

Variáveis Aleatórias Contínuas: Função Densidade de Probabilidade e Função de Distribuição

Prof. Alejandra, Cristian, Paula e Rafael

Introdução à Probabilidade e Estatística



Estes slides tem por objetivo apenas complementar e esclarecer os textos principais indicados para a disciplina, e não devem ser usados como único texto de referência.

Junto à essas notas o aluno deve ler os textos além da referência principal do curso.

Dado um experimento aleatório com espaço amostral S :

- Uma **variável aleatória** (v.a.) X é uma função $X : S \rightarrow \mathbb{R}$, que associa a cada resultado $\omega \in S$ um valor real $X(\omega)$.
- Os **valores que assume** a v.a. X são os $a \in \mathbb{R}$ tais que existe $\omega \in S$ tal que $X(\omega) = a$.
- Se X assume uma quantidade finita ou infinita contável (enumerável) de valores dizemos que X é uma **v.a. discreta**.
- Dentre as variáveis que assumem uma quantidade infinita e não contável de valores estão as **v.a.'s contínuas**.
- Em geral, uma v.a. contínua assume quaisquer valor em um intervalo dos reais. Ou seja, se X é uma v.a. contínua, então o conjunto de valores assumidos por X é do tipo $\mathbb{R}$, $[a, b]$, $[a, +\infty)$ e similares.

Neste módulo vamos nos ocupar de estudar v.a. contínuas. Começemos fazendo, de novo, com uma pequena lista de exemplos deste tipo de variáveis aleatórias:

- Uma válvula eletrônica é instalada em um circuito, sendo X o período de tempo que a válvula funciona. A v.a. X é contínua;
- Escolhemos aleatoriamente uma certa estação de trem. O horário de chegada H de um trem nesta estação numa certa data é uma v.a. contínua;
- Considere o experimento de ligar uma lâmpada e esperar que ela apague sozinha. A variável T que mede o tempo de vida da lâmpada é uma v.a. contínua.
- Lançamos um dardo em um alvo circular. A variável R que mede a distância do dardo ao centro do alvo é uma v.a. contínua.

Para este tipo de v.a. (que pode assumir uma quantidade não contável de valores) não é suficiente (e nem útil) obter as probabilidades dos eventos unitários $\{X = a\}$.

Tome, por exemplo, uma v.a. contínua X que assume todos os valores no intervalo $[a, b]$ de tal forma que

$$P(X = x) = P(X = y) \leq k, \text{ para quaisquer par de pontos } x, y \in [a, b].$$

Observemos que, independentemente do valor de k , é sempre possível encontrar $x_1, x_2, \dots, x_n \in [a, b]$ de tal forma que

$$\sum_{i=1}^n P(X = x_i) \geq nk > 1.$$

Esta última igualdade entra em claro conflito com a definição de função de probabilidade, e por isso não faria sentido definir função de probabilidade como fizemos nas variáveis aleatórias discretas.

De fato, chegamos a comentar anteriormente que se X é uma v.a. contínua, então $P(X = a) = 0$ para todo $a \in \mathbb{R}$.

Não chegamos a justificar essa afirmativa. Mais do que isso, nunca chegamos a dar uma definição real do que é uma variável aleatória contínua.

Entendemos algumas particularidades dela, como o fato de assumir valores em intervalos ou outros conjuntos não enumeráveis, e chegamos inclusive a calcular a função distribuição acumulada em um ou dois exemplos. Mas nunca chegamos a definir o que é uma variável contínua.

Assim, sem mais delongas, passemos então para a definição de uma v.a. contínua!!

Definição 1

Diremos que X é uma v.a. contínua se existe uma função $f : \mathbb{R} \rightarrow [0, +\infty)$, conhecida por **função de densidade de probabilidade (f.d.p.)** da v.a. X , tal que

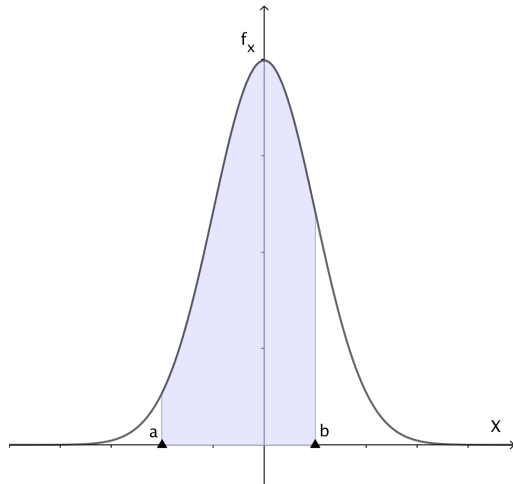
$$P(a \leq X \leq b) = \int_a^b f(x) dx, \quad (1)$$

com $a, b \in \mathbb{R}$, com $a < b$.

Colocando em palavras, a equação (1) diz que a probabilidade de que X assumia valores entre a e b , pode ser obtida integrando-se a função densidade de probabilidade ao longo do intervalo (a, b) .

Função Densidade de Probabilidade

Geometricamente, isso significa que a $P(a \leq X \leq b)$ é igual à área limitada pelo gráfico de $y = f(x)$, as retas $x = a$ e $x = b$ e o eixo x , como na figura abaixo.



Com essa definição em mãos podemos finalmente justificar por que $P(X = a) = 0$.

Se tomarmos $a \in \mathbb{R}$, então

$$P(X = a) = P(a \leq X \leq a) = \int_a^a f(x) dx = 0. \quad (2)$$

Ou seja:

Se X é uma v.a. contínua então $P(X = a) = 0$, para qualquer $a \in \mathbb{R}$.

Isto é quase que o oposto do que acontece com v.a.'s discretas, com as quais é suficiente conhecer a probabilidade dos conjuntos do tipo $\{X = a\}$.

Como consequência da igualdade (2), podemos ver que

$$P(a \leq X \leq b) = P(a < X < b) = P(a < X \leq b) = P(a \leq X < b).$$

Estudando uma destas igualdades, vemos que, de fato

$$P(a \leq X \leq b) = \underbrace{P(X = a)}_0 + P(a < X < b) + \underbrace{P(X = b)}_0 = P(a < X < b).$$

Para eventos do tipo $\{X \geq a\}$ ou $\{X < b\}$, as definições são semelhantes. Vale, por exemplo, que

$$P(X \geq a) = \int_a^{+\infty} f(x) dx. \quad (3)$$

Mas o que significa esta última integral? As integrais deste tipo são chamadas improprias e são uma forma simplificada de escrever um limite. Neste caso, temos que

$$\int_a^{+\infty} f(x) dx = \lim_{t \rightarrow +\infty} \int_a^t f(x) dx.$$

Veremos integrais improprias com muita frequência daqui em diante, então temos de nos acostumar com seu uso. Só para ilustrar, vejamos alguns exemplos.

- Suponhamos que uma v.a. X tem f.d.p. dada por

$$f(x) = \begin{cases} 1, & \text{para } x \in [0, 1], \\ 0, & \text{em qualquer outro caso.} \end{cases}$$

Calcule $P(X \geq 3/5)$.

$$P(X \geq 3/5) = \int_{\frac{3}{5}}^{+\infty} f(x) dx = \int_{\frac{3}{5}}^1 1 dx = \frac{2}{5}.$$

- Seja Y uma v.a. com f.d.p. dada por

$$g(x) = \begin{cases} e^{-x}, & \text{para } x \in (0, +\infty), \\ 0, & \text{em qualquer outro caso.} \end{cases}$$

Calcule $P(X \geq 1)$.

$$P(X \geq 1) = \int_1^{+\infty} e^{-x} dx = \lim_{a \rightarrow +\infty} [-e^{-x}]_1^a = \lim_{a \rightarrow +\infty} (-e^{-a} + e^{-1}) = e^{-1}.$$

Finalmente, note que se $f(x)$ é a função densidade de probabilidade de uma v.a. X , então

$$\int_{-\infty}^{+\infty} f(x) dx = P(-\infty < X < +\infty) = 1. \quad (4)$$

Mas em particular, se todos os valores que pode assumir a v.a. X estão no intervalo (a, b) então

$$\int_a^b f(x) dx = P(a < X < b) = 1. \quad (5)$$

Proposição 1

Sejam X uma v.a. contínua e f_X sua função de densidade de probabilidade. Então:

- ❶ Para todo $a \in \mathbb{R}$, $f_X(a) \geq 0$;
- ❷ $\int_{-\infty}^{+\infty} f_X(x) dx = P(X \in \mathbb{R}) = 1$.
- ❸ $P(a \leq X \leq b) = P(a < X < b) = \int_a^b f_X(x) dx$;
- ❹ Para todo $a \in \mathbb{R}$, $P(X = a) = 0$;

Suponhamos que uma certa v.a. X tenha f.d.p. dada por

$$f(x) = \begin{cases} cx, & \text{para } x \in (0, 4), \\ 0, & \text{em qualquer outro caso.} \end{cases}$$

Calcule c , $P(1 \leq X \leq 2)$ e $P(X > 2)$.

Solução: Para calcular c devemos usar o fato de que f é uma f.d.p. e portanto

$$1 = \int_{-\infty}^{+\infty} f(x) dx = \int_0^4 cx dx = c \frac{x^2}{2} \Big|_0^4 = 8c.$$

Deste modo, $c = 1/8$, e portanto

$$P(1 \leq X \leq 2) = \int_1^2 f(x) dx = \frac{1}{8} \int_1^2 x dx = \frac{1}{8} \frac{x^2}{2} \Big|_1^2 = \frac{3}{16} \text{ e}$$

$$P(X > 2) = \int_2^{+\infty} f(x) dx = \frac{1}{8} \int_2^4 x dx = \frac{1}{8} \frac{x^2}{2} \Big|_2^4 = \frac{3}{4}.$$

Lembramos que se X é uma v.a. então a sua função de distribuição acumulada F_X é definida como $F_X(a) = P(X \leq a)$, para $a \in \mathbb{R}$.

Em particular:

Se X é uma v.a. contínua e sua função densidade de probabilidade é f_X então

$$F_X(a) = P(X \leq a) = \int_{-\infty}^a f_X dx.$$

A voltagem X em um determinado sistema elétrico é uma variável aleatória com uma distribuição contínua cuja f.d.p. pode se aproximar pela função

$$f(x) = \begin{cases} 0, & \text{para } x \in (-\infty, 0], \\ \frac{1}{(1+x)^2}, & \text{para } x \in (0, +\infty). \end{cases}$$

Para encontrar a função a função de distribuição acumulada $F(a)$ devemos considerar dois casos:

- Se $a \leq 0$, então $F(a) = P(X \leq a) = \int_{-\infty}^a f dx = \int_{-\infty}^a 0 dx = 0$.
- Se $a > 0$, então $F(a) = P(X \leq a) = \int_{-\infty}^a f dx = \int_0^a \frac{1}{(1+x)^2} dx = 1 - \frac{1}{1+a}$.

Portanto,

$$F(a) = \begin{cases} 0, & \text{para } a \in (-\infty, 0], \\ 1 - \frac{1}{(1+a)}, & \text{para } a \in (0, +\infty). \end{cases}$$

Função de Distribuição Acumulada

Vamos agora relembrar algumas propriedades da função distribuição, mas no contexto das variáveis aleatórias contínuas.

Se $a_1 < a_2$ usando as propriedades da integral podemos escrever

$$F_X(a_2) = P(X \leq a_2) = \int_{-\infty}^{a_2} f_X dx = \underbrace{\int_{-\infty}^{a_1} f_X dx}_{F_X(a_1)} + \underbrace{\int_{a_1}^{a_2} f_X dx}_{P(a_1 \leq X \leq a_2)} . \quad (6)$$

Como $P(a_1 \leq X \leq a_2) \geq 0$ então podemos afirmar que

$$F(a_1) \leq F(a_2).$$

Ou seja, F_X é uma função não-decrescente.

Observação: Lembre que, como já vimos, não é necessário usar integrais para concluir isto! Basta observar que para $a_1 < a_2$, $\{X \leq a_1\} \subseteq \{X \leq a_2\}$ e usar as propriedades de uma probabilidade para mostrar que

$$\underbrace{P(X \leq a_1)}_{F_X(a_1)} \leq \underbrace{P(X \leq a_2)}_{F_X(a_2)}.$$

Vamos usar de novo a igualdade (6)

$$F_X(a_2) = P(X \leq a_2) = \int_{-\infty}^{a_2} f_X dx = \underbrace{\int_{-\infty}^{a_1} f_X dx}_{F_X(a_1)} + \underbrace{\int_{a_1}^{a_2} f_X dx}_{P(a_1 \leq X \leq a_2)},$$

e a vamos reescreve-la da seguinte forma

$$P(a_1 \leq X \leq a_2) = F_X(a_2) - F_X(a_1).$$

Resumimos no que segue as propriedades que acabamos de discutir e outras, que já discutimos anteriormente.

Propriedades

Se $F_X(x)$ é a função distribuição acumulada de uma v.a. contínua X com f.d.p. f_X , então vale que:

- 1 $F_X(a) = P(X \leq a) = P(X < a) = \int_{-\infty}^a f_X dx$
- 2 $P(a \leq X \leq b) = F_X(b) - F_X(a);$
- 3 $F_X(x)$ é não-decrescente. Ou seja, se $a < b$, então $F_X(a) \leq F_X(b);$
- 4 $\lim_{x \rightarrow \infty} F_X(x) = 1$ e $\lim_{x \rightarrow -\infty} F_X(x) = 0.$

Observação:

Vale enfatizar mais uma vez que quando estamos lidando com uma variável aleatória contínua X , com função distribuição $F_X(x)$, então

$$P(a < X < b) = P(a \leq X \leq b) = F_X(b) - F_X(a),$$

e ainda

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

Da mesma forma

$$P(X < a) = P(X \leq a) = F_X(a)$$

e

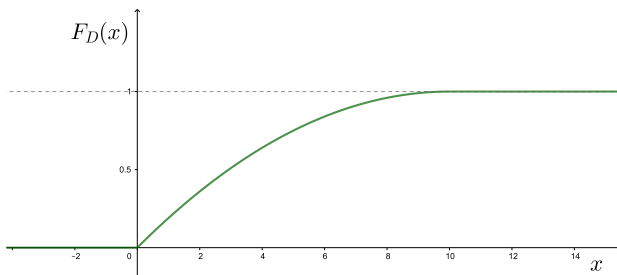
$$P(X \geq a) = P(X > a) = 1 - F_X(a).$$

Exemplo

Lançamos uma moeda ao acaso em um chão quadriculado, coberto por azulejos de 20 cm de lado. Seja D a distância do centro da moeda para a borda de azulejo mais próxima. A função distribuição da v.a. D , como vimos no Módulo “Introdução às variáveis aleatórias discretas”, é dada por:

$$F_D(x) = \begin{cases} 0 & \text{se } x < 0; \\ \frac{20x - x^2}{100} & \text{se } 0 \leq x < 10; \\ 1 & \text{se } x \geq 10; \end{cases}$$

cujo gráfico é dado abaixo.



Observando o gráfico de F_D vemos que F_D é uma função não-decrescente,

$$\lim_{x \rightarrow \infty} F_D(x) = 1 \quad \text{e que} \quad \lim_{x \rightarrow -\infty} F_D(x) = 0,$$

e se quisermos agora calcular $P(D \geq 5)$ e $P(2 < D < 4)$ fazemos:

$$P(D \geq 5) = 1 - P(D < 5) = 1 - F_D(5) = 1 - \frac{75}{100} = \frac{25}{100},$$

e

$$P(2 < D < 4) = F_D(4) - F_D(2) = \frac{64}{100} - \frac{36}{100} = \frac{28}{100}.$$

Mas quem é a função densidade de probabilidade de D ?

Lembrando, sabemos que a função distribuição é dada por

$$F_D(x) = \begin{cases} 0 & \text{se } x < 0; \\ \frac{20x - x^2}{100} & \text{se } 0 \leq x < 10; \\ 1 & \text{se } x \geq 10; \end{cases}$$

Mas sabemos também que se D tem f.d.p $f(x)$, então

$$F_D(t) = \int_{-\infty}^t f(x) dx,$$

e pelo Teorema Fundamental do Cálculo, segue que

$$f(t) = \frac{dF_D}{dt}(t),$$

e portanto

$$f(x) = \begin{cases} 0 & \text{se } x < 0 \text{ ou } x \geq 10; \\ \frac{10-x}{50} & \text{se } 0 \leq x < 10; \end{cases}$$

No caso discreto, é simples encontrar a função de probabilidade de uma v.a. conhecendo a sua função de distribuição: fazemos isto olhando nos saltos da distribuição.

Mas como vimos, para as v.a. contínuas a passagem da função distribuição à função densidade é mais “sofisticada” e se dá, via o Teorema Fundamental do Cálculo, pela seguinte igualdade:

$$\frac{dF_X(t)}{dt} = f_X(t), \text{ para todo } t \text{ onde } f_X \text{ é contínua.} \quad (7)$$

A equação anterior diz que derivando função de distribuição acumulada obtemos a função de densidade de probabilidade.

Escolhemos ao acaso um ponto em um círculo S de raio 1. Seja R a distância do ponto escolhido ao centro da circunferência. Suponhamos que a probabilidade do ponto observado estar em uma região $K \subset S$ do círculo é dada por

$$P(K) = \frac{\text{Área}(K)}{\text{Área}(S)} = \frac{\text{Área}(K)}{\pi}.$$

Encontre a f.d.p da v.a. R .

Solução: Sabemos que a v.a. R assume valores no conjunto $[0, 1]$, e já vimos em módulos anteriores que, para $0 \leq r \leq 1$, $P(R \leq r) = \frac{\pi r^2}{\pi} = r^2$. Portanto,

$$F_R(r) = \begin{cases} 0, & \text{para } x \in (-\infty, 0), \\ r^2, & \text{para } x \in [0, 1], \\ 1, & \text{para } x \in [1, +\infty). \end{cases}$$

Logo

$$f_R(r) = \begin{cases} 2r, & \text{para } x \in [0, 1], \\ 0, & \text{para } x \in (-\infty, 0) \cup (1, +\infty). \end{cases}$$

Suponhamos que uma v.a. contínua X tem função de distribuição acumulada dada por

$$F(x) = \begin{cases} 0, & \text{para } x \in (-\infty, 0), \\ x^{2/3}, & \text{para } x \in [0, 1], \\ 1, & \text{para } x \in (1, +\infty). \end{cases}$$

Observemos que F é contínua em $\mathbb{R}$ e diferenciável em $\mathbb{R} \setminus \{0, 1\}$. Portanto, a f.d.p. de X pode ser encontrada, em qualquer ponto diferente de 0 e 1, derivando F . Desta forma, temos que

$$f(x) = \frac{2}{3}x^{-\frac{1}{3}}, \text{ para } x \in (0, 1) \quad \text{e} \quad f(x) = 0, \text{ para } x \in (-\infty, 0) \cup (1, +\infty).$$

O valor de f em 0 e 1 pode ser atribuído arbitrariamente.

Encontre a f.d.p. de uma v.a. T cuja função distribuição acumulada é dada por

$$F_T(t) = \begin{cases} 0 & \text{se } t < 0; \\ 1 - e^{-2t} & \text{se } t \geq 0. \end{cases}$$

Para aqueles que já cursaram FVV, deixamos o seguinte desafio.

Encontre o valor da constante $c \in \mathbb{R}$ para que a função

$$\phi(x) = c e^{-\frac{x^2}{2}}, \quad x \in \mathbb{R}$$

possa ser a função densidade de probabilidade de alguma variável aleatória contínua X .