que $\{x\}$ é Boreliano e

$$P_X(x) = F_X(x) - F_X(x^-).^1 \quad (1)$$

Ou seja, a probabilidade da variável aleatória X assumir o valor x é igual ao salto da função de distribuição acumulada F_X no ponto x .

¹ $F_X(a^-) = \lim_{x \rightarrow a^-} F_X(x)$ é o limite de $F_X(x)$ quando x tende a a por valores menores que a , ou seja, o limite a esquerda de $F_X(x)$ quando x tende a a .

Exercício

Exemplo

Determine quais das seguintes funções são funções de distribuição acumuladas, especificando a propriedade que não for satisfeita caso a função não seja uma distribuição acumulada.

(a) $\frac{e^x}{1+e^x}$

(b) $I_{[0,\infty)}(x) + [1 - I_{[0,\infty)}(x)](1 + e^x)/2$

(c) $e^{-|x|}$

(d) $I_{[0,\infty)}(x)$

(e) $I_{(0,\infty)}(x)$

Exercício

Exemplo

Considere a seguinte função $G(x)$.

$$G(x) = \begin{cases} a - 2b & \text{se } x < 0, \\ ax & \text{se } 0 \leq x < 1, \\ a + b(x - 1) & \text{se } 1 \leq x < 2, \\ 1 & \text{se } x \geq 2. \end{cases}$$

- (a) Determine as restrições que as constantes a e b devem satisfazer para que a função $G(x)$ seja função de distribuição acumulada de alguma variável aleatória X .
- (b) Determine o valor de $P(1/2 \leq X \leq 3/2)$ em função de a e b .

Exercício

Exemplo

Seja K o número de íons emitidos por uma fonte em um tempo T . Se $F_K(1) - F_K(1/2) = 0,1$, qual o valor de $P(K = 1)$?

Exemplo

Uma sequência de 10 bytes independentes foi recebida. É sabido que a probabilidade é igual a 0,3 que o primeiro símbolo de um byte sea igual a 0. Seja K o número de bytes recebidos tendo 0 como primeiro símbolo.

- (a) Calcule $P(K = 2)$
- (b) Calcule $F_K(1)$

Calculando Probabilidades

Vamos agora mostrar como usar a função de distribuição acumulada para calcular probabilidades de intervalos.

Lembrando que

$$F_X(x) = P(X \leq x) = P_X((-\infty, x]).$$

$$\begin{aligned} \text{(a)} \quad (-\infty, b] &= (-\infty, a] \cup (a, b], \quad a \leq b \Rightarrow \\ P_X((-\infty, b]) &= P_X((-\infty, a]) + P_X((a, b]) \Rightarrow \\ P_X((a, b]) &= P_X((-\infty, b]) - P_X((-\infty, a]) = F_X(b) - F_X(a) \Rightarrow \end{aligned}$$

$$P(a < X \leq b) = F_X(b) - F_X(a). \quad (2)$$

$$\begin{aligned} \text{(b)} \quad (a, b) \cup \{b\} &= (a, b] \Rightarrow \\ P_X((a, b)) + P_X(\{b\}) &= P_X((a, b]) \Rightarrow \\ P_X((a, b)) &= P_X((a, b]) - P_X(b) = F_X(b) - F_X(a) - P(X = b) \Rightarrow \end{aligned}$$

$$P(a < X < b) = F_X(b^-) - F_X(a). \quad (3)$$

O resultado em 3 foi obtido usando 1 e 2.

Exemplo

$$\begin{aligned}
 \text{(c)} \quad (a, b] \cup \{a\} &= [a, b] \Rightarrow \\
 P_X((a, b]) + P_X(a) &= P_X([a, b]) \Rightarrow \\
 P_X([a, b]) &= P_X((a, b]) + P(X = a) = F_X(b) - F_X(a) + P(X = a) \Rightarrow
 \end{aligned}$$

$$P(a \leq X \leq b) = F_X(b) - F_X(a^-). \quad (4)$$

O resultado em 4 foi obtido usando 1 e 2.

$$\begin{aligned}
 \text{(d)} \quad [a, b) &= (a, b) \cup \{a\} \Rightarrow \\
 P_X([a, b)) &= P_X((a, b)) + P_X(a) = F_X(b) - F_X(a) - P(X = b) + P(X = a) \Rightarrow
 \end{aligned}$$

$$P(a \leq X < b) = F_X(b^-) - F_X(a^-). \quad (5)$$

5 foi obtida a partir de 1 e 2.

Exemplo

$$(e) \quad (-\infty, b] = (-\infty, b) \cup \{b\} \Rightarrow$$

$$P_X((-\infty, b]) = P_X((-\infty, b)) + P(X = b) \Rightarrow$$

$$P_X((-\infty, b)) = P_X((-\infty, b]) - P(X = b) \Rightarrow$$

$$P(-\infty < X < b) = F_X(b^-). \quad (6)$$

6 foi obtida a partir de 1 e 2.

Tipos de VA

Existem três tipos de variáveis aleatórias: **discreta**, **contínua** e **singular**.

VA Discreta

Definição

Uma variável aleatória X é *discreta* se assume valores num conjunto enumerável com probabilidade 1, ou seja, se existe um conjunto enumerável $\{x_1, x_2, \dots\} \subseteq \mathbf{R}$ tal que $P(X = x_i) \geq 0, \forall i \geq 1$ e $P(X \in \{x_1, x_2, \dots\}) = 1$.

A função $p(\cdot)$ definida por

$$p(x_i) = P_X(\{x_i\}), i = 1, 2, \dots$$

e

$$p(x) = 0, x \notin \{x_1, x_2, \dots\},$$

é chamada de *função probabilidade (de massa)* de X . Toda função probabilidade é uma função real e assume valores entre 0 e 1, sendo positiva para uma quantidade enumerável de pontos e tal que $\sum_i p(x_i) = 1$. De modo geral escreve-se

$$0 \leq p(x_i) \leq 1,$$

$$\sum p(x_i) = 1.$$

VA Discreta

O conjunto de pontos

$$(x_i, p(x_i)), i = 1, 2, \dots,$$

é usualmente denotado na literatura por *distribuição de probabilidade* da variável aleatória X .

Para esta variável aleatória tem-se que

$$F_X(x) = \sum_{i: x_i \leq x} p(x_i),$$

de modo que F_X é uma função degrau. Seja $p: \mathbf{R} \rightarrow [0, 1]$, sendo p positiva para uma quantidade enumerável de pontos $\{x_1, x_2, \dots\}$ e satisfazendo $\sum_i p(x_i) = 1$ e seja

$$P(B) = \sum_{x_i \in B} p(x_i), \forall B \in \mathcal{B}.$$

Prova-se que $P(B)$ é uma probabilidade em $(\mathbf{R}, \mathcal{B})$ (P satisfaz os axiomas de Kolmogorov). Logo, a distribuição de uma variável aleatória discreta X pode ser determinada tanto pela função de distribuição acumulada F_X quanto pela sua função de probabilidade p .

Exemplo

Exemplo

Este exemplo mostra como calcular a função de distribuição acumulada para uma variável aleatória discreta. Seja X assumindo os valores 0, 1, 2 com igual probabilidade. Portanto,

$$x < 0 \Rightarrow F_X(x) = 0,$$

$$0 \leq x < 1 \Rightarrow F_X(x) = P(X = 0) = \frac{1}{3},$$

$$1 \leq x < 2 \Rightarrow F_X(x) = P(X = 0) + P(X = 1) = \frac{1}{3} + \frac{1}{3} = \frac{2}{3},$$

$$x \geq 2 \Rightarrow F_X(x) = P(X = 0) + P(X = 1) + P(X = 2) = 1.$$

Assim,

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



Exemplo

Exemplo

Este exemplo mostra como calcular a função probabilidade de X a partir do conhecimento da função de distribuição acumulada. Seja

$$F_X(x) = \begin{cases} 0, & x < -1, \\ \frac{1}{6}, & -1 \leq x < 1, \\ \frac{2}{3}, & 1 \leq x < 3, \\ 1, & x \geq 3. \end{cases}$$

logo,

$$P(X = -1) = F_X(-1) - F_X(-1^-) = \frac{1}{6} - 0 = \frac{1}{6},$$

$$P(X = 1) = F_X(1) - F_X(1^-) = \frac{2}{3} - \frac{1}{6} = \frac{1}{2},$$

$$P(X = 3) = F_X(3) - F_X(3^-) = 1 - \frac{2}{3} = \frac{1}{3}$$

$$P(X = k) = 0, \forall k \neq -1, 1, 3.$$



VA Contínua

Definição

Uma variável aleatória X é *contínua* se existe uma função real $f_X \geq 0$ tal que

$$F_X(x) = \int_{-\infty}^x f_X(t) dt, \forall x \in \mathbb{R}.$$

A função f_X é chamada de *função densidade de probabilidade* de X . F_X é contínua e

$$f_X(x) = F'_X(x).$$

Função Densidade

Uma função $f \geq 0$ é densidade de alguma variável aleatória se e somente se, $\int_{-\infty}^{\infty} f(x)dx = 1$, sendo neste caso fácil provar que a função F definida por $\int_{-\infty}^x f(t)dt$ satisfaz às condições F1, F2, e F3. Portanto, pelo Teorema ??, F é uma função de distribuição acumulada. Então, para variável aleatória contínua, a distribuição de uma variável aleatória contínua X pode ser determinada tanto pela função de distribuição acumulada F_X quanto pela sua função densidade f_X .

Uma variável aleatória X tem densidade se F_X é a integral (de Lebesgue) de sua derivada; sendo, neste caso, a derivada de F_X uma função densidade para X . Em quase todos os casos encontrados na prática, uma variável aleatória X tem densidade se F_X é (i) contínua e (ii) derivável por partes, ou seja, se F_X é derivável no interior de um número finito ou enumerável de intervalos cuja união é \mathbf{R} .

VA Contínua

Por exemplo, seja

$$F_X(x) = \begin{cases} 0, & \text{se } x < 0, \\ x^2, & \text{se } 0 \leq x < 1, \\ 1, & \text{se } x \geq 1. \end{cases}$$

Então X tem densidade pois F_X é contínua e derivável em todos os pontos da reta exceto em $\{0, 1\}$.

Quando X é uma variável aleatória contínua,

$$\begin{aligned} P(X < b) &= F_X(b) - P(X = b) \\ &= F_X(b) - (F_X(b) - F_X(b^-)) \\ &= F_X(b) \\ &= P(X \leq b). \end{aligned}$$

Exemplo

Exemplo

Seja X uma variável aleatória com densidade

$$f_X(x) = \begin{cases} 2(m^2 + 1)x + (k + 1)m + (k + 1)^2, & 0 < x < 1, \\ 0, & x \notin (0, 1), \end{cases}$$

onde m e k são números reais.

- (a) Determine os valores de m e k especificando então $f_X(x)$.
- (b) Obtenha $F_X(x)$.

Exemplo

Solução:

(a)

$$\int_0^1 (2(m^2 + 1)x + (k + 1)m + (k + 1)^2) dx = 1 \Rightarrow$$

$$m^2 + m(k + 1) + (k + 1)^2 = 0 \Rightarrow$$

$$m = \frac{-(k + 1) \pm \sqrt{(k + 1)^2 - 4(k + 1)^2}}{2}.$$

Como

$$k \in \mathbf{R} \Rightarrow -3(k + 1)^2 \geq 0 \Rightarrow k = -1.$$

Logo, $m = 0$. Portanto,

$$f_X(x) = \begin{cases} 2x, & 0 < x < 1, \\ 0, & x \notin (0, 1), \end{cases}$$

Exemplo

(b)

$$x < 0, F_X(x) = 0,$$

$$0 \leq x < 1, F_X(x) = \int_0^x 2t dt = x^2,$$

$$x > 1, F_X(x) = 1.$$

Logo,

$$F_X(x) = \begin{cases} 0, & x < 0, \\ x^2, & x \in [0, 1), \\ 1, & x \geq 1. \end{cases}$$



VA Singular

Definição

Uma variável aleatória X é *singular* se F_X é uma função contínua cujos pontos de crescimento formam um conjunto de comprimento (medida de Lebesgue) nulo.

Pode-se provar que toda função de distribuição de probabilidade acumulada F_X pode ser decomposta na soma de no máximo três funções de distribuição de probabilidade acumuladas, sendo uma discreta, uma contínua e outra singular. Uma variável aleatória que possui apenas partes discreta e contínua é conhecida como uma variável aleatória **mista**.

VA Mista

Na prática, é pouco provável que surja uma variável aleatória singular. Portanto, quase todas as variáveis aleatórias são discretas, contínuas ou mistas.

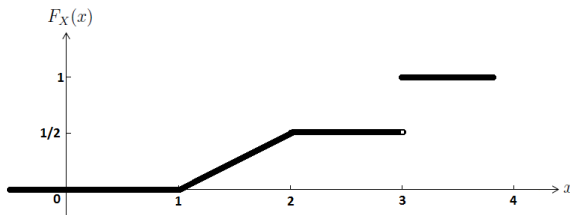


Figura: Gráfico de função de distribuição acumulada para uma variável aleatória mista

Exemplo

Exemplo

Exemplo de cálculo de probabilidades com uma variável aleatória mista.
Seja

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

Exemplo

- (a) $P(X = 0.5)$
- (b) $P(X = 1.5)$
- (c) $P(X = 2)$
- (d) $P(1 < X < 2)$
- (e) $P(1.5 \leq X \leq 2.5)$
- (f) $P(2 \leq X \leq 2.5)$
- (g) $P(2 < X \leq 3)$
- (h) $P(2 \leq X \leq 3)$
- (i) $P(X < 3.7)$
- (j) $P(X < 2)$
- (k) $P(X \leq 1)$
- (l) $P(X \leq 3)$

Decomposição de uma Variável Aleatória Mista

Seja X uma variável aleatória mista qualquer e seja F sua função de distribuição. Se $J = \{x_1, x_2, \dots\}$ é o conjunto dos pontos de salto de F , indiquemos com p_i o salto no ponto x_i , ou seja,

$$p_i = F(x_i) - F(x_i^-).$$

Definimos $F_d(x) = \sum_{i: x_i \leq x} p_i$. F_d é uma função degrau não-decrescente: a *parte discreta* de F . Como uma função monótona possui derivada em quase toda parte, seja

$$f(x) = \begin{cases} F'(x) & \text{se } F \text{ é diferenciável em } x, \\ 0 & \text{se } F \text{ não é diferenciável em } x. \end{cases}$$

Seja $F_{ac}(x) = \int_{-\infty}^x f(t)dt$. F_{ac} é não-decrescente, pois a integral indefinida de uma função não-negativa ($f \geq 0$ porque F é não-decrescente). A sua derivada é igual a f em quase toda parte, de modo que F_{ac} é absolutamente contínua: F_{ac} é a *parte absolutamente contínua* de F .

Exemplo

Suponha que $X \sim U[0, 1]$ e $Y = \min(X, 1/2)$. Note que

$$F_Y(x) = \begin{cases} 0 & \text{se } x < 0, \\ x & \text{se } 0 \leq x < 1/2, \\ 1 & \text{se } x \geq 1/2. \end{cases}$$

F_Y tem apenas um salto em $x = 1/2$ e $p_1 = 1/2$. Logo, $F_d(x) = 0$ se $x < 1/2$ e $F_d(x) = 1/2$ se $x \geq 1/2$. Diferenciando F_Y , temos

$$F'_Y(x) = \begin{cases} 0 & \text{se } x < 0 \text{ ou } x > 1/2, \\ 1 & \text{se } 0 < x < 1/2. \end{cases}$$

Exemplo

Logo, por definição,

$$f(x) = \begin{cases} 0 & \text{se } x \leq 0 \text{ ou } x \geq 1/2, \\ 1 & \text{se } 0 < x < 1/2. \end{cases}$$

Portanto,

$$F_{ac}(x) = \int_{-\infty}^x f(t)dt = \begin{cases} 0 & \text{se } x < 0, \\ x & \text{se } 0 \leq x \leq 1/2, \\ 1/2 & \text{se } x > 1/2. \end{cases}$$

Como $F_d + F_{ac} = F_Y$, temos que não há parte singular.

Distribuições de caudas-pesadas

Uma outra classificação importante de distribuições de variáveis aleatórias diz respeito a magnitude da probabilidade de ocorrência de valores extremos. Intuitivamente, distribuições que possuem uma maior probabilidade de ocorrência de valores extremos é chamada de uma *distribuição de cauda pesada*. Formalmente,

Definição

Uma variável aleatória X segue uma distribuição de caudas-pesadas se

$$\lim_{x \rightarrow \infty} e^{\lambda x} P(X > x) = \infty, \forall \lambda > 0.$$

Exemplo

Seja X tal que

$$f_X(x) = \alpha k^\alpha x^{-(\alpha+1)}, k \in \mathbf{R}^+, x \geq k, \alpha > 0.$$

Então X tem caudas-pesadas porque

$$F_X(x) = 1 - \left(\frac{k}{x}\right)^\alpha, x \geq k \Rightarrow P(X > x) = \left(\frac{k}{x}\right)^\alpha.$$

Esta é a distribuição de Pareto, que será vista mais adiante no curso.

Pareto versus Exponencial

A Tabela 1 mostra resultados para a distribuição de Pareto, quando $k = \alpha = 1$ e outra variável aleatória² de densidade $f_X(x) = e^{-x}$, $x > 0$, a qual não tem caudas-pesadas.

Tabela: Resultados: Pareto *versus* Exponencial.

x	$P(X > x) = e^{-x}$	$P(X > x) = x^{-1}$
2	0.1353352...	0.5
5	0.0067397...	0.2
10	0.0000453...	0.1

²que também será estudada mais adiante no curso.

Pareto versus Exponencial

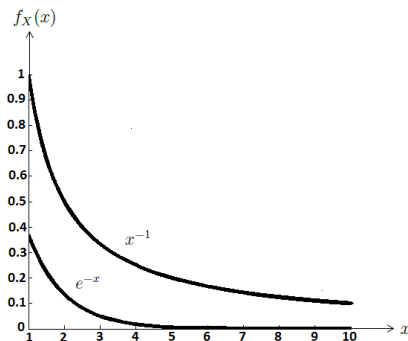


Figura: Comparação entre distribuição Exponencial e distribuição de Pareto

Distribuição de Caudas Pesadas

Observa-se que as probabilidades da variável que não tem caudas-pesadas rapidamente tendem a zero. Não são apenas variáveis aleatórias contínuas que têm caudas-pesadas. Segundo Fine (2006), a Zeta que é discreta também tem caudas-pesadas. Variáveis de caudas-pesadas têm sido usadas na modelagem de tráfego de redes de computadores porque, como o tráfego nas redes é variado (som, imagem, pacotes, dados, etc) qualquer que seja a variável observada, esta apresentará grande variabilidade o que implica no aumento da variância e consequentemente probabilidades altas nas caudas da distribuição (Adler, Feldman, Taqqu, (1998)).

Funções de Variáveis Aleatórias

Muitas vezes é dada a distribuição de probabilidade que descreve o comportamento de uma variável aleatória X definida no espaço mensurável (Ω, \mathcal{A}) , mas o interesse é na descrição de uma função $Y = H(X)$. Por exemplo, X pode ser uma mensagem enviada em um canal de telecomunicações e Y ser a mensagem recebida.

Uma pergunta inicial é: se X é uma variável aleatória \sqrt{X} , $\log X$, X^2 , $2X - 3$ são variáveis aleatórias? Se sim, (o que é verdade), sendo conhecida a distribuição de probabilidade de X , como esse fato pode ser usado para encontrar a lei de probabilidade de \sqrt{X} , $\log X$, X^2 ou $2X - 3$?

Funções de Variáveis Aleatórias

O problema é determinar $P(Y \in C)$, onde C é um evento Boreliano. Para determinar essa probabilidade, a imagem inversa da função H é fundamental, ou seja, a probabilidade do evento $\{Y \in C\}$ será por definição igual a probabilidade do evento $\{X \in H^{-1}(C)\}$, onde $H^{-1}(C) = \{x \in \mathbf{R} : H(x) \in C\}$. Para que esta probabilidade esteja bem definida, é preciso restringir H tal que $H^{-1}(C)$ seja um evento Boreliano para todo C Boreliano, caso contrário não é possível determinar $P(\{X \in H^{-1}(C)\})$; uma função que satisfaz esta condição é conhecida como *mensurável com respeito a \mathcal{B}* . Note que Y também pode ser vista como uma função do espaço amostral Ω , $Y(\omega) = H(X(\omega))$ para todo $\omega \in \Omega$. Vista dessa maneira Y é uma variável aleatória definida em (Ω, \mathcal{A}) , pois para todo Boreliano C , $Y^{-1}(C) = X^{-1}(H^{-1}(C))$ e como por suposição $H^{-1}(C)$ é Boreliano porque X é uma variável aleatória, tem-se que $X^{-1}(H^{-1}(C)) \in \mathcal{A}$ e portanto satisfaz a definição de uma variável aleatória. A figura a seguir exibe os espaços mensuráveis e as transformações entre eles.

Funções de Variáveis Aleatórias

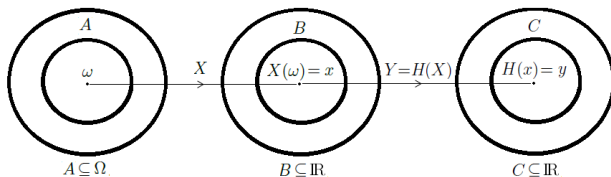


Figura: Mapeamento realizado por funções de variáveis aleatórias

Funções de Variáveis Aleatórias

Seja $A = \{\omega \in \Omega : X(\omega) \in B\}$. Portanto, como já mencionado anteriormente, a probabilidade induzida pela variável aleatória X é tal que

$$P_X(B) = P(X^{-1}(B)) = P(A).$$

De forma similar, sendo

$$B = Y^{-1}(C) = \{x \in \mathbf{R} : H(x) \in C\}$$

então,

$$P_Y(C) = P_{H(X)}(C) = P_X(\{x \in \mathbf{R} : H(x) \in C\}) = P(\{\omega \in \Omega : H(X(\omega)) \in C\}),$$

e assim,

$$P_Y(C) = P_X(Y^{-1}(C)).$$

Logo,

$$P_Y(C) = P_X(B) = P(A).$$

Funções de Variáveis Aleatórias

A função H da variável aleatória X define uma variável aleatória no espaço de probabilidade $(\mathbf{R}, \mathcal{B}, P_Y)$, onde a medida de probabilidade P_Y é induzida pela variável aleatória $Y = H(X)$. P_Y está bem definida pois

$$Y^{-1}(C) = B \in \mathcal{B},$$

o que mostra que a imagem inversa do conjunto mensurável C é o conjunto mensurável B . Adicionalmente, pode-se mostrar que P_Y satisfaz os axiomas de Kolmogorov e, portanto, todas as propriedades de probabilidade.

Funções de Variáveis Aleatórias

Vale ressaltar que se X for uma variável aleatória discreta, $Y = H(X)$ também o será. Os exemplos a seguir ilustram como calcular a distribuição de probabilidade de uma função de variável aleatória. Ressalta-se a importância fundamental da função de distribuição acumulada, F , e de gráficos para visualizar as regiões C e B .

Exemplos

X , **discreta**; $Y = H(X)$, **discreta**. Admita-se que X tenha os valores possíveis $1, 2, 3, \dots$ e que $P(X = n) = (1/2)^n$. Seja $Y = 1$ se X for par e $Y = -1$ se X for ímpar. Determinar a distribuição de Y .

Exemplos

Solução: Então,

$$P(Y = 1) = \sum_{n=1}^{\infty} (1/2)^{2n} = \sum_{n=1}^{\infty} (1/4)^n = \frac{1/4}{1 - 1/4} = 1/3.$$

Consequentemente,

$$P(Y = -1) = 1 - P(Y = 1) = 2/3.$$



Caso Discreto

De modo geral, suponha que X assume os valores x_1, x_2, \dots e que H uma função real tal que $Y = H(X)$ assume os valores y_1, y_2, \dots . Agrupando os valores que X assume de acordo os valores de suas imagens quando se aplica a função H , ou seja, denotando por $x_{i1}, x_{i2}, x_{i3}, \dots$ os valores de X tal que $H(x_{ij}) = y_i$ para todo j , tem-se que

$$P(Y = y_i) = P(X \in \{x_{i1}, x_{i2}, x_{i3}, \dots\}) = \sum_{j=1}^{\infty} P(X = x_{ij}) = \sum_{j=1}^{\infty} p_X(x_{ij}),$$

ou seja, para calcular a probabilidade do evento $\{Y = y_i\}$, acha-se o evento equivalente em termos de X , isto é, todos os valores x_{ij} de X tal que $H(x_{ij}) = y_i$ e somam-se as probabilidades de X assumir cada um desses valores. ■

Exemplos

Exemplo

X , **discreta**; $Y = H(X)$, **discreta**. Seja X como no exemplo anterior e $Y = H(X) = X^2$. Determinar a lei de Y .

Solução: O contradomínio da variável Y , R_Y , e as respectivas probabilidades são:

$$R_Y = \{1, 4, \dots, n^2, \dots\},$$

$$P(Y = 1) = P(X = 1) = \frac{1}{2},$$

$$P(Y = 4) = P(X = 2) = \frac{1}{4},$$

...

$$P(Y = n^2) = P(X = n) = \frac{1}{2^n}.$$

...



Exemplos

Exemplo

X , contínua; $Y = H(X)$, discreta. Seja $f_X(x) = 2x$, $0 < x < 1$ e $Y = H(X)$ definida por $Y = 0$ se $X < \frac{1}{3}$, $Y = 1$, se $\frac{1}{3} \leq X < \frac{2}{3}$ e $Y = 2$, se $X \geq \frac{2}{3}$. Determinar a distribuição de Y .

Solução: Em termos de eventos equivalentes tem-se que:

$$C_1 = \{Y = 0\} \equiv B_1 = \{X < \frac{1}{3}\}, C_2 = \{Y = 1\} \equiv B_2 = \{\frac{1}{3} \leq X < \frac{2}{3}\},$$

$$C_3 = \{Y = 2\} \equiv B_3 = \{X \geq \frac{2}{3}\}.$$

Logo,

$$P(Y = 0) = P(X < \frac{1}{3}) = \int_0^{\frac{1}{3}} 2x dx = \frac{1}{9},$$

$$P(Y = 1) = P(\frac{1}{3} \leq X < \frac{2}{3}) = \int_{\frac{1}{3}}^{\frac{2}{3}} 2x dx = \frac{3}{9},$$

$$P(Y = 2) = P(X \geq \frac{2}{3}) = \int_{\frac{2}{3}}^1 2x dx = \frac{5}{9}.$$

Exemplos

X , **contínua**; $Y = H(X)$, **contínua**. Seja a densidade de X como no exemplo anterior e $Y = H(X) = e^{-X}$. Determinar a distribuição de Y .

Solução: O evento onde a densidade de X é não nula é $B = \{0 < X < 1\}$.

Portanto, a densidade de Y está concentrada em

$C = \{y = H(x) : x \in B\} = \{e^{-1} < y < 1\}$ e

$$\begin{aligned}
 F_Y(y) &= P(Y \leq y) \\
 &= P(e^{-X} \leq y) \\
 &= P(-X \leq \ln y) \\
 &= P(X \geq -\ln y) \\
 &= \int_{-\ln y}^1 2x dx \\
 &= 1 - (-\ln y)^2 \Rightarrow f_Y(y) = \frac{-2 \ln y}{y}.
 \end{aligned}$$

Exemplos

Logo,

$$f_Y(y) = \begin{cases} \frac{-2 \ln y}{y}, & y \in (e^{-1}, 1), \\ 0, & y \notin (e^{-1}, 1). \end{cases}$$



Exemplos

Exemplo

Se $f_X(x) = 1$, $0 < x < 1$, e zero para quaisquer outros valores, qual a distribuição de $Y = -\log(X)$?

Solução: Como

$$0 < Y < \infty \Leftrightarrow 0 < X < 1$$

e $P(0 < X < 1) = 1$, tem-se $F_Y(y) = 0, y \leq 0$. Se $y > 0$, então

$$P(Y \leq y) = P(-\log(X) \leq y) = P(X \geq e^{-y}) = 1 - e^{-y},$$

ou seja, $Y \sim \text{Exp}(1)$, isto é, uma Exponencial (que será vista depois) de parâmetro 1. ■

Exemplos

Seja $f_X(x) = \frac{1}{3}x^2$, $-1 < x < 2$ e zero para quaisquer outros valores de x .
 Encontrar a função densidade da variável aleatória $Y = X^2$.

Exemplos

Solução: Como pode ser visto na figura abaixo,

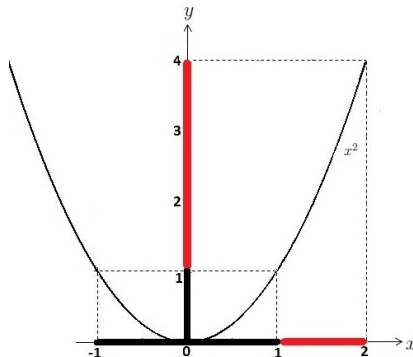


Figura: Mapeamento realizado por funções de variáveis aleatórias

$$-1 < x < 1 \Rightarrow 0 < y < 1$$

Exemplos

Portanto,

$$f_Y(y) = \begin{cases} \frac{\sqrt{y}}{3}, & y \in (0, 1), \\ \frac{\sqrt{y}}{6}, & y \in [1, 4). \\ 0, & y \notin (0, 4). \end{cases}$$



Caso contínuo

No caso de X e Y serem contínuas, tem-se o teorema seguinte.

Teorema

Seja H uma função diferenciável, crescente ou decrescente em um dado intervalo I , $H(I)$ o contradomínio de H , H^{-1} a função inversa de H e X uma variável aleatória contínua com função densidade $f_X(x) > 0$, se $x \in I$ e $f_X(x) = 0$, se $x \notin I$. Então, $Y = H(X)$ tem função densidade de probabilidade dada por:

$$f_Y(y) = \begin{cases} 0, & y \notin H(I), \\ f_X(H^{-1}(y)) \left| \frac{dH^{-1}(y)}{dy} \right|, & y \in H(I). \end{cases}$$

Caso contínuo

Prova:

(a) H é crescente. Logo, H^{-1} também é crescente em I . Portanto,

$$\begin{aligned} F_Y(y) &= P(Y \leq y) \\ &= P(H(X) \leq y) \\ &= P(X \leq H^{-1}(y)) \\ &= F_X(H^{-1}(y)). \end{aligned}$$

Logo,

$$\frac{d}{dy} F_Y(y) = \frac{d}{dy} F_X(H^{-1}(y)) = \frac{dF_X(H^{-1}(y))}{dx} \frac{dx}{dy},$$

onde $x = H^{-1}(y)$.

Caso contínuo

Mas,

$$\frac{d}{dy} F_Y(y) = F'_Y(y) = f_Y(y).$$

Portanto,

$$\frac{dF_X(H^{-1}(y))}{dx} \frac{dx}{dy} = F'_X(H^{-1}(y)) \frac{dH^{-1}(y)}{dy}.$$

Logo,

$$f_Y(y) = f_X(H^{-1}(y)) \frac{dH^{-1}(y)}{dy}, y \in H(I).$$

Caso contínuo

- (b) H é decrescente em I . Então H^{-1} também é decrescente em I .
Logo,

$$\begin{aligned}
 F_Y(y) &= P(Y \leq y) \\
 &= P(H(X) \leq y) \\
 &= P(X \geq H^{-1}(y)) \\
 &= 1 - F_X(H^{-1}(y)) + P(X = H^{-1}(y)) \\
 &= 1 - F_X(H^{-1}(y)).
 \end{aligned}$$

Porque $P(X = H^{-1}(y)) = 0$ e seguindo o procedimento visto em (a),

$$F'_Y(y) = -F'_X(H^{-1}(y)) \frac{dH^{-1}(y)}{dy}$$

e assim

$$f_Y(y) = -f_X(H^{-1}(y)) \frac{dH^{-1}(y)}{dy}, y \in H(I).$$



Exemplos

Seja X com densidade $f_X(x)$ e $Y = X^2$. Então

$$\begin{aligned}
 F_Y(y) &= P(Y \leq y) \\
 &= P(X^2 \leq y) \\
 &= P(-\sqrt{y} \leq X \leq \sqrt{y}) \\
 &= F_X(\sqrt{y}) - F_X(-\sqrt{y}) + P(X = -\sqrt{y}) \\
 &= F_X(\sqrt{y}) - F_X(-\sqrt{y}),
 \end{aligned}$$

porque $P(X = -\sqrt{y}) = 0$. Logo,

$$\frac{d}{dy} F_Y(y) = \frac{d}{dy} (F_X(\sqrt{y}) - F_X(-\sqrt{y})) = \frac{d}{dy} F_X(\sqrt{y}) - \frac{d}{dy} F_X(-\sqrt{y}).$$

Exemplos

Mas,

$$\begin{aligned}\frac{d}{dy} F_Y(y) &= f_Y(y), \\ \frac{d}{dy} F_X(\sqrt{y}) &= \frac{dF_X(\sqrt{y})}{dx_1} \frac{dx_1}{dy}, \quad x_1 = \sqrt{y}, \\ \frac{dF_X(\sqrt{y})}{dx_1} &= f_X(\sqrt{y}), \\ \frac{dx_1}{dy} &= \frac{1}{2\sqrt{y}}, \\ \frac{d}{dy} F_X(-\sqrt{y}) &= \frac{dF_X(-\sqrt{y})}{dx_2} \frac{dx_2}{dy}, \quad x_2 = -\sqrt{y}, \\ \frac{dF_X(-\sqrt{y})}{dx_2} &= f_X(-\sqrt{y}), \\ \frac{dx_2}{dy} &= -\frac{1}{2\sqrt{y}}.\end{aligned}$$

Logo,

$$f_Y(y) = \begin{cases} \frac{1}{2\sqrt{y}} (f_X(\sqrt{y}) + f_X(-\sqrt{y})), & y \geq 0, \\ 0, & y < 0, \end{cases}$$