

介绍

- 动机和历史背景
 - 蒙特卡罗方法是一种通过随机数解决确定性问题的方法。它使用反复随机抽样来估算数值答案，适用于许多难以分析或数值难以解决的问题。
 - 蒙特卡罗方法由恩里科·费米、约翰·冯·诺依曼和斯坦尼斯拉夫·乌拉姆在洛斯阿拉莫斯国家实验室曼哈顿计划期间开发。
 - 主要应用领域：物理、化学、生物学、自然科学、社会科学、机器学习、贝叶斯统计、信号处理、优化理论等。

逆变换方法

逆变换方法的目的是通过已知的均匀分布随机数生成其他复杂分布的随机数。这种方法常用于模拟和蒙特卡罗方法中，特别是在计算机科学和工程领域，用来解决各种复杂的概率和统计问题。

核心原理 逆变换方法基于以下核心原理：

- 均匀分布**：假设我们可以从均匀分布 $U(0,1)$ 生成随机值。均匀分布的概率密度函数 (PDF) 为 $f(x)$ ，在 $(0,1)$ 区间内为常数。
- 累积分布函数 (CDF)**：给定任意连续随机变量 X ，其累积分布函数 $F(x)$ 定义为 $F(x)=P(X\leq x)$ 。累积分布函数 $F(x)$ 将 x 的支撑区间映射到 $[0,1]$ 区间。
- 逆变换**：通过 $X = F^{-1}(U)$ ，我们可以将均匀分布的随机变量 U 转换为符合目标分布 $F(x)$ 的随机变量 X 。

实现思路 实现逆变换方法的步骤如下：

- 生成均匀分布随机数**：从 $U(0,1)$ 间生成随机数 U 。
- 计算逆累积分布函数**：通过累积分布函数 $F(x)$ 的逆函数 $F^{-1}(u)$ ，将均匀分布随机数 U 转换为目标分布的随机数 X 。
- 输出结果**：最终得到的 X 是符合目标分布的随机数。

数学推导 假设 X 是目标随机变量， $F(x)$ 是其累积分布函数。我们希望通过均匀分布的随机变量 U 生成 X 。具体推导如下：

- $U \sim U(0,1)$ ，即 U 在 $(0,1)$ 间均匀分布。
- $X = F^{-1}(U)$ ，即 X 是通过 U 逆变换得到的。
- 由于 U 是均匀分布， $P(U \leq u) = u$ 。
- 因此， $P(X \leq x) = P(F^{-1}(U) \leq x) = P(U \leq F(x)) = F(x)$ 。

举例说明 假设我们希望生成指数分布 X 的随机变量， X 的累积分布函数为： $F(x) = 1 - e^{-\lambda x}$ 其逆函数为： $F^{-1}(u) = -\frac{1}{\lambda} \ln(1 - u)$

实现步骤：

- 从均匀分布 $U(0,1)$ 生成随机数 U 。
- 计算 $X = F^{-1}(U) = -\frac{1}{\lambda} \ln(1 - U)$ 。
- 输出 X 。

通过上述步骤，我们将均匀分布的随机数 U 转换为指数分布的随机数 X 。

优势

简单易懂，易于实现，适用于各种分布，生成随机数非常高效。

- **简单高效**：一旦获得逆累积分布函数，生成随机数非常简便。
- **通用性强**：适用于多种连续分布，只要目标分布的累积分布函数 $F(x)$ 已知。

局限性

- **依赖逆函数**：需要目标分布的累积分布函数有解析形式的逆函数，对于一些复杂分布，逆函数可能难以求出。
- **计算复杂性**：对于某些分布，计算逆函数可能涉及复杂的数值计算。

拒绝采样方法

拒绝采样方法是一种从复杂目标分布 $f(x)$ 生成随机样本的技术，使用一个更容易采样的提议分布 $g(x)$ 来实现。这种方法广泛应用于统计计算和蒙特卡罗模拟中，特别是在难以直接从目标分布中采样的情况下。

核心原理 拒绝采样方法基于以下核心原理：

1. **提议分布 $g(x)$** ：选择一个易于采样的分布 $g(x)$ ，称为提议分布，要求它的支撑包含目标分布 $f(x)$ 的支撑。
2. **缩放因子 M** ：找到一个常数 M ，使得对所有 x ，都有 $Mg(x) \geq f(x)$ 。这个缩放因子确保 $g(x)$ 的某个倍数完全包住目标分布 $f(x)$ 。
3. **接受率**：对于每个从提议分布 $g(x)$ 中生成的样本 x ，根据 $f(x)/(Mg(x))$ 的比值决定是否接受该样本。

实现思路 实现拒绝采样方法的步骤如下：

1. **选择提议分布 $g(x)$** ：选择一个易于采样的分布 $g(x)$ 作为提议分布。
2. **确定缩放因子 M** ：找到一个常数 M ，使得 $Mg(x) \geq f(x)$ 对所有 x 都成立。
3. 生成样本并计算接受率：
 1. 从提议分布 $g(x)$ 中生成一个样本 x 。
 2. 从均匀分布 $U(0, Mg(x))$ 中生成一个随机数 y 。
 3. 如果 $y \leq f(x)$ ，则接受 x 作为目标分布 $f(x)$ 的样本；否则，拒绝 x 并返回步骤 (3.1)。

数学推导 假设目标分布为 $f(x)$ ，提议分布为 $g(x)$ ，并且满足 $Mg(x) \geq f(x)$ 对所有 x 都成立。我们通过以下步骤进行推导：

1. 从提议分布 $g(x)$ 生成样本 x 。
2. 计算接受率 $\alpha(x) = \frac{f(x)}{Mg(x)}$ 。
3. 从均匀分布 $U(0, Mg(x))$ 中生成随机数 y 。
4. 如果 $y \leq f(x)$ ，接受 x 作为样本；否则拒绝 x 并重新生成。

举例说明 假设我们希望生成符合目标分布 $f(x) = \frac{1}{2}e^{-|x|}$ 的随机样本，该分布在 $(-\infty, \infty)$ 上具有双指数形状。我们选择提议分布 $g(x) = e^{-x}$ （单侧指数分布），并确定缩放因子 M 如下：

1. 目标分布 $f(x)$ 在 $x \geq 0$ 时为 $\frac{1}{2}e^{-x}$ ，而提议分布 $g(x) = e^{-x}$ 。因此，在 $x \geq 0$ 时，有 $Mg(x) \geq f(x)$ 成立，其中 $M = 2$ 。
2. 对于 $x < 0$ ，我们有 $f(x) = \frac{1}{2}e^x$ ，而 $g(x) = e^{-x}$ 不再有效。因此，我们需要找到一个不同的提议分布 $g(x)$ ，使得 $Mg(x) \geq f(x)$ 在 $x < 0$ 成立。可以选择 $g(x) = e^x$ ，对应的缩放因子 $M = 2$ 。

算法实现

1. 选择提议分布 $g(x)$ 和确定缩放因子 M 。

- 从提议分布 $g(x)$ 生成样本 x 。
- 从均匀分布 $U(0, Mg(x))$ 生成随机数 y 。
- 如果 $y \leq f(x)$, 接受 x 作为样本; 否则拒绝 x 并返回步骤 (2)。

优势

- 灵活性:** 适用于任何目标分布, 只要提议分布 $g(x)$ 和缩放因子 M 满足条件。
- 易于实现:** 实现简单, 只需选择适当的提议分布和确定缩放因子。

局限性

- 效率低下:** 当目标分布和提议分布相差较大时, 接受率可能非常低, 导致大量样本被拒绝。
- 高维问题:** 在高维空间中找到合适的提议分布和缩放因子 M 变得更加困难和复杂。

总结 拒绝采样方法是一种强大且灵活的随机数生成技术, 特别适用于无法直接采样的复杂目标分布。尽管在某些情况下效率可能较低, 但其简单易实现的特性使其在许多应用中广泛使用。通过合理选择提议分布和缩放因子, 可以显著提高采样效率。

经典蒙特卡罗积分

经典蒙特卡罗积分的目标是通过随机抽样来近似计算复杂函数的积分值, 特别是在高维空间或不规则区域上, 这些问题通常很难通过解析或传统数值方法求解。

核心原理 蒙特卡罗积分基于以下核心原理:

- 随机采样:** 使用随机数生成器从定义域内生成一系列随机样本。
- 平均值估计:** 计算函数在这些随机样本点上的值, 并求其平均值来估计积分值。
- 大数定律:** 随着样本数量的增加, 样本平均值将收敛于积分的真实值。

实现思路 实现蒙特卡罗积分的方法步骤如下:

- 定义积分问题: 假设需要计算函数 $h(x)$ 在区间 $[a, b]$ 上的积分:

$$I = \int_a^b h(x) dx$$

- 随机采样:** 在区间 $[a, b]$ 内随机生成 n 个样本点 x_1, x_2, \dots, x_n , 这些样本点均匀分布。
- 函数评估:** 计算每个样本点上的函数值 $h(x_1), h(x_2), \dots, h(x_n)$ 。
- 估计积分值: 根据大数定律, 积分值可以通过以下公式估计:

$$I \approx \frac{b-a}{n} \sum_{i=1}^n h(x_i)$$

其中, $\frac{b-a}{n}$ 是区间长度的缩放因子。

数学推导 假设需要计算函数 $h(x)$ 在区间 $[a, b]$ 上的积分 I :

$$I = \int_a^b h(x) dx$$

将积分表示为期望值形式:

$$I = (b-a) \mathbb{E}[h(X)]$$

其中, X 是均匀分布在 $[a, b]$ 上的随机变量。因此, 通过生成 X 的样本 x_1, x_2, \dots, x_n , 并计算函数 $h(x)$ 的平均值, 可以近似估计积分值:

$$I \approx (b-a) \frac{1}{n} \sum_{i=1}^n h(x_i)$$

举例说明 假设需要计算函数 $h(x) = x^2$ 在区间 $[0,1]$ 上的积分：

$$I = \int_0^1 x^2 dx$$

通过蒙特卡罗积分方法：

- 随机采样**：在 $[0,1]$ 内随机生成 n 个样本点 x_1, x_2, \dots, x_n 。
- 函数评估**：计算每个样本点上的函数值 $h(x_i) = x_i^2$ 。
- 估计积分值**：根据公式 $I \approx \frac{1}{n} \sum_{i=1}^n x_i^2$ 。

优势

- 处理高维积分**：蒙特卡罗积分在处理高维积分时具有显著优势，避免了传统数值积分方法在高维空间中的计算爆炸。
- 适应复杂形状**：随机采样方法对积分域的形状不敏感，可以处理不规则形状的积分域。
- 易于实现**：算法简单，实现方便，只需随机数生成和函数评估。

局限性

- 收敛速度慢**：蒙特卡罗积分的收敛速度较慢，通常需要大量样本才能获得较高精度的估计。
- 高方差问题**：在某些情况下，随机样本可能会导致较高的估计方差，影响结果的稳定性。

总结 经典蒙特卡罗积分是一种基于随机采样的数值积分方法，通过生成均匀分布的随机样本并计算函数值的平均来估计积分值。尽管在处理高维积分和不规则积分域方面具有显著优势，但其收敛速度较慢，需要大量样本来提高估计精度。蒙特卡罗积分广泛应用于科学、工程和金融等领域，特别是在传统数值方法难以应用的复杂积分问题中。

重要性采样

重要性采样是一种用于估计期望值的方法，特别适用于当直接从目标分布 $f(x)$ 采样不可行或效率低下的情况。通过从更易采样的提议分布 $g(x)$ 生成样本，并对这些样本进行加权，来估计目标分布下的期望值。

核心原理 重要性采样基于以下核心原理：

- 提议分布 $g(x)$** ：选择一个更易于采样的分布 $g(x)$ ，要求其支撑包含目标分布 $f(x)$ 的支撑。
- 加权样本**：使用样本在目标分布 $f(x)$ 和提议分布 $g(x)$ 下的概率密度比值来调整样本的权重，从而修正因使用提议分布采样引入的偏差。
- 加权平均值**：通过加权平均值来估计目标分布下的期望值。

实现思路 实现重要性采样的方法步骤如下：

- 选择提议分布 $g(x)$** ：选择一个易于采样的分布 $g(x)$ 作为提议分布，要求其支撑包含目标分布 $f(x)$ 的支撑。
- 生成样本**：从提议分布 $g(x)$ 中生成 n 个样本 x_1, x_2, \dots, x_n 。
- 计算权重**：计算每个样本的权重 $w(x_i) = \frac{f(x_i)}{g(x_i)}$ 。
- 估计期望值**：使用加权平均值来估计目标分布下的期望值 $\mathbb{E}_f[h(X)]$ ，其中， $h(x)$ 是需要估计其期望值的函数：

$$\mathbb{E}_f[h(X)] \approx \frac{1}{n} \sum_{i=1}^n h(x_i) w(x_i)$$

数学推导 假设我们要估计目标分布 $f(x)$ 下函数 $h(x)$ 的期望值：

$$\mathbb{E}_f[h(X)] = \int h(x) f(x) dx$$

引入提议分布 $g(x)$ ，并重写期望值的表达式：

$$\mathbb{E}_f[h(X)] = \int h(x) \frac{f(x)}{g(x)} g(x) dx = \mathbb{E}_g \left[h(X) \frac{f(X)}{g(X)} \right]$$

这样，通过从 $g(x)$ 生成样本，并计算每个样本的权重 $\frac{f(x)}{g(x)}$ ，我们可以使用加权平均值来估计期望值：

$$\mathbb{E}_f[h(X)] \approx \frac{1}{n} \sum_{i=1}^n h(x_i) \frac{f(x_i)}{g(x_i)}$$

举例说明 假设我们希望估计目标分布 $f(x) = \frac{1}{2}e^{-|x|}$ 下函数 $h(x) = x^2$ 的期望值。我们选择提议分布 $g(x) = e^{-x}$ (单侧指数分布)，并进行如下步骤：

- 选择提议分布 $g(x)$ ：**选取单侧指数分布 $g(x) = e^{-x}$ ，其支撑为 $[0, \infty)$ 。
- 生成样本：**从 $g(x)$ 中生成 n 个样本 x_1, x_2, \dots, x_n 。
- 计算权重：**对于每个样本 x_i ，计算权重 $w(x_i) = \frac{f(x_i)}{g(x_i)} = \frac{\frac{1}{2}e^{-x_i}}{e^{-x_i}} = \frac{1}{2}$ 。
- 估计期望值：使用加权平均值估计期望值：

$$\mathbb{E}_f[h(X)] \approx \frac{1}{n} \sum_{i=1}^n x_i^2 \cdot \frac{1}{2} = \frac{1}{2n} \sum_{i=1}^n x_i^2$$

优势

- 减少方差：**通过选择合适的提议分布 $g(x)$ ，可以集中样本在 $f(x)h(x)$ 重要的区域，从而减少方差，提高估计精度。
- 灵活性：**适用于任何目标分布 $f(x)$ ，只要能找到合适的提议分布 $g(x)$ 。

局限性

- 需要深度知识：**选择合适的提议分布 $g(x)$ 需要对目标分布 $f(x)$ 和函数 $h(x)$ 有深刻的理解。
- 计算复杂性：**计算权重 $\frac{f(x)}{g(x)}$ 可能涉及复杂的计算，特别是对于高维问题。

总结 重要性采样是一种有效的蒙特卡罗方法，通过从更易采样的提议分布生成样本，并对这些样本进行加权来估计目标分布下的期望值。尽管在某些情况下需要深厚的领域知识和复杂的计算，但其减少方差、提高估计精度的特性使其在许多应用中广受欢迎，特别是在高维积分和稀有事件模拟中。

马尔科夫链蒙特卡罗 (MCMC) 方法

马尔科夫链蒙特卡罗 (MCMC) 方法的目標是通过构建一个马尔科夫链，其平稳分布与目标分布一致，从而生成符合目标分布的样本。这种方法广泛应用于贝叶斯统计、计算物理和机器学习等领域，尤其是在高维复杂分布的采样中。

核心原理 MCMC方法基于以下核心原理：

- 马尔科夫链：**MCMC方法通过设计一个马尔科夫链，使其状态随着时间的推移逐渐覆盖目标分布的各个区域。

2. **平稳分布**：马尔科夫链的转移概率设计得当，使得随着时间的推移，链的状态分布趋向于目标分布，即达到平稳状态。
3. **依赖性采样**：与独立采样不同，MCMC方法利用前一个状态的信息生成下一个状态，从而提高采样效率。

实现思路 实现MCMC方法的步骤如下：

1. **初始化状态**：选择一个初始状态 x_0 。
2. **构建马尔科夫链**：根据转移概率从当前状态生成下一个状态，重复此过程，构建状态序列。
3. **采样**：当链达到平稳分布后，从序列中提取样本用于估计。

常见的MCMC算法包括Metropolis-Hastings算法和Gibbs采样。

Metropolis-Hastings算法

1. **目标**：生成符合目标分布 $\pi(x)$ 的样本。
2. **步骤**：
 1. 选择初始状态 x_0 。
 2. 对于第 i 步：
 1. 从提议分布 $q(x'|x_i)$ 中生成候选样本 x' 。
 2. 计算接受率： $\alpha = \min\left(1, \frac{\pi(x')q(x_i|x')}{\pi(x_i)q(x'|x_i)}\right)$
 3. 生成均匀随机数 $u \sim U(0, 1)$ 。
 4. 如果 $u < \alpha$ ，接受 x' 作为新状态 x_{i+1} ；否则，保留当前状态 $x_{i+1} = x_i$ 。
3. **优点**：易于实现，适用于各种复杂分布。
4. **缺点**：在高维空间中，选择合适的提议分布较为困难，收敛速度可能较慢。

Gibbs采样

1. **目标**：在多维空间中逐个更新每个变量，使得每个变量的条件分布达到目标分布。
2. **步骤**：
 1. 选择初始状态 $x^{(0)} = (x_1^{(0)}, x_2^{(0)}, \dots, x_d^{(0)})$ 。
 2. 对于第 t 次迭代：
 1. 从条件分布 $\pi(x_1|x_2^{(t-1)}, x_3^{(t-1)}, \dots, x_d^{(t-1)})$ 中生成 $x_1^{(t)}$ 。
 2. 从条件分布 $\pi(x_2|x_1^{(t)}, x_3^{(t-1)}, \dots, x_d^{(t-1)})$ 中生成 $x_2^{(t)}$ 。
 3. 重复上述过程，直到从条件分布 $\pi(x_d|x_1^{(t)}, x_2^{(t)}, \dots, x_{d-1}^{(t)})$ 中生成 $x_d^{(t)}$ 。
3. **优点**：每步采样容易实现，适合高维空间。
4. **缺点**：条件分布的推导和计算可能复杂，变量之间存在高度相关性时收敛速度较慢。

数学推导 假设目标分布为 $\pi(x)$ ，我们希望通过MCMC方法生成符合 $\pi(x)$ 的样本。设马尔科夫链的状态转移矩阵为 P ，我们需要确保 P 使得 $\pi(x)$ 是其平稳分布，即满足平稳分布方程：

$$\pi(x)P = \pi(x)$$

为了实现这一目标，Metropolis-Hastings算法引入了接受率 α ，确保详细平衡条件成立：

$$\pi(x)P(x \rightarrow x') = \pi(x')P(x' \rightarrow x)\pi(x)$$

具体步骤如下：

1. 从提议分布 $q(x'|x)$ 中生成候选样本 x' 。
2. 计算接受率: $\alpha = \min\left(1, \frac{\pi(x')q(x|x')}{\pi(x)q(x'|x)}\right)$
3. 生成均匀随机数 $u \sim U(0, 1)$ 。
4. 如果 $u < \alpha$, 接受 x' 作为新状态; 否则, 保留当前状态。

举例说明 假设我们希望从目标分布 $\pi(x)$ 中生成样本, 该分布是两个正态分布的混合:

$$\pi(x) = 0.5\mathcal{N}(-2, 1) + 0.5\mathcal{N}(2, 1)$$

使用Metropolis-Hastings算法的步骤如下:

1. 选择初始状态 $x_0 = 0$ 。
2. 设提议分布为 $q(x'|x) = \mathcal{N}(x, \sigma^2)$, 其中 σ 是步长。
3. 生成候选样本 x' 。
4. 计算接受率 α 。
5. 决定是否接受候选样本。

吉布斯采样

假设我们有二维正态分布 $\pi(x, y)$ 的目标分布, 条件分布为:

$$\pi(x|y) \sim N(\mu_{x|y}, \sigma_{x|y}^2), \quad \pi(y|x) \sim N(\mu_{y|x}, \sigma_{y|x}^2)$$

1. 初始化 (x_0, y_0) 。
2. 对于 $t = 1, 2, \dots, T$:
 1. 从 $\pi(x|y_t)$ 生成 x_{t+1} 。
 2. 从 $\pi(y|x_{t+1})$ 生成 y_{t+1} 。
3. 记录样本 (x_{t+1}, y_{t+1}) 。

优势

- **处理高维问题**: MCMC 方法特别适合处理高维复杂概率分布的问题。
- **灵活性**: 适用于多种分布, 只需设计合适的马尔科夫链。
- **渐进收敛**: 随着样本数量的增加, 样本分布会收敛到目标分布。

局限性

- **收敛速度慢**: 初始样本可能需要较长时间才能达到平衡分布。
- **依赖提议分布**: 选择合适的提议分布和参数调优是关键, 尤其在高维问题中。
- **混合性问题**: 链的混合性较差时, 可能会导致样本相关性高, 影响结果的有效性。

优势

- **适用性广泛**: MCMC方法可以应用于复杂、高维的概率分布采样。
- **灵活性强**: 不同的MCMC算法可以适应不同类型的目标分布和应用场景。

局限性

- **收敛性问题**: 判断马尔科夫链是否达到平稳分布较为困难, 收敛速度可能较慢。
- **依赖性问题**: 生成的样本之间存在相关性, 需要更多的样本来确保独立性。

总结 马尔科夫链蒙特卡罗 (MCMC) 方法是一种强大的采样技术，通过构建一个马尔科夫链，使其平稳分布与目标分布一致，从而生成符合目标分布的样本。尽管其实现和应用可能面临一些挑战，但在处理复杂、高维的概率分布问题时，MCMC方法仍然是一种不可或缺的工具。