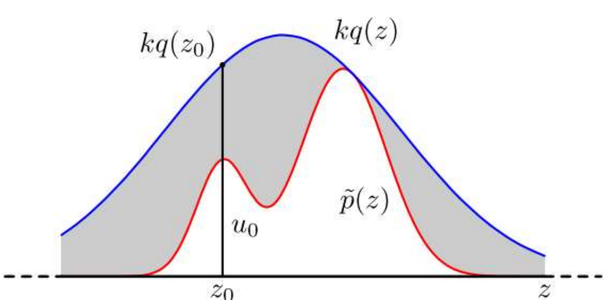
采样方法初步

> Written by Native.S1mple

1.**蒙特卡罗方法**：赌场名，随机模拟的方法，类似赌场扔骰子的过程，最早用于求解微积分：$\theta = \int_a^b f(x)dx$
简单近似$[a,b]$取一点x_0,$\theta = (b-a)f(x_0)$,当然仅有一个点过于粗糙，我们可采样$[a,b]$区间上n个点$\Rightarrow \theta = \frac{b-a}{n}\sum_{i=0}^{n-1} f(x_i)$
上述办法仅仅适用于x在$[a,b]$上均匀分布，不均匀则误差大。
加入我们知道x在$[a,b]$的概率分布函数$p(x) \Rightarrow \theta = \int_a^b f(x)dx = \int_a^b \frac{f(x)}{p(x)}p(x)dx \approx \frac{1}{n}\sum_{i=0}^{n-1} \frac{f(x_i)}{p(x_i)}$.
若x在$[a,b]$均匀分布，可代入$p(x) = \frac{1}{b-a}$即之前式子。
2.**概率分布采样**
如果有x的概率分布函数，如何采样:对于常见的均匀分布$uniform(0,1)$随机生成器即可，其他常见的概率分布，无论离散连续，均可通过$uniform(0,1)$样本转换得到。
例：二维正态分布样本(Z_1,Z_2)可通过独立采样得到的$uniform(0,1)$样本对(x_1,x_2)转化:$Z_1 = \sqrt{-2lnX_1}cos(2\pi X_2) \quad Z_2 = \sqrt{-2lnX_1}sin(2\pi X_2)$可调库。
3.**接受-拒绝采样**(ACP-REJ)
若$p(x)$不常见甚至过于复杂，我们设定一个可采样的分布$q(x)$，然后按照一定的方法拒绝某些样本，以达到接近$p(x)$分布的目的，其中$q(x)$称为$proposal\ distribution$
　设定一个方便采样的函数$q(x)$以及一个常量k，使得$p(x)$总在$kq(x)$之下。首先采样得到的$q(x)$的一个样本z,采样方法如上一节，然后从均匀分布$(0.kq(z_0)$中采样得到一个值u,若u落在下图灰色区域则拒绝，否则接受。重复操作接受n个样本$z_0,z_1 \ldots\ldots z_{n-1}$则最后蒙特卡罗求解结果为:
$\frac{1}{n}\sum_{i=0}^{n-1}\frac{f(z_i)}{p(z_i)} \rightarrow kq(z_i)$



4.**缺陷**

- 对于一些二维分布$p(x,y)$容易得$p(x|y),p(y|x)$，但是难以得到$p(x,y)$一般形式。
- 对于一些高维度复杂非常见分布$p(x_1,x_2,\ldots\ldots x_n)$找一个合适的$q(x)$和$k$很困难。