Modelação e Física Estatística

Conceitos de probabilidade e teoria da informação[1]

António Luís Ferreira

March 8, 2021

[1] slides baseados no Cap. 1 de Information, Physics and Computation, Oxford University press, 2009

Temas

1. Variáveis aleatórias
2. Teoria dos grandes desvios
3. Processos Estocásticos e Cadeias de Markov
4. Introdução à teoria de informação
   - compressão de dados
   - transmissão de Dados

variáveis aleatórias

- X - variável aleatória discreta
- $\mathscr{X}$ - conjunto de valores tomado pela variável, $0 \leq p_X(x) = prob(X = x) \leq 1$, Normalização - $\sum_{x \in \mathscr{X}} p_X(x) = 1$
- Valor médio - $\overline{X} = \sum_{x \in \mathscr{X}} x\, p_X(x)$. $\overline{f(x)} = \sum_{x \in X} f(x)\, P_X(x)$
- Variância - $Var\, X = \overline{(X - \bar{X})^2} = \overline{X^2} - \overline{X}^2 = \sum_{x \in \mathscr{X}} (x - \overline{X})^2 p_X(x)$
- Acontecimento $\mathscr{A} \subseteq \mathscr{X}$ (contido em $\mathscr{X}$) tem probabilidade $prob(\mathscr{A}) = \sum_{x \in \mathscr{A}} p(x)$

$$\overline{f+g} = \bar{f} + \bar{g}$$

$$(x-\bar{x})^2 = x^2 - 2x\bar{x} + \bar{x}^2$$

$$\overline{(x-\bar{x})^2} = \overline{x^2} - \overline{2\bar{x}\,x} + \overline{\bar{x}^2}$$

$$= \overline{x^2} - 2\bar{x}\,\bar{x} + \bar{x}^2$$

$$= \overline{x^2} - \bar{x}^2$$

exemplos de variáveis aleatórias

$P_X(x) \geq 0$

$P_X(x) \doteq$ densidade de probabilidade

$P_X(x)\,dx$ = probabilidade de $x \in [x, x+dx]$

$\mathscr{A} \doteq [a, b]$

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

- Variáveis contínuas - $prob(X \in \mathscr{A}) = \int_{x \in \mathscr{A}} p_X(x)\,dx = \int_{\mathscr{X}} I(x \in \mathscr{A}) p_X(x)\,dx$, onde $I(s)$ é a função indicadora igual a 1 se a *afirmação s* é verdadeira e 0 se é falsa; $p_X(x) \geq 0$ é a densidade de probabilidade que está normalizada: $\int_{x \in \mathscr{X}} p_X(x)\,dx = 1$
- Exemplos: números pseudo-aleatórios
  - distribuição uniforme - $x \in \mathscr{X}$=[0,1], $p_X(x) = 1$. rand(1)
  - distribuição discreta uniforme: $x \in \mathscr{X}$={1,2, ..., M}, $p_X(x) = 1/M$ randi( )
  - distribuição Gaussiana- $x \in \mathscr{X} = [-\infty, \infty]$,$p_X(x) = \frac{1}{\sqrt{2\pi}\sigma} exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$; $\overline{X} = \mu$; $Var\, X = \sigma^2$

    $\mu = 0$
    $\sigma = 1$

    randn( )

mais exemplos

- distribuição exponencial - $x \in \mathscr{X} = [0, \infty], p_X(x) = \lambda \exp(-\lambda x)$. ver exercicio 2
- distribuição Bernoulli - $x \in \mathscr{X} = \{0,1\}$, $p_X(1) = p$, $p_X(0) = 1 - p$; $\overline{X} = p$; $Var\, X = p(1-p)$
- distribuição Binomial: $k \in \mathscr{X} = \{0, 1, \ldots, N\}$, $p_X(k) = \frac{N!}{k!(N-k)!} p^k (1-p)^{N-k}$. ver exercicio 11
- distribuição Poisson: $k \in \mathscr{X} = \{0, 1, 2, \ldots, \infty\}$, $p_X(k) = \exp(-\lambda) \frac{\lambda^k}{k!}$ $\bar{k} = \lambda$ ver exercicio 9

λ = taxa de decaimento radioativo

$\lambda\, dt$ = probabilidade de decaimento.

$$\overline{k} = \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} k = e^{-\lambda} \lambda \sum_{k=0}^{\infty} \frac{\lambda^{k-1}}{k!} k$$

$$k! = k(k-1)!$$

$$= e^{-\lambda} \lambda \left(\overset{(k=0)}{0} + \underbrace{\overset{(k=1)}{\frac{1}{0!}} + \overset{(k=2)}{\frac{\lambda}{1!}} + \overset{(k=3)}{\frac{\lambda^2}{2!}} \dots}_{e^{\lambda}} \right)$$

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$

$$\overline{k} = \lambda \qquad \simeq \text{ Gaussiana } \mu = \lambda$$

$$\text{Var } k = \overline{k^2} - \overline{k}^2 = \lambda \qquad \sigma^2 = \lambda$$

$$\overline{k^2} = \sum_{k=0}^{\infty} k^2 \frac{\lambda^k}{k!} e^{-\lambda}$$

$$= e^{-\lambda} \sum_{k=0}^{\infty} k^2 \frac{\lambda^k}{k!} = e^{-\lambda} \lambda \frac{d}{d\lambda} \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!}$$

$$\frac{d}{d\lambda} \lambda^k = k \lambda^{k-1}$$

$$= e^{-\lambda} \lambda \frac{d}{d\lambda} \left[\lambda \frac{d}{d\lambda} \underbrace{\sum_{k=0}^{\infty} \frac{\lambda^k}{k!}}_{e^{\lambda}} \right]$$

$$\lambda \frac{d}{d\lambda} \lambda^k = k \lambda^k$$

$$\overline{k^2} = e^{-\lambda} \lambda \frac{d}{d\lambda} \left[\lambda \frac{d}{d\lambda} e^{\lambda} \right] = e^{-\lambda} \lambda \frac{d}{d\lambda} \left[\lambda e^{\lambda} \right]$$

$$= e^{-\lambda} \lambda \left(1 \times e^{\lambda} + \lambda e^{\lambda} \right) = \lambda + \lambda^2$$

$$\mathrm{Var}\, k = \overline{k^2} - \overline{k}^2 = \lambda + \cancel{\lambda^2} - \cancel{\lambda^2} = \lambda$$

Histograma

$k = 0, 1, 2, \dots, \infty$

$k \equiv$ numero de nucleos que decairam num dado intervalo de tempo t

$0 \leq P_k \leq 1$

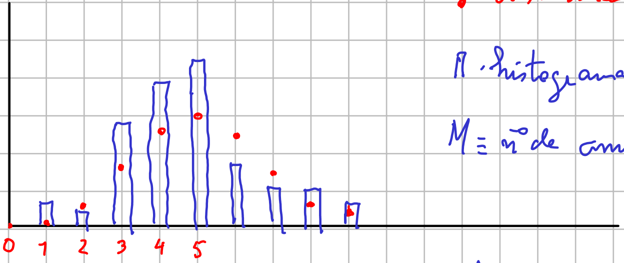
$$P_k = \frac{(\lambda t)^k}{k!} e^{-\lambda t}$$

$\lambda \equiv$ taxa de decaimento radioactivo

$\lambda \equiv \lambda t$

$\overline{k} = \lambda t$

Fazer M contagens $\quad k_v = k_1, k_2, k_3, \dots, k_M$



• distribuicao Poisson

▯ histograma normalizado

M = nº de amostras

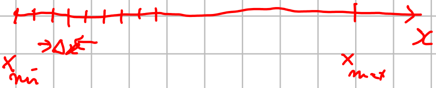
$[hv] = hist(k_v, 1:km)$

$hn = hv/sum(hv)$

↑

normalizado

bins



$P_X(x)$

↑

Variável contínua

$h_n = \left(hv/sum(hv)\right)/deltax$

$\Delta x = (x_{max} - x_{min})/N_{bins}$

métodos numéricos para gerar números aleatórios

- efeito de uma transformação de variáveis $Y = f(X)$
  - $p_Y(y) = \left|\frac{dx}{df}\right| p_X(x)$. Qual a função $f(x)$ que devemos usar para produzir $p_Y(y)$ desejado?
- método da transformação de variáveis

  $P_x(x)$

  - Definido a distribuição de probabilidade cumulativa, $F(y) = prob(X \leq y) = \int_{-\infty}^{y} p_X(x)\,dx$, calcula-se $y = F^{-1}(u)$ com U com distribuição uniforme - $u \in \mathscr{U}$=[0,1] e obtém-se Y com distribuição $p_X(y)$.
    - demonstração: Fazendo $u = F(y)$ temos $p_U(u)du = p_Y(y)dy$ e portanto $p_Y(y) = \left|\frac{du}{dy}\right| p_U(u)$. Dado que $\left|\frac{du}{dy}\right| = \left|\frac{d}{dy}F(y)\right|$, $p_U(u) = 1$ e $F(y) = \int_{-\infty}^{y} p_X(x)\,dx$ obtém-se $p_Y(y) = p_X(y)$ como se pretendia.

$u = F(y)$

$\boxed{y = F^{-1}(u)}$

métodos numéricos para gerar números aleatórios

- método da aceitação / rejeição

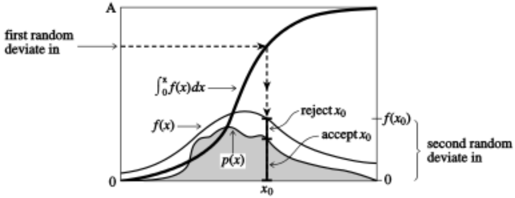
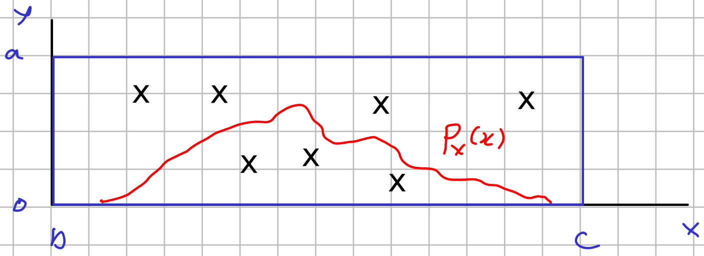


Figure 7.3.2. Rejection method for generating a random deviate x from a known probability distribution $p(x)$ that is everywhere less than some other function $f(x)$. The transformation method is first used to generate a random deviate x of the distribution f (compare Figure 7.3.1). A second uniform deviate is used to decide whether to accept or reject that x. If it is rejected, a new deviate of f is found, and so on. The ratio of accepted to rejected points is the ratio of the area under p to the area between p and f.

Figure: do livro Numerical Recipes

Metodo de aceitação /rejeição



$x = b + (c-b) * rand(1)$ $\qquad b \le x \le c$

$y = a * rand(1)$

Coordenada x dos pontos abaixo da curva $P_x(x)$ (aceites) tem densidade de probabilidade $P_x(x)$

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.