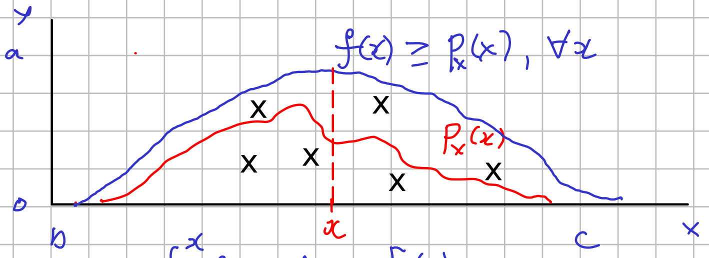
Metodo de aceitação /rejeição



$$u = \int_b^x f(x)\,dx = F(x)$$

$b \leq x \leq c$

Transf. Variáveis

$$x = F^{-1}(u)$$

$$x = F^{-1}\{\text{rand}(1)) \qquad y = f(x)\ \text{rand}(1)$$

Coordenada x dos pontos abaixo da curva $P_x(x)$ (aceites) tem densidade de probabilidade $P_x(x)$. Aceitar se $y \leq P_x(x)$.

métodos numéricos para gerar números aleatórios

- método da aceitação / rejeição
  - se gerarmos um ponto com distribuição uniforme sob a curva $p_X(x)$, então a coordenada x desse ponto tem distribuição $p_X(x)$. Para fazer isso podemos gerar pontos uniformemente no retângulo mas é bastante ineficiente porque os pontos acima da curva seriam rejeitados. É mais eficiente encontrar uma função, f(x), tal que $p_X(x) \leq f(x)$, calcular $F(x_0) = \int_{-\infty}^{x_0} f(x)\, dx$ e obter $x_0 = F^{-1}(u_1)$, com u_1, uniforme em $[0, F(\infty)]$. A seguir geramos u_2 uniforme em $[0, f(x_0)]$. Os pontos (u_2, x_0) têm distribuição uniforme sob a curva f(x) . Se rejeitarmos os x_0 tais que $u_2 > p_X(x_0)$ ficamos com valores de x_0 com a distribuição pretendida.

(x, y)

(x_0, u_2)

entropia

- Entropia de uma variável aleatória com distribuição de probabilidade $p_X(x)$

$$S_X = S(p_X(x)) = -\sum_{x\in\mathscr{X}} p_X(x)\log_2 p_X(x) = \overline{\log_2 \frac{1}{p_X(X)}}$$

com $0\log_2 0 = 0$.

- Propriedades:
  - $S_X \geq 0$
  - S_X toma um valor máximo para a distribuição uniforme.
- Bernoulli variable, X: $S_X = -p\log_2(p) - (1-p)\log_2(1-p)$ tem máximo para $p = 1/2$,

Condição de máximo

$S(p_X(x))$ Restrição $\sum_x p_X(x) = 1$

α é ajustado no final de maneira a obedecer à restrição

$x = x_1, x_2, \ldots \quad \sum_x p_X(x) \log p_X(x) = p_X(x_1) \log p_X(x_1) + p_X(x_2) \log p_X(x_2) + \cdots$

$$\frac{d}{dp_X(x)} \left[-\sum_x p_X(x) \log p_X(x) + \alpha \left(\sum_x p_X(x) - 1 \right) \right] = 0$$

$$-\log p_X(x) - p_X(x) \frac{1}{p_X(x)} + \alpha = 0$$

$\log p_X(x) = \alpha - 1 \Rightarrow p_X(x) = e^{\alpha - 1}$ é independente de x

$x = x_1, x_2, \ldots$

$$\sum_x p_X(x) \log p_X(x) = p_X(x_1) \log p_X(x_1) + p_X(x_2) \log p_X(x_2) + \cdots$$

$$\frac{d}{d p_X(x_1)} \sum_x p_X(x) \log p_X(x) = \log p_X(x_1) + p_X(x_1) \frac{1}{p_X(x_1)}.$$

$$\sum_x p_X(x) = 1 \iff \sum_x e^{\alpha - 1} = 1 \iff e^{\alpha-1} \underbrace{\sum_x 1}_{M \,=\, \text{nº de valores} \neq s \text{ que } x \text{ toma}} = 1$$

$$e^{\alpha - 1} = \frac{1}{M}$$

$$p_X(x) = e^{\alpha - 1} = \frac{1}{M}$$ distribuição uniforme

desigualdade de Jensen

- Para uma função convexa (equivalente a $\frac{d^2f}{dx^2} > 0$),temos, $f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha) f(y)$ com $\alpha \in [0,1]$;

Ex:

$f(x) = x^2$

y_r

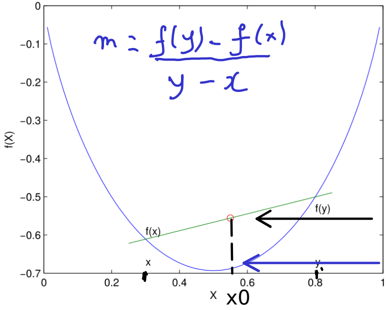
$\boxed{x_0 = \alpha x + (1-\alpha) y}$

$\alpha = 0 \quad y$

$\alpha = 1 \quad x$

$0 \leq \alpha \leq 1$

$m = \frac{f(y) - f(x)}{y - x}$



$y_r = m(x_0 - x) + f(x)$

x_0 é variável

$x \leq x_0 \leq y$

$\alpha f(x) + (1-\alpha) f(y)$

$f(x_0)$

Figure: Função convexa, f(x). A reta secante que passa pelos pontos (x,f(x)) e (y,f(y)) está sempre acima da função entre esses pontos

Equação da reta

$$y_r = \frac{f(y) - f(x)}{y - x}(x_0 - x) + f(x)$$

$$x_0 = \alpha x + (1-\alpha) y \qquad 0 \leq \alpha \leq 1$$

$$y_r \leq f(x_0)$$

$$y_r = \frac{f(y) - f(x)}{y - x}\left(\alpha x + (1-\alpha) y - x\right) + f(x)$$

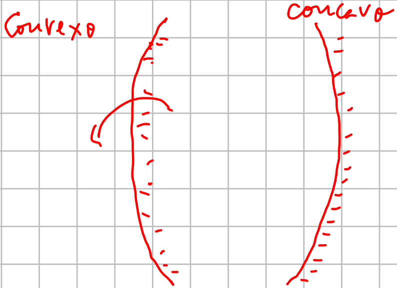
$$= \frac{f(y) - f(x)}{y - x}\underbrace{\left[(\alpha - 1) x + (1-\alpha) y\right]}_{(1-\alpha)(y-x)} + f(x)$$

$$y_r = \left(f(y) - f(x)\right)(1-\alpha) + f(x)$$

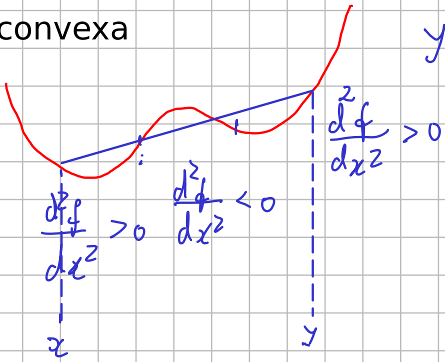
$$= \alpha f(x) + (1-\alpha) f(y)$$

$$y_r = \alpha f(x) + (1-\alpha) f(y) \geq f\left(\alpha x + (1-\alpha) y\right)$$

$$0 \leq \alpha \leq 1$$



função não convexa



$y_r \leq f(x_0)$

não é verdade
para qualquer
x_0 entre x e y

desigualdade de Jensen

$\alpha x + (1-\alpha) y$ é um ponto entre x e y $0 \leq \alpha \leq 1$

- Eq. da reta secante

$$\begin{aligned} z(\alpha) &= f(x) + \frac{f(y)-f(x)}{y-x}(\alpha x + (1-\alpha)y - x) \\ &= \alpha f(x) + (1-\alpha) f(y) \geq f(\alpha x + (1-\alpha) y) \end{aligned}$$

- Se f(x) é uma função convexa então, $\boxed{\overline{f(X)} \geq f(\overline{X})}$ $\forall\, P_X(x)$
- Demonstração
  - Se $\mathscr{X} = \{x_1, x_2\}$ então $\overline{f(X)} = p_1 f(x_1) + p_2 f(x_2)$ com $p_1 + p_2 = 1$ e $\overline{f(X)} \geq f(p_1 x_1 + p_2 x_2)$ ou seja $\overline{f(X)} \geq f(\overline{X})$ com $\alpha = p_1$

$$\overline{f(x)} = \sum_x P_X(x)\, f(x) = \underbrace{P_X(x_1)}_{P_1} f(x_1) + \underbrace{P_X(x_2)}_{P_2} f(x_2)$$

$$\overline{X} = \sum_x P_X(x)\, x = x_1 P_1 + x_2 P_2$$

desigualdade de Jensen

- Se $\mathscr{X} = \{x_1, x_2, x_3\}$ então $\overline{f(X)} = p_1 f(x_1) + p_2 f(x_2) + p_3 f(x_3)$ com $p_1 + p_2 + p_3 = 1$
  - $\overline{f(X)} = (p_1 + p_2)\left[\frac{p_1}{p_1+p_2} f(x_1) + \frac{p_2}{p_1+p_2} f(x_2)\right] + p_3 f(x_3)$ com $\alpha = \frac{p_1}{p_1+p_2}$
    - $\overline{f(X)} \geq (p_1 + p_2) f(\frac{p_1}{p_1+p_2} x_1 + \frac{p_2}{p_1+p_2} x_2) + p_3 f(x_3) \geq f\left((p_1 + p_2)\left[\frac{p_1}{p_1+p_2} x_1 + \frac{p_2}{p_1+p_2} x_2\right] + p_3 x_3\right)$ com $\alpha = p_1 + p_2$
    - $\overline{f(X)} \geq f(p_1 x_1 + p_2 x_2 + p_3 x_3) = f(\overline{X})$

Raciocineo indutivo: É válido para $\mathscr{X} = \{x_1, x_2\}$ implica que é válido para $\mathscr{X} = \{x_1, x_2, x_3\}$ então é válido para $\mathscr{X} = \{x_1, x_2, x_3, \ldots, x_m\}$ c/ n arbitrário

divergência de Kulback-Leibler

compara duas distribuições de probabilidade

- Divergência de Kulback-Leibler entre duas distribuições de probabilidade $q_x(x)$ $p_x(x)$
  - $D(q_X||p_X) = \sum_{x\in\mathscr{X}} q_X(x)\log\frac{q_X(x)}{p_X(x)}$
- Propriedades
  - $D(q_X||p_X) = 0$ se $q_X(x) = p_X(x)$;
  - $D(q_X||p_X) \geq 0$. $\forall\, q_x(x)$ ou $\forall\, p_x(x)$
    - Como $-\log(x)$ é uma função convexa $D(q_X||p_X) = \overline{-\log\frac{p_X(x)}{q_X(x)}} \geq -\log\overline{\frac{p_X(x)}{q_X(x)}} = 0$ dado que $\overline{\frac{p_X(x)}{q_X(x)}} = 1$

$$\overline{\frac{p_x(x)}{q_x(x)}} = \sum_x \cancel{q_x(x)}\frac{p_x(x)}{\cancel{q_x(x)}} = \sum_x p_x(x) = 1$$

$$y = \log(x) \qquad \frac{dy}{dx} = \frac{1}{x}$$

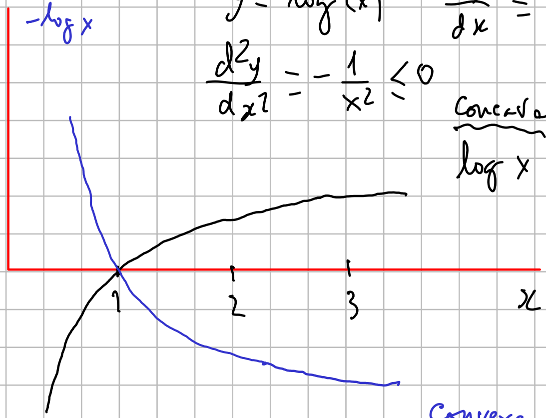
$$\frac{d^2y}{dx^2} = -\frac{1}{x^2} \leq 0$$

Concava

log x

y

$Z = -\log x$



Convexa

desigualdade de Chebyshev

$-\infty$ $\quad \leftarrow \varepsilon \rightarrow \leftarrow \varepsilon \rightarrow$ $\quad +\infty$

$|x-\bar{x}| \geq \varepsilon$ $\quad \bar{X}-x > \varepsilon \quad \bar{X} \quad x-\bar{X} > \varepsilon \quad x$

domínio de integração

- Para qualquer variável aleatória, X, com variância, $Var\,X$, finita verifica-se a desigualdade, para qualquer $\varepsilon > 0$:
  - $prob\left(\left|X-\overline{X}\right| \geq \varepsilon\right) \leq \frac{var\,X}{\varepsilon^2}$.
  - demonstração

$$\frac{(x-\bar{X})^2}{\varepsilon^2} \geq 1$$

- como $\boxed{prob\left(\left|X-\overline{X}\right| \geq \varepsilon\right) = \int_{\left|X-\overline{X}\right| \geq \varepsilon} p_X(x)dx}$ e $\frac{\left|x-\overline{X}\right|}{\varepsilon} \geq 1$ temos $prob\left(\left|X-\overline{X}\right| \geq \varepsilon\right) \leq \frac{1}{\varepsilon^2}\int_{\left|x-\overline{X}\right| \geq \varepsilon}\left(x-\overline{X}\right)^2 p_X(x)dx$
- Como $\frac{1}{\varepsilon^2}\int_{\left|x-\overline{X}\right| \geq \varepsilon}\left(x-\overline{X}\right)^2 p_X(x)dx \leq \frac{1}{\varepsilon^2}\int_{x\in\mathscr{X}}\left(x-\overline{X}\right)^2 p_X(x)dx = \frac{var\,X}{\varepsilon^2}$ prova-se a desigualdade.

várias variáveis

$p_{X,Y}(x,y)$

probabilidade conjunta

- probabilidade condicional (Fórmula de Bayes)

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

  marginal

  $p_Y(y) = \sum_{x\in\mathscr{X}} p_{X,Y}(x,y)$ e portanto $\sum_{x\in\mathscr{X}} p_{X|Y}(x|y) = 1$
- variáveis independentes

  $p_{X,Y}(x,y) = p_X(x)p_Y(y)$ ou seja

  $p_{X|Y}(X = x|Y = y) = p_X(x)$;

  $Var(X+Y) = Var(X) + Var(Y)$
- independentes e idênticamente distribuídas -

  $p_N(x_1, x_2, \cdots x_N) = \prod_{i=1,N} p(x_i)$;
- Acontecimentos mutuamente exclusivos, $\mathscr{A}_1 \cap \mathscr{A}_2 \equiv \oslash$ temos

  $prob(A_1 \cup \mathscr{A}_2) = prob(\mathscr{A}_1) + prob(\mathscr{A}_2)$

  Ex: $X \in \mathscr{X} \equiv \mathscr{R}$ e queremos calcular $prob(|X| \geq \varepsilon)$ com $\varepsilon > 0$.

  Podemos escrever $prob(|X| \geq \varepsilon) = \int_{\varepsilon}^{\infty} p_X(x)dx + \int_{-\infty}^{-\varepsilon} p_X(x)dx$