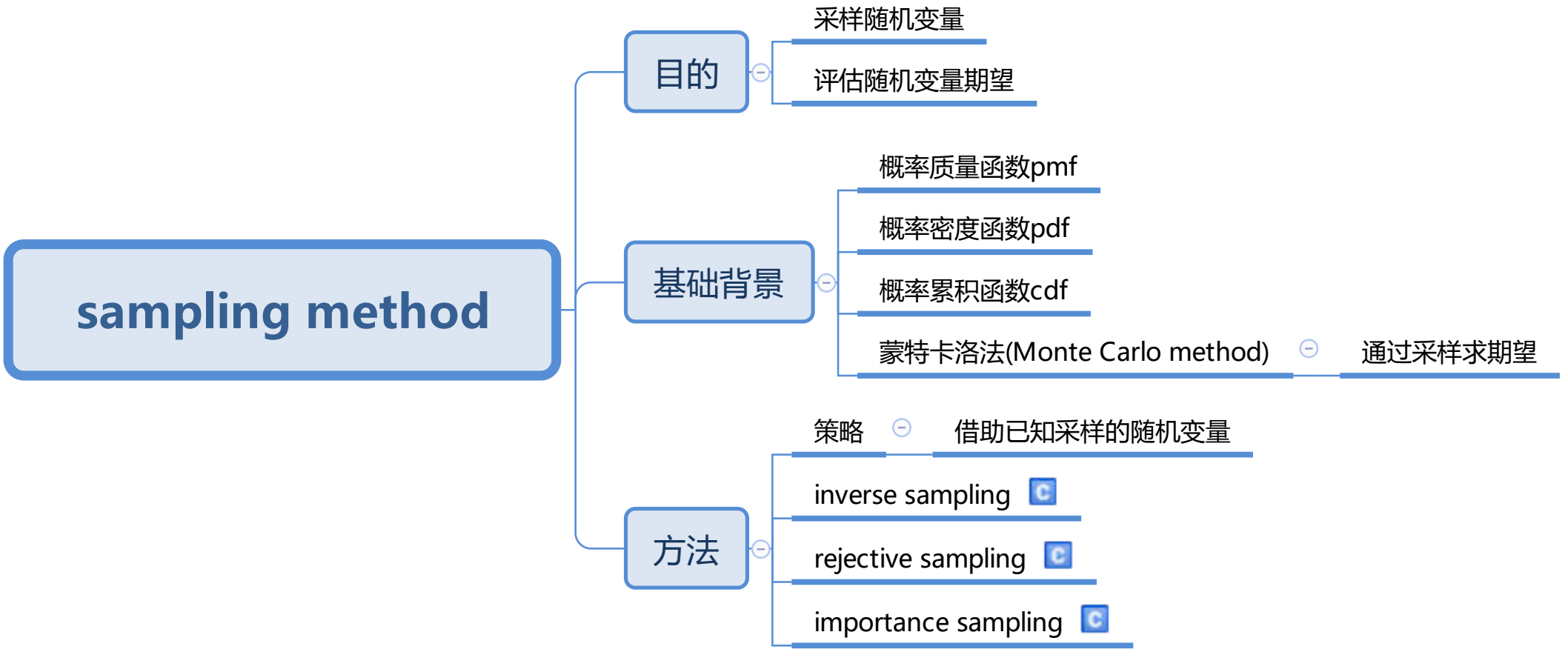

sampling method
目的
采样随机变量
评估随机变量期望
基础背景
概率质量函数pmf
概率密度函数pdf
概率累积函数cdf
蒙特卡洛法(Monte Carlo method)
通过采样求期望
方法
策略
借助已知采样的随机变量
inverse sampling
rejective sampling
importance sampling

inverse sampling

描述

- T 采样随机变量b
- S 已知b的cdf(F(x))
- I 如果b=T(a), a为已知采样的随机变量，则T实现
- A 选取已知采样a
  - 求可逆，单调递增函数Q(x)
  - 使H(Q(x))=F(x), 其中H(x)为a的cdf
- R Q(-1)(a) = b，即采样b，为采样a，然后变换为Q(-1)(a)
- P 需要知道b的cdf
  - 需要求满足条件的Q(x)

推导

已知 $P(a \le x) = H(x), P(b \le x) = F(x), Q(x)$单调递增，且可逆，$H(Q(x)) = F(x)$

则 $P(Q^{-1}(a) \le x) = P(a \le Q(x)) = H(Q(x)) = F(x) = P(b \le x)$

即$Q^{-1}(a) = b$

实例

$$P(a \le x) = H(x) = \begin{cases} 1 & ,x \ge 1 \\ x & ,0 \le x \le 1 \\ 0 & ,x < 0 \end{cases}$$

a为[0,1]均匀分布随机变量

Q(x)为b的cdf，F(x) $H(F(x)) = F(x)$

对于已知cdf且cdf可逆的随机变量b，都可依靠[0,1]均匀分布随机变量进行采样

rejective sampling

描述

- T 采样随机变量b
- S 已知b的pdf，f(x)的值或是等比例值
- I f(xi)的值越大，则采样值出现xi的可能性越大
- A
  - 选取已知采样a，a的值空间Va包含b的值空间Vb
  - 求常数Z，使得Z*h(xj)> =f(xj),xj属于Vb，h(xi)为a的pdf在xi的值
  - 对a采样xi
- R 依f(xi)/Z/h(xi)的概率决定是否接受xi为b的采样
- P
  - 如果f(xi)比h(xi)大很多，会导致对b很重要的xi采样缺失
    即在整个值空间，需要使f(x)不比h(xi)大很多
  - 如果f(xi)/Z/h(xi)太小，则xi总是被拒绝，
    即在整个值空间，需要使f(x)比Z*h(xi)不要小太多

推导

根据蒙特卡洛法，随机变量的充分采样的采样直方图外轮廓曲线近似于随机变量的pdf或pmf曲线即（图1）

当随机变量a在任意点的采样次数都多于随机变量b时，将a，b采样直方图画于同一（图2）中。

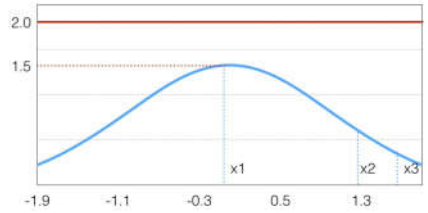
由图2可知，采样a得到xi，有e(x)/d(x) 的可能性xi也为b的采样

由蒙特卡洛法可知,$f(x) = Z_a * e(x), h(x) = Z_b * d(x), Z_x, Z_a$为常数

$$\frac{e(x)}{d(x)} = \frac{f(x)}{Zh(x)}, Z = Z_b/Z_b$$

实例

a 为值空间内的均匀分布



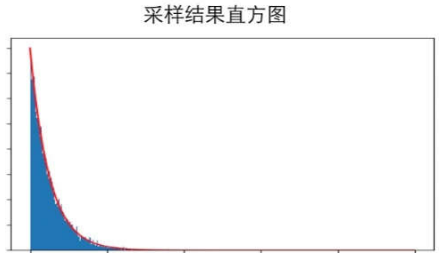


图1

蓝色的采样直方图y坐标和红色的pdf曲线值为不同尺度。

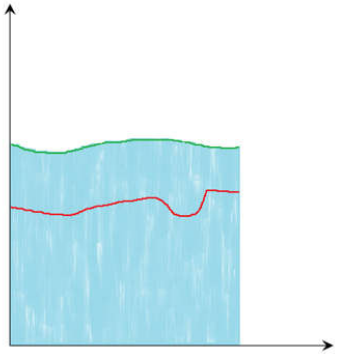


图2

绿线为a的采样轮廓线d(x)，红线为b的采样轮廓线e(x)，d(x)>b(x)。

importance sampling

- 描述
  - T
    - 求随机变量b函数的期望
  - S
    - 已知b的pdf，f(x)的值或是等比例值Z*f(x),Z为未知常数
  - I
    - $\mathrm{E}(\mathrm{g}(\mathrm{x})) = \int g(x)f(x)dx = \int g(x)\frac{\tilde{f}(x)}{Zh(x)}h(x)dx, \tilde{f}(x) = Zf(x)$
    - $\mathrm{Z} = \int \tilde{f}(x)dx = \int \frac{\tilde{f}(x)}{h(x)}h(x)dx$
  - A
    - 选取随机变量a
    - 采样a
  - R
    - 随机变量 a 的 $E\left(g(x)\frac{\tilde{f}(x)}{E\left(\frac{\tilde{f}(x)}{h(x)}\right)h(x)}\right)$
  - P
    - 随机变量a的pdf,h(x)需要满足下列要求
    - 在值域范围内，$g(x)\tilde{f}(x)$或$\tilde{f}(x)$的值较大时，h(x)的值不能太小
    - 不然导致采样得到的期望误差太大
- 推导
  - 需要推导，怎么选取h(x)，暂时空缺
- 实例
  - 机器学习中，softmax函数输出特别多，反向传播求导时其中一部分为pmf求期望，a选用n_gram模型