Teoria da Informação

Charles Casimiro Cavalcante

charles@gtel.ufc.br

Grupo de Pesquisa em Telecomunicações Sem Fio – GTEL
Programa de Pós-Graduação em Engenharia de Teleinformática
Universidade Federal do Ceará – UFC
http://www.gtel.ufc.br/∼charles

"A principal função de um sistema de comunicação é reproduzir, exatamente ou de forma aproximada, uma informação proveniente de outro ponto diferente."

Claude Shannon, 1948

Conteúdo do curso

1. Revisão de probabilidade
2. Informação e Entropia
3. Codificação de fontes
4. Codificação e capacidade de canal
5. Complexidade de Kolmogorov
6. Funções de otimização
7. *Independent Component Analysis*

Parte I

Revisão de Probabilidade

Evento

Definição: Qualquer subconjunto do espaço amostral $\mathcal{S}$ que constitui um campo de Borel $\mathcal{F}$

Eventos mutuamente exclusivos

Quando a ocorrência de um impossibilita a ocorrência do outro

Exemplo: Dado

$$\left.\begin{array}{r} A = \{\text{par}\} \\ B = \{\text{impar}\} \end{array}\right\rangle A \cdot B = \emptyset \quad \textbf{(eventos mutuamente exclusivos)}$$

Probabilidade (Definição Axiomática)

É qualquer função real definida na classe $\mathcal{F}$ tal que

1. $\Pr(A) \geq 0$
2. $\Pr(\mathcal{S}) = 1$
3. Se $A \cdot B = \emptyset \Rightarrow \Pr(A + B) = \Pr(A) + \Pr(B)$ (eventos mutuamente exclusivos)

Assim,

$$\Pr(\cdot) : \mathcal{F} \to \mathbb{R}$$

Probabilidade condicional

Probabilidade de ocorrência de A dado que ocorreu B

$$\Pr(A|B) \triangleq \frac{\Pr(AB)}{\Pr(B)}, \quad \Pr(B) > 0 \tag{1}$$

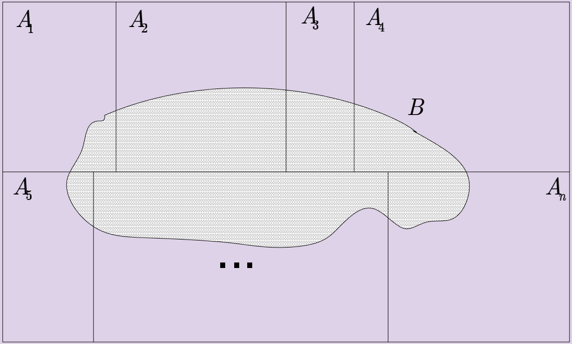
$\Pr(A|B)$ é probabilidade, pois

- $\Pr(A|B) = \frac{\Pr(AB)}{\Pr(B)} \geq 0$
- $\Pr(\mathcal{S}|B) = 1$
- Para $A \cdot C = \emptyset \Rightarrow \Pr\left[(A + C)|B\right] = \Pr(A|B) + \Pr(C|B)$

Teorema da probabilidade total

Sejam $A_1, A_2, A_3, \ldots, A_n$ eventos mutuamente exclusivos

- $\Pr(A_i) > 0, \quad i = 1, 2, \ldots, n$
- $B \subset \{A_1 + A_2 + \cdots + A_n\}$



Teorema da probabilidade total - cont.

- $\Pr(B) = \Pr(BA_1) + \Pr(BA_2) + \cdots + \Pr(BA_n)$ pois
  $B = \underbrace{BA_1 + BA_2 + \cdots + BA_n}_{\text{mutuamente exclusivos}}$
- $\Pr(B) = \Pr(B|A_1)\Pr(A_1) + \cdots + \Pr(B|A_n)\Pr(A_n)$

Probabilidade total

$$\Pr(B) = \sum_{i=1}^{n} \Pr(B|A_i)\Pr(A_i) \tag{2}$$

Regra de Bayes

Inverso do conceito da probabilidade total

$$\Pr(A_j|B) = \frac{\Pr(B|A_j) \cdot \Pr(A_j)}{\sum_{i=1}^{n} \Pr(B|A_i) \cdot \Pr(A_i)} \tag{3}$$

Também chamada de *probabilidade a posteriori*

Eventos independentes

Dois eventos A e B são independentes se

$$\Pr(A \cdot B) = \Pr(A) \cdot \Pr(B)$$

Generalizando (para três eventos): A, B e C

$$\left.\begin{array}{r}\Pr(AB) = \Pr(A)\Pr(B) \\ \Pr(AC) = \Pr(A)\Pr(C) \\ \Pr(BC) = \Pr(B)\Pr(C)\end{array}\right\} \Pr(ABC) = \Pr(A)\Pr(B)\Pr(C)$$

Propriedades de eventos independentes

1. $\Pr(A|B) = \Pr(A)$
2. $\Pr(\overline{A}B) = \Pr(\overline{A}) \cdot \Pr(B)$
3. $\Pr(A\overline{B}) = \Pr(A) \cdot \Pr(\overline{B})$ e $\Pr(\overline{AB}) = \Pr(\overline{A}) \cdot \Pr(\overline{B})$

Ou seja Se A e B são independentes, A e $\overline{B}$ são independentes e $\overline{A}$ e $\overline{B}$ também o são

Eventos conjuntos

Dado $\mathcal{S}$ (espaço amostral), podemos atribuir diferentes atributos aos eventos pertencentes a diferentes classes de Borel

$\mathcal{S} = \{x_1, x_2, \ldots, x_n\}$

$$\begin{cases} A_1, A_2, \cdots, A_n \in \mathcal{F}_1 \\ B_1, B_2, \cdots, B_n \in \mathcal{F}_2 \end{cases}$$

Exemplo:

$\mathcal{S} = \{$João, José, Maria $\}$

(idade,altura)

Rescrevendo:

$\mathcal{S} = \{(10, 1.50), (30, 1.80), (32, 1.65)\}$

Probabilidade marginal

$$A_1 + A_2 + \cdots + A_n = \mathcal{S}_1$$

$$B_1 + B_2 + \cdots + B_n = \mathcal{S}_2$$

$$\left.\begin{aligned} \Pr(A_i) &= \sum_{j=1}^{n} \Pr(A_i, B_j) \\ \Pr(B_j) &= \sum_{i=1}^{n} \Pr(A_i, B_j) \end{aligned}\right\} \sum_{i=1}^{n} \sum_{j=1}^{n} \Pr(A_i, B_j) = 1 \tag{4}$$

Definição

Variável aleatória (v.a.) é qualquer função definida no espaço amostral $\mathcal{S}$ tal que:

$$\{X : \mathcal{S} \to \mathbb{R}, X(w) \in (-\infty, x], w \in \mathcal{S}\} \in \mathcal{F} \tag{5}$$

Exemplo

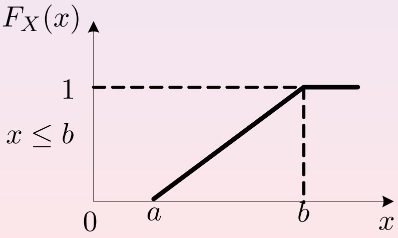
- Moeda:
  $\mathcal{S} = \{\text{cara}, \text{coroa}\}$
  $X(\text{cara}) = 0$
  $X(\text{coroa}) = 1$

Definição

$$\boxed{F_X(x) \triangleq \Pr\{X \leq x\}} \tag{6}$$

Exemplo: *Distribuição uniforme*

$$F_X(x) = \begin{cases} 0, & \text{se } x \leq a \\ x - a, & \text{se } a < x \leq b \\ 1, & \text{se } x > b \end{cases}$$



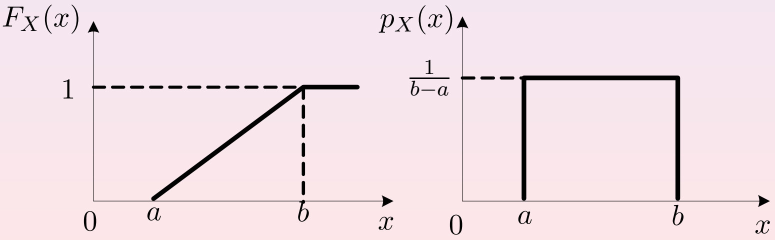
Propriedades da fdc

1. $F_X(-\infty) = 0$
2. $F_X(\infty) = 1$
3. $\Pr\{x_1 \leq X \leq x_2\} = F_X(x_2) - F_X(x_1)$
4. Se $x_1 \leq x_2 \Rightarrow F_X(x_1) \leq F_X(x_2)$, ou seja, $F_X(x)$ é *monotônico não-decrescente*

Definição

$$\boxed{p_X(x) \triangleq \frac{d}{dx} F_X(x)} \tag{7}$$

Exemplo: *Distribuição uniforme*

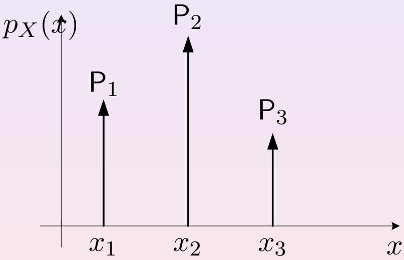


Propriedades

1. $F_X(x) = \int\limits_{-\infty}^{x} p_X(\xi)\; d\xi$
2. $F_X(x)$ é monotônico não-decrescente $\Rightarrow p_X(x) \geq 0$
3. $\Pr\{X > x\} = 1 - F_X(x) = 1 - \Pr\{X \leq x\}$
4. $\Pr\{x_1 < X \leq x_2\} = F_X(x_2) - F_X(x_1) = \int\limits_{x_1}^{x_2} p_X(\xi)\; d\xi$

Dificuldade

Neste caso, as variáveis admitem valores somente em determinados instantes de tempo. O que ocorre com as probabilidades?



Funções $\delta(\cdot)$ de Dirac (impulsos)

$$\delta(t) \begin{cases} \int_{-\infty}^{\infty} f(t)\delta(t-t_0)\ dt = f(t_0) \\ \int_{-\infty}^{\infty} \delta(t)\ dt = 1 \end{cases}$$

$$\begin{aligned} \delta(t) &= \text{função impulsiva de Dirac} \\ &= \frac{d}{dt}u(t), \text{em que } u(t) \text{é a função degrau unitário} \end{aligned}$$

Logo, teremos

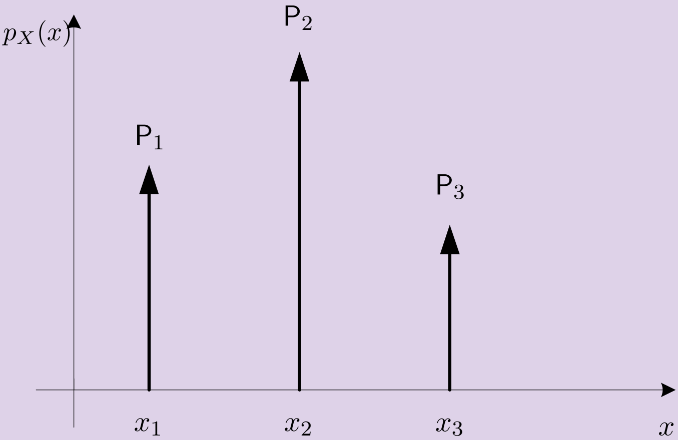
$$p_X(x) = \sum_{i=1}^{N} \Pr\{X = x_i\} \cdot \delta(x - x_i)$$

$$F_X(x) = \int_{-\infty}^{x} p_X(\xi)\ d\xi = \sum_{i=1}^{N} \Pr\{X = x_i\} \cdot \int_{-\infty}^{x} \delta(\xi - x_i)\ d\xi$$

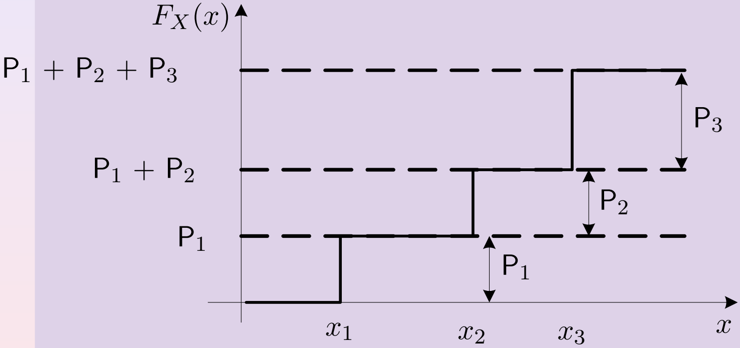
Mas sabe-se que

$$\int_{-\infty}^{x} \delta(\xi - x_i)\ d\xi = \begin{cases} 0, & \text{se } x < x_i \\ 1, & \text{se } x \geq x_i \end{cases}$$

pdf



fdc



Definição

Seja X uma v.a., X é dito ter distribuição de probabilidade gaussiana, ou *normal*, se sua densidade de probabilidade pode ser escrita da seguinte forma

$$\boxed{p_X(x) = \frac{1}{\sqrt{2\pi\sigma}} \cdot \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)} \tag{8}$$

Notação usual:

$$X \sim \mathcal{N}\left(\mu, \sigma^2\right)$$

$$\text{Parâmetros} \begin{cases} \mu \rightarrow & \text{média} \\ \sigma^2 \rightarrow & \text{variância} \end{cases}$$

Normalização

$$Z \sim \mathcal{N}(0,1)$$

$$\boxed{p_Z(z) = \frac{1}{2\pi} \cdot \exp\left(-\frac{z^2}{2}\right)}$$

Função erro

Função de distribuição cumulativa da função gaussiana

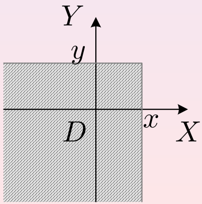
$$\begin{aligned} \mathsf{erf}(x) &= F_Z(x) = \Pr\{Z \leq x\} \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{z^2}{2}\right) \, dz \end{aligned} \tag{9}$$

Função distribuição de probabilidade

$$F_{X,Y}(x,y) = \Pr \underbrace{\{X \le x, \quad Y \le y\}}_{\text{intersecção}} \tag{10}$$

$$\underbrace{\{X \le x, \quad Y \le y\}}_{\text{evento}} = \{ \; w \in \mathcal{S} | [X(w), Y(w)] \in D \; \}$$

em que $D = \{ \; (X,Y) | X \in (-\infty, x], \quad Y \in (-\infty, y] \; \}$



$$\boxed{p_{X,Y}(x,y) = \frac{\partial^2}{\partial x \, \partial y} F_{X,Y}(x,y)} \tag{11}$$

Propriedades

1. $F_{X,Y}(-\infty, y) = 0$
2. $F_{X,Y}(x, -\infty) = 0$
3. $F_{X,Y}(\infty, \infty) = 1$
4. $$\left.\begin{array}{l} F_{X,Y}(x, \infty) = F_X(x) \\ F_{X,Y}(\infty, y) = F_Y(y) \end{array}\right\}\text{distribuições marginais}$$
5. $$F_{X,Y}(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} p_{X,Y}(\alpha, \beta)\ d\alpha\ d\beta$$
6. $$p_X(x) = \int_{-\infty}^{\infty} p_{X,Y}(x, y)\ dy$$
   $$p_Y(y) = \int_{-\infty}^{\infty} p_{X,Y}(x, y)\ dx$$

Definição

Sejam X e Y v.a.'s. Elas são independentes se:

$$\boxed{F_{X,Y}(x,y) = F_X(x) \cdot F_Y(y)} \tag{12}$$

$$\begin{aligned} F_{X,Y}(x,y) &= \Pr\{\, X \leq x, \quad Y \leq y \,\} \\ &= \Pr\{X \leq x\} \cdot \Pr\{Y \leq y\} \end{aligned}$$

Logo:

$$\boxed{p_{X,Y}(x,y) = p_X(x) \cdot p_Y(y)} \quad \text{pois}$$

$$p_{X,Y}(x,y) = \frac{\partial^2}{\partial x\, \partial y} F_{X,Y}(x,y)$$

$$= \frac{\partial^2}{\partial x\, \partial y} \left(F_X(x) \cdot F_Y(y)\right) \overbrace{=}^{?} p_X(x) \cdot \; p_Y(y)$$

Momentos

São *estatísticas* de uma variável aleatória capazes de representar seu comportamento probabilístico. Os infinitos momentos estatísticos definem a função de densidade de probabilidade.

Média

Também chamada de esperança matemática, valor esperado, momento de 1^a ordem, é definido como

$$\boxed{\mu = \mathbb{E}\{X\} \triangleq \int_{-\infty}^{\infty} x \cdot p_X(x)\ dx} \tag{13}$$

para $X \sim$ v.a. contínua

Média

Para X discreta

$$\boxed{\mu = \mathbb{E}\{X\} = \sum_{i=-\infty}^{\infty} x_i \cdot \Pr\{X = x_i\}} \tag{14}$$

OBS: Se $p_X(x)$ for simétrico em relação a um valor $x = a \Rightarrow \mathbb{E}\{X\} = a$

Pergunta: Num jogo de moeda, qual a média? E no caso de uma distribuição uniforme entre $[0, 1]$?

Propriedades da média

1. Linearidade - $\mathbb{E}\{X+Y\} = \mathbb{E}\{X\} + \mathbb{E}\{Y\}$ ou ainda,
$$\mathbb{E}\left\{\sum_{i=1}^{m} a_i X_i\right\} = \sum_{i=1}^{m} a_i \mathbb{E}\left\{X_i\right\}$$
2. $\mathbb{E}\{X \cdot Y\} = \mathbb{E}\{X\} \cdot \mathbb{E}\{Y\}$ se X e Y são v.a.'s independentes
3. Transformação linear - $\mathbb{E}\left\{\mathbf{Ax}\right\} = \mathbf{A}\mathbb{E}\left\{\mathbf{x}\right\}$
4. $\mathbb{E}\{f(X,Y)\} = \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} f(x,y)p_{X,Y}(x,y)\; dxdy$

Propriedades da média - cont.

5. Invariância à transformação - Seja $\mathbf{y} = g(\mathbf{x})$, então

$$\int_{-\infty}^{\infty} \mathbf{y} p_Y(\mathbf{y})\, d\mathbf{y} = \int_{-\infty}^{\infty} g(\mathbf{x}) p_X(\mathbf{x})\, d\mathbf{x}$$

6. Se X e Y são independentes
   $\mathbb{E}\{f(X) \cdot g(Y)\} = \mathbb{E}\{f(X)\} \cdot \mathbb{E}\{g(Y)\}$

Momentos de ordem k

$$\boxed{\mu_k = \mathbb{E}\left\{X^k\right\} = \int_{-\infty}^{\infty} x^k \cdot p_X(x)\ dx} \tag{15}$$

- Os momentos de uma variável aleatória são uma representação da pdf da variável
- A coletânea dos *infinitos* momentos da v.a. definem sua pdf
- Algumas distribuições possuem alguns momentos nulos
- A estimativa de momentos cresce em complexidade e decresce em precisão com o aumento direto de k

Momentos centrados

Uma importante medida estatística é avaliar o comportamento da v.a. *em torno* da média. Assim, define-se o *momento centrado de ordem k* como sendo

$$c_k = \mathbb{E}\left\{(X-\mu)^k\right\} = \int_{-\infty}^{\infty} (x-\mu)^k \cdot p_X(x)\, dx \tag{16}$$

De particular interesse: variância $\left(\sigma^2 = \mathbb{E}\left\{(X-\mu)^2\right\} \geq 0\right)$

OBS: Se $p_X(x)$ é simétrica em relação a média $c_k = 0$ para $\forall k$ ímpar!

Meta

Caracterização da distribuição de probabilidade de uma variável aleatória condicionada a ocorrência de outra variável aleatória ou evento

Definição distribuição cumulativa condicionada

$$\begin{aligned} F_X(x|A) &\triangleq \Pr\{X \leq x|A\} \\ &= \frac{\Pr\{X \leq x, A\}}{\Pr\{A\}} \end{aligned} \tag{17}$$

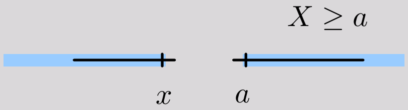
- $$p_X(x|A) \triangleq \frac{d}{dx} F_X(x|A)$$

Analisando...

Seja $A = \{x \geq a\}$ (evento)

$$\begin{aligned} F_X(x|A) &= F_X(x|X \geq a) \\ &= \Pr\{X \leq x | X \geq a\} \end{aligned}$$

(a) $x < a$

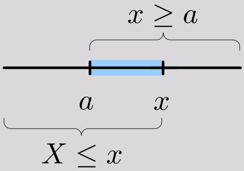


$$\Rightarrow F_X(x|x \geq a) = 0$$

Variáveis aleatórias

Distribuições e densidades condicionais - cont.

(a) $x \geq a$



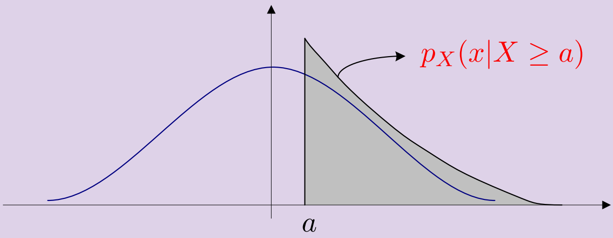
$$F_X(x|x \geq a) = \frac{\Pr\{X \leq x, X \geq a\}}{\Pr\{X \geq a\}} = \frac{\int_a^x p_X(\alpha)\ d\alpha}{\int_a^\infty p_X(\beta)\ d\beta}$$

$$= \frac{F_X(x) - F_X(a)}{1 - F_X(a)}$$

$$p_X(x|X \geq a) = \frac{d}{dx} F_X(x|X \geq a) = \frac{\frac{d}{dx} F_X(x)}{1 - F_X(a)} = \frac{p_X(x)}{1 - F_X(a)}$$

Resumindo

$$p_X(x|X \geq a) = \frac{p_X(x)}{\int\limits_0^{\infty} p_X(\beta)\ d\beta} U(x-a)$$



Observações

1. $$\int_{-\infty}^{\infty} p_X(x|X \geq a)\ dx = \int_{-\infty}^{\infty} \left[\frac{p_X(x)}{\int_a^{\infty} p_X(\beta)\ d\beta} \right] \cdot U(x-a)\ dx$$
$$= \frac{\int_a^{\infty} p_X(x)\ dx}{\int_a^{\infty} p_X(\beta)\ d\beta} = 1$$

Observações - cont.

2. $\mathbb{E}\{X|A\} = ?$

$$A = \{X \geq a\}$$

$$\mathbb{E}\{X|X \geq a\} = \int_{-\infty}^{\infty} x \cdot p_X(x|X \geq a)\ dx$$

$$= \int_{\infty}^{\infty} \frac{x p_X(x)}{\int_a^{\infty} p_X(\beta)\ d\beta} \cdot U(x-a)\ dx$$

$$\boxed{\mathbb{E}\{X|X \geq a\} = \frac{\int_a^{\infty} x \cdot p_X(x)\ dx}{\int_a^{\infty} p_X(\beta)\ d\beta}}$$

Observações - cont.

3. Caso em que $A = \{X = a\}$

$$F_X(x|X=a) = \frac{\Pr\{X \leq x, X = a\}}{\underbrace{\Pr\{X = a\}}_{\text{v.a. contínua} \Rightarrow \Pr\{X=a\}=0}}$$

Relaxando "um pouco" $X = a$ para

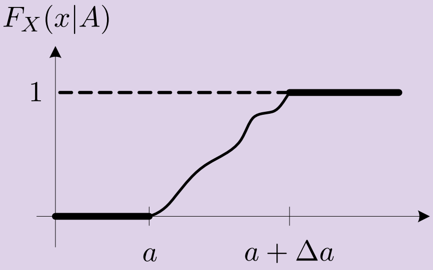
$$a \leq X \leq a + \Delta a, \quad \Delta a \to 0 \qquad \text{(depois)}$$

$$F_X(x|A) = F_X(x|a \leq X \leq a + \Delta a) = \frac{\Pr\{X \leq x, a \leq X \leq a + \Delta a\}}{\Pr\{a \leq X \leq a + \Delta a\}}$$

Observações - cont.

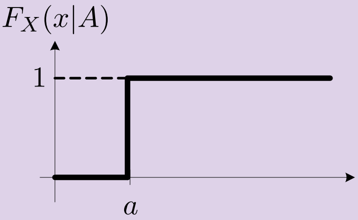
Temos 3 situações

- $X < a \Rightarrow F_X(x|A) = 0$
- $a \leq X \leq a + \Delta a \Rightarrow F_X(x|A) = \frac{\Pr\{a \leq X \leq x\}}{\Pr\{a \leq X \leq a + \Delta a\}}$
- $X > a + \Delta a \Rightarrow F_X(x|A) = 1$



Observações - cont.

Fazendo $\Delta \to 0$:



$$F_X(x|A) = \underbrace{U(x-a)}_{\text{função degrau unitário}}$$

$$\Rightarrow F_X(x|X=a) = U(x-a)$$

$$\boxed{p_X(x|X=a) = \frac{d}{dx}F_X(x|X=a) = \delta(x-a)}$$

Função densidade condicional de duas variáveis aleatórias

$$p_Y(y|x) = ?$$

$$F_Y(y|x \leq X \leq x + \Delta x) = \frac{\Pr\{Y \leq y, x \leq X \leq x + \Delta x\}}{\Pr\{x \leq X \leq x + \Delta x\}}$$

Ainda

$$\Pr\{Y \leq y, x \leq X \leq x + \Delta x\} = \int_{-\infty}^{y} \int_{x}^{x+\Delta x} p_{X,Y}(\alpha, \beta)\ d\alpha d\beta$$

$$\cong p_{X,Y}(x, \beta)\Delta x \quad \text{método de Euler}$$

$$= \int_{-\infty}^{y} p_{X,Y}(x, \beta)\ d\beta \cdot \Delta x$$

Função densidade condicional de duas variáveis aleatórias - cont.

$$\Pr\{x \leq X \leq x + \Delta x\} = \int\limits_{x}^{x+\Delta x} p_X(\gamma)\ d\gamma \cong p_X(x) \cdot \Delta x$$

$$F_Y(y|x) \cong \frac{\int\limits_{-\infty}^{y} p_{X,Y}(x, \beta)\ d\beta \cdot \Delta x}{p_X(x) \cdot \Delta x} = \frac{\int\limits_{-\infty}^{y} p_{X,Y}(x, \beta\ d\beta}{p_X(x)}$$

$$p_Y(y|x) = \frac{dF_Y(y|x)}{dy} = \frac{p_{X,Y}(x, y)}{p_X(x)}$$

Se X e Y forem independentes

$$\begin{aligned} p_{X,Y}(x, y) &= p_X(x) \cdot p_Y(y) \\ \Rightarrow p_Y(y|x) &= p_Y(y) \end{aligned} \tag{18}$$

- $p_{X|Y}(x|y) = \dfrac{p_{X,Y}(x,y)}{p_Y(y)}$
- $p_{Y|X}(y|x) = \dfrac{p_{X,Y}(x,y)}{p_X(x)}$
- $p_{Y|X}(y|x) = \dfrac{p_{X|Y}(x|y)p_Y(y)}{p_X(x)}$

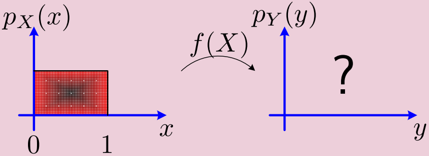
Função de uma variável aleatória

Dada $X \sim$ v.a. e uma função

$$Y = f(X)$$

a questão é como determinar $p_Y(y)$ conhecendo-se $p_X(x)$.
Por exemplo, seja X uma v.a. uniforme em $[0, 1]$

$$Y = \frac{1}{\lambda} \cdot \ln\left(\frac{1}{X}\right)$$



Vamos analisar dois casos

1o. caso

$$X \sim \text{v.a.} \quad p_X(x)\text{conhecido}$$

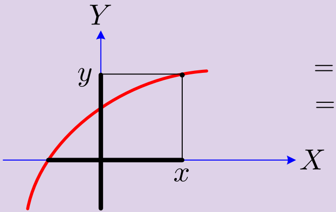
$$\begin{cases} y = f(x) \\ \frac{df(x)}{dx} > 0 \Rightarrow f(x) \text{ é monotônico crescente} \end{cases}$$

$$\rightarrow f \text{ é biunívoca}[p_Y(y), F_Y(y)]$$

$$f_y(Y) = \Pr\{Y \leq y\} = \Pr\{f(X) \leq Y\}$$

Como

$\frac{df(x)}{dx} > 0 \Rightarrow$



$$= \Pr\{X \leq f^{-1}(y)\}$$
$$= F_X\left[f^{-1}(y)\right]$$

1o. caso - cont.

$$p_Y(y) = \frac{dF_Y(y)}{dy} = \frac{dF_X\left[f^{-1}(y)\right]}{dy}$$
$$= \frac{d}{dy}\int_{-\infty}^{f^{-1}(y)} p_X(x)\ dx$$
$$= p_X\left[f^{-1}(y)\right]\cdot\frac{df^{-1}(y)}{dy}$$

Mas: $\frac{df^{-1}(y)}{dy} = \frac{dx}{dy} = \frac{1}{\frac{dy}{dx}} = \frac{1}{\frac{df(x)}{dx}}$ Logo,

$$\boxed{p_Y(y) = p_X(x)\cdot\left.\frac{1}{\frac{df(x)}{dx}}\right|_{x=f^{-1}(y)} \qquad \text{para } \frac{df(x)}{dx} > 0}$$

2o. caso

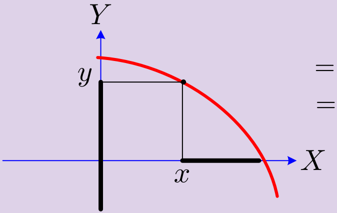
$$X \sim \text{v.a.} \quad p_X(x) \text{conhecido}$$

$$\begin{cases} y = f(x) \\ \frac{df(x)}{dx} < 0 \Rightarrow f(x) \text{ é monotônico decrescente} \end{cases}$$

$$\rightarrow f \text{ é biunívoca} [p_Y(y), F_Y(y)]$$

$$F_Y(y) = \Pr\{Y \leq y\}$$

$$F_Y(y) = \Pr\{f(X) \leq y\}$$



$$= \Pr\{X > f^{-1}(y)\}$$
$$= 1 - F_X\left[f^{-1}(y)\right]$$

2o. caso - cont.

$$p_Y(y) = \frac{F_Y(y)}{dy} = \left.\frac{-p_X(x)}{\frac{df(x)}{dx}}\right|_{x=f^{-1}(y)} \quad \text{com} \frac{df(x)}{dx} < 0$$

$$= \frac{p_X(x)}{\frac{df(x)}{dx}}$$

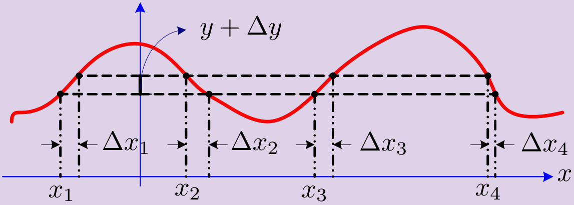
O sinal desaparece pois $\frac{df(x)}{dx}$ também é negativo.

Resumindo, para funções de uma variável aleatória temos:

Para encontrar $p_Y(y)$

$$\boxed{p_Y(y) = \left.\frac{p_X(x)}{\left|\frac{df(x)}{dx}\right|}\right|_{x=f^{-1}(y)}} \tag{19}$$

Funções não-biunívocas



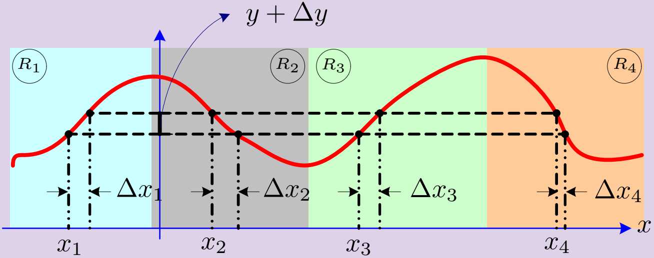
$$
\begin{aligned}
F_Y(y) &= \Pr\{Y \le y\} \\
&= \Pr\{y \le Y \le y + \Delta y\}
\end{aligned}
$$

Como há contribuições de diferentes intervalos de x, então

$$
\begin{aligned}
F_Y(y) = &\Pr\{x_1 \le X_1 \le x_1 + \Delta x_1\} + \Pr\{x_2 - \Delta x_2 \le X_2 \le x_2\} \\
&+ \Pr\{x_3 \le X_3 \le x_3 + \Delta x_3\} + \Pr\{x_4 - \Delta x_4 \le X_4 \le x_4\}
\end{aligned}
$$

Funções não-biunívocas - cont.

Nestes casos, dividimos o espaço amostral em regiões biunívocas. Ou seja, teremos



Funções não-biunívocas - cont.

Nestes casos, chamando de

$$y = f_i(x_i) \quad \text{em cada região } R_i$$

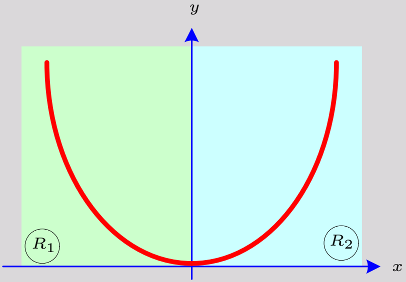
$p_Y(y)$

$$\boxed{p_Y(y) = \sum_{R_i} \left. \frac{p_X(x_i)}{\left| \frac{df_i(x_i)}{dx_i} \right|} \right|_{x_i = f^{-1}(y)}} \tag{20}$$

Exemplo

$$X \sim \text{v.a.} \quad N(0, \sigma^2)$$

$$Y = X^2$$



$$p_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \cdot \exp\left(-\frac{x^2}{2\sigma^2}\right)$$

Exemplo - cont.

$$p_Y(y) = \underbrace{\left.\frac{p_X(x)}{\left|\frac{df_1(x)}{dx}\right|}\right|_{x=f_1^{-1}(y)}}_{R_1} + \underbrace{\left.\frac{p_X(x)}{\left|\frac{df_2(x)}{dx}\right|}\right|_{x=f_2^{-1}(y)}}_{R_2}$$

$$\frac{df_1(x)}{dx} = \frac{dx^2}{dx} = 2x \quad x = -\sqrt{y}$$

$$\frac{df_2(x)}{dx} = \frac{dx^2}{dx} = 2x \quad x = \sqrt{y}$$

$$p_Y(y) = p_X(x)|_{x=-\sqrt{y}} \cdot \frac{1}{2\sqrt{y}} + p_X(x)|_{x=\sqrt{y}} \cdot \frac{1}{2\sqrt{y}}$$

$$= \begin{cases} \frac{1}{\sqrt{2\pi}\sigma} \cdot \exp\left(-\frac{y}{2\sigma^2}\right) \cdot \frac{1}{\sqrt{y}}, & y \geq 0 \\ 0, \quad y < 0 \end{cases}$$

Problema

$$X \sim \text{v.a.} \qquad Y \sim \text{v.a.}$$

$$\begin{cases} U = f(X,Y) \\ V = g(X,Y) \end{cases}$$

Conhecido $p_{X,Y}(x,y)$, como achar $p_{U,V}(u,v)$?

$p_{U,V}(u,v)$

Em regiões biunívocas:

$$\boxed{p_{U,V}(u,v) = \left.\frac{p_{X,Y}(x,y)}{\left|J\left(\frac{u,v}{x,y}\right)\right|}\right|_{\substack{x=\mathfrak{f}(u,v)\\ y=\mathfrak{g}(u,v)}}} \tag{21}$$

em que $\mathfrak{f}(\cdot)$ e $\mathfrak{g}(\cdot)$ são funções inversas. E

$$J\left(\frac{u,v}{x,y}\right) = \det\begin{bmatrix} \frac{du}{dx} & \frac{du}{dy} \\ \frac{dv}{dx} & \frac{dv}{dy} \end{bmatrix} = \frac{1}{\det\begin{bmatrix} \frac{d\mathfrak{f}}{du} & \frac{d\mathfrak{f}}{dv} \\ \frac{d\mathfrak{g}}{du} & \frac{d\mathfrak{g}}{dv} \end{bmatrix}}$$

é chamado *Jacobiano de* u, v *em relação a* x, y

Caso particular

$$Z = X + Y, \qquad X \sim \text{v.a.}, Y \sim \text{v.a.}$$

$$p_{X,Y}(x,y) \text{ conhecido}$$

$$p_X(z) = ?$$

Definir então

$$\begin{cases} z = x + y \\ w = x \end{cases} \Rightarrow p_{Z,W}(z,w)$$

$$J\left(\frac{z,w}{x,y}\right) = \det \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = -1$$

Logo,

$$p_{Z,W}(z,w) = \left.\frac{p_{X,Y}(x,y)}{|-1|}\right|_{\substack{x = \mathfrak{f}(z,w) \\ y = \mathfrak{g}(z,w)}}$$

Caso particular - cont.

$$x = w, \qquad y = z - w$$

$$p_{Z,W}(z, w) = p_{X,Y}(w, z - w)$$

Então, para achar a densidade marginal $p_Z(z)$, temos

$$p_Z(z) = \int_{-\infty}^{\infty} p_{X,Y}(w, z - w)\ dw$$

Caso particular - cont.

Se supormos que X e Y são independentes

$$p_{X,Y}(w, z-w) = p_X(w) \cdot p_Y(z-w)$$
$$\Rightarrow p_{Z,W}(z, w) = p_X(w) \cdot p_Y(z-w)$$

$$p_Z(z) = \int_{-\infty}^{\infty} p_{Z,W}(z, w)\; dw = \underbrace{\int_{-\infty}^{\infty} p_X(w) \cdot p_Y(z-w)\; dw}_{\text{convolução}}$$

Assim:

$$\boxed{p_Z(z) = p_x(x) \star p_Y(y)} \qquad (22)$$