

蒙特卡罗方法

续本达

清华大学 工程物理系

2023-11-01

蒙特卡罗方法

续本达

引言

复习与概念
引入

定义

舍选法

马尔可夫链法

逆函数法

总结

引言



两弹一星题材《马兰花开》剧照



《马兰花开》脚本

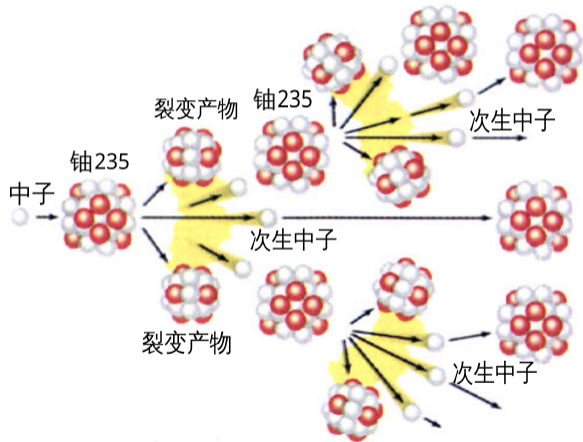
理论组：老邓，第八次计算结果与第九次完全吻合！

邓稼先：中国的第一颗原子弹，就从这里开始吧！

九次计算耗时半年，最终结果一致。
纠正了苏联专家的计算错误。

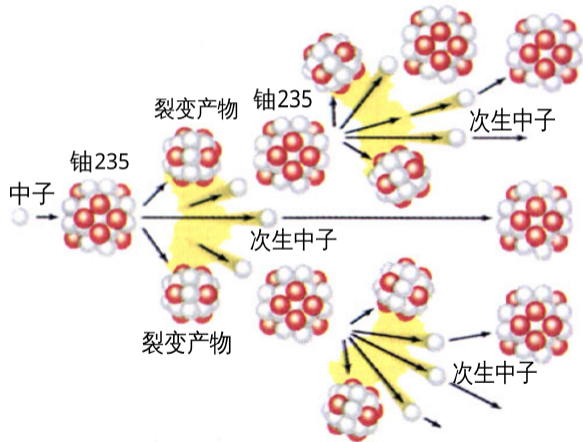
- 核裂变链式反应过程复杂，各步伴**随机性**；
- 中子在介质里输运被**随机**散射损耗能量；
- 最终得到**确定**的“试爆成功”结果。

掌握和运用核物理，需要大量的**计算**。



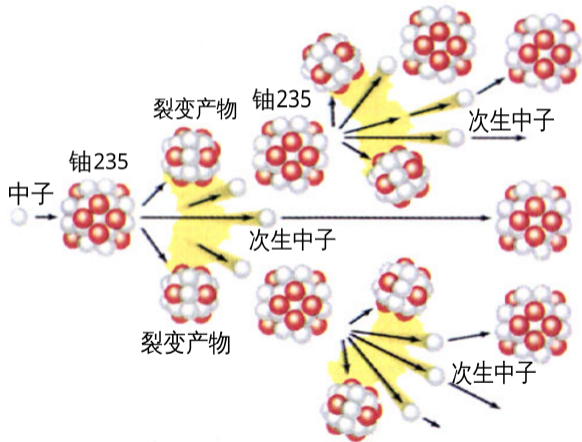
- 核裂变链式反应过程复杂，各步伴**随机性**；
- 中子在介质里输运被**随机**散射损耗能量；
- 最终得到**确定**的“试爆成功”结果。

掌握和运用核物理，需要大量的**计算**。



- 核裂变链式反应过程复杂，各步伴**随机性**；
- 中子在介质里输运被**随机**散射损耗能量；
- 最终得到**确定**的“试爆成功”结果。

掌握和运用核物理，需要大量的**计算**。



蒙特卡罗方法

续本达

引言

复习与概念
引入

定义

舍选法

马尔可夫链法

逆函数法

总结

复习与概念引入

算术平均值依概率收敛于数学期望

设 X_1, X_2, \dots, X_n 是独立同分布的随机变量序列，那么

$$\underbrace{\frac{1}{n} \sum_{i=1}^n g(X_i)}_{\text{算术平均值}} \xrightarrow[n \rightarrow \infty]{P} \underbrace{\int g(x) f_X(x) dx}_{g(X)\text{的数学期望}}$$

$f_X(x)$ 是 X_i 的概率密度函数， $g(x)$ 是 x 的任意函数。

$$\left. \begin{array}{l} g(x) = \sin x, \quad x \in [0, \pi] \\ f_X(x) = \frac{1}{\pi}, \quad \text{均匀分布} \end{array} \right\} \implies \int_0^{\pi} g(x) f_X(x) dx = \frac{2}{\pi} = 0.64$$

生成 $[0, 1]$ 上均匀分布的随机数 X_i , 计算 $\frac{1}{n} \sum_{i=1}^n g(X_i)$

- 算术平均值依概率收敛于数学期望

$$\frac{1}{n} \sum_{i=1}^n \sin X_i \xrightarrow[n \rightarrow \infty]{P} \int_0^{\pi} \sin x \underbrace{f_X(x)}_{1/\pi} dx = \frac{2}{\pi}$$

- 反向应用：能否把待求量构造成数学期望，使用算术平均近似？

$$\pi \approx \frac{2n}{\sum_{i=1}^n \sin X_i}$$

定义 (蒙特卡罗方法)

使用随机数解决计算问题的方法统称为蒙特卡罗方法。

- 算术平均值依概率收敛于数学期望

$$\frac{1}{n} \sum_{i=1}^n \sin X_i \xrightarrow[n \rightarrow \infty]{P} \int_0^{\pi} \sin x \underbrace{f_X(x)}_{1/\pi} dx = \frac{2}{\pi}$$

- 反向应用：能否把待求量构造成数学期望，使用算术平均近似？

$$\pi \approx \frac{2n}{\sum_{i=1}^n \sin X_i}$$

定义 (蒙特卡罗方法)

使用随机数解决计算问题的方法统称为蒙特卡罗方法。

蒙特卡罗方法

续本达

引言

复习与概念
引入

定义

舍选法

马尔可夫链法

逆函数法

总结

定义

使用 **随机数** 来解决 **计算问题** 方法的统称。

系统性地“盲人摸象”

- ① 随机变量的算术均值：用几个典型数 X_i ，刻画 \sin 函数的特征。
- ② 在可能的取值非常多，函数非常复杂时，随机取典型数，比确定地取要好。



延伸阅读：为什么蒙特卡罗方法有效？

- 姚（期智）准则 (Yao's principle)：给算法合理加入随机因素，会有效提升效率。（参考文献 4）
 - 蒙特卡罗方法为物理学中 **确定性的困难问题** 提供了 **仅有的可操作** 解法。

使用 **随机数** 来解决 **计算问题** 方法的统称。

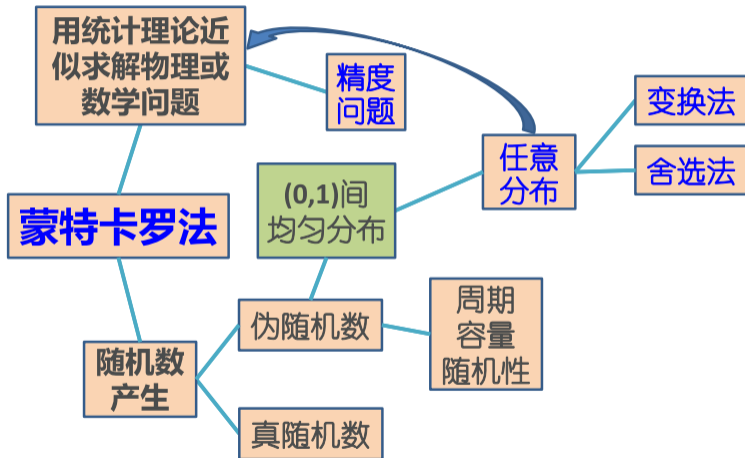
系统性地“盲人摸象”

- ① 随机变量的算术均值：用几个典型数 X_i ，刻画 \sin 函数的特征。
- ② 在可能的取值非常多，函数非常复杂时，随机取典型数，比确定地取要好。



延伸阅读：为什么蒙特卡罗方法有效？

- 姚（期智）准则 (Yao's principle)：给算法合理加入随机因素，会有效提升效率。（参考文献 4）
 - 蒙特卡罗方法为物理学中 **确定性的困难问题** 提供了 **仅有的可操作** 解法。



蒙特卡罗方法

续本达

引言

复习与概念
引入

定义

舍选法

马尔可夫链法

逆函数法

总结

舍选法

$$\pi \approx \frac{2n}{\sum_{i=1}^n \sin X_i}, X_i \in [0, \pi] \text{均匀}$$

- 估算 π 需要事先知道 $\sin(\cdot)$ 函数的计算方法

$$\pi \approx \frac{2n}{\sum_{i=1}^n \sin X_i}, X_i \in [0, \pi] \text{均匀}$$

- 估算 π 需要事先知道 $\sin(\cdot)$ 函数的计算方法



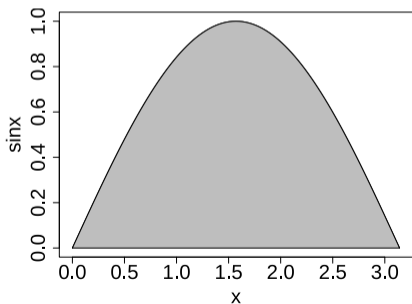
$$\pi \approx \frac{2n}{\sum_{i=1}^n \sin X_i}, X_i \in [0, \pi] \text{均匀}$$

- 估算 π 需要事先知道 $\sin(\cdot)$ 函数的计算方法



$$\pi \approx \frac{2n}{\sum_{i=1}^n \sin X_i}, X_i \in [0, \pi] \text{ 均匀}$$

- 计算 $\sin X_i$ 不易，如何简化问题？



- $\int_0^{\pi} \sin x dx$ 正是 $\sin x$ 与 x 轴围成的面积
- 每次均匀生成一对随机数 (X_i, Y_i)

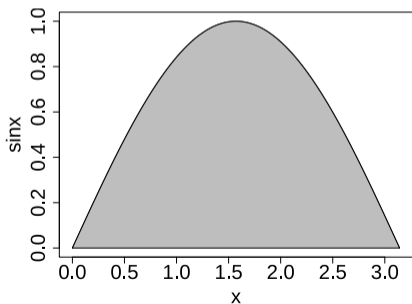
$$X_i \in [0, \pi], Y_i \in [0, 1]$$

- 定义函数 $h(\cdot, \cdot)$ ，如果它落在阴影中，则记 $h(X_i, Y_i) = 1$ ，否则记 $h(X_i, Y_i) = 0$ 。

$$\frac{1}{n} \sum_{i=1}^n h(X_i, Y_i) \approx \frac{2}{\pi}$$

$$\pi \approx \frac{2n}{\sum_{i=1}^n \sin X_i}, X_i \in [0, \pi] \text{ 均匀}$$

- 计算 $\sin X_i$ 不易，如何简化问题？

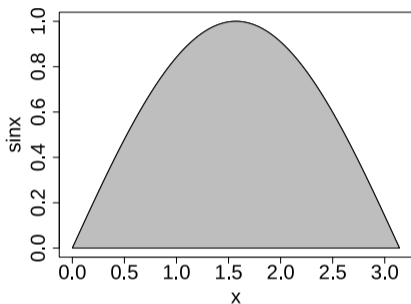


- $\int_0^{\pi} \sin x dx$ 正是 $\sin x$ 与 x 轴围成的面积
- 每次均匀生成一对随机数 (X_i, Y_i)

$$X_i \in [0, \pi], Y_i \in [0, 1]$$

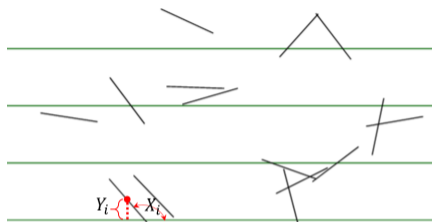
- 定义函数 $h(\cdot, \cdot)$ ，如果它落在阴影中，则记 $h(X_i, Y_i) = 1$ ，否则记 $h(X_i, Y_i) = 0$ 。

$$\frac{1}{n} \sum_{i=1}^n h(X_i, Y_i) \approx \frac{2}{\pi}$$

什么叫做“如果 (X_i, Y_i) 落在阴影中”？

$$h(X_i, Y_i) = \begin{cases} 1 & \sin X_i \leq Y_i \\ 0 & \sin X_i > Y_i \end{cases}$$

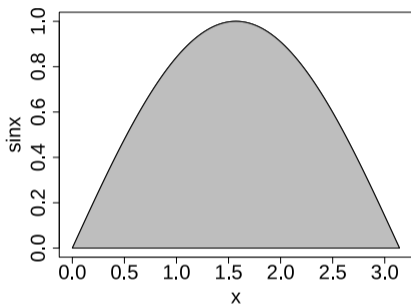
还需要算 \sin 函数吗？



- X_i 为针与平行线的夹角
- Y_i 为针的中点与最近平行线的距离
- 针长与平行线间隔相等

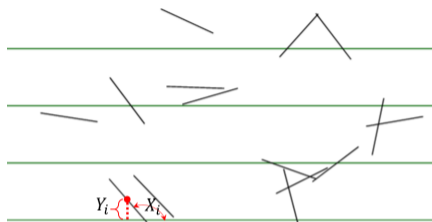
那么

$$\begin{cases} \sin X_i > Y_i & \Leftrightarrow \text{针与直线相交} \\ \sin X_i \leq Y_i & \Leftrightarrow \text{两者不相交} \end{cases}$$

什么叫做“如果 (X_i, Y_i) 落在阴影中”？

$$h(X_i, Y_i) = \begin{cases} 1 & \sin X_i \leq Y_i \\ 0 & \sin X_i > Y_i \end{cases}$$

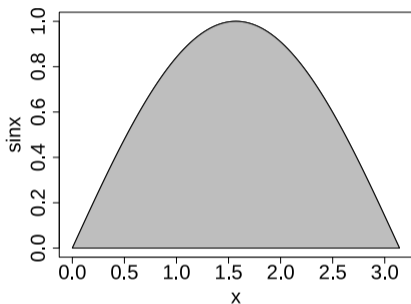
还需要算 \sin 函数吗？



- X_i 为针与平行线的夹角
- Y_i 为针的中点与最近平行线的距离
- 针长与平行线间隔相等

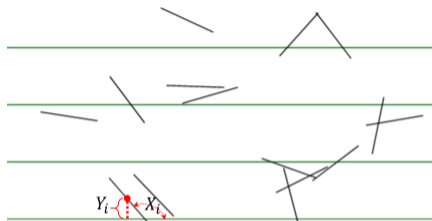
那么

$$\begin{cases} \sin X_i > Y_i & \Leftrightarrow \text{针与直线相交} \\ \sin X_i \leq Y_i & \Leftrightarrow \text{两者不相交} \end{cases}$$

什么叫做“如果 (X_i, Y_i) 落在阴影中”？

$$h(X_i, Y_i) = \begin{cases} 1 & \sin X_i \leq Y_i \\ 0 & \sin X_i > Y_i \end{cases}$$

还需要算 \sin 函数吗？



- X_i 为针与平行线的夹角
- Y_i 为针的中点与最近平行线的距离
- 针长与平行线间隔相等

那么

$$\begin{cases} \sin X_i > Y_i & \Leftrightarrow \text{针与直线相交} \\ \sin X_i \leq Y_i & \Leftrightarrow \text{两者不相交} \end{cases}$$

蒲丰（1707–1788 法国）投针方案 $\pi \approx \frac{2n}{\sum_{i=1}^n h_i}$

引言

复习与概念
引入

定义

舍选法

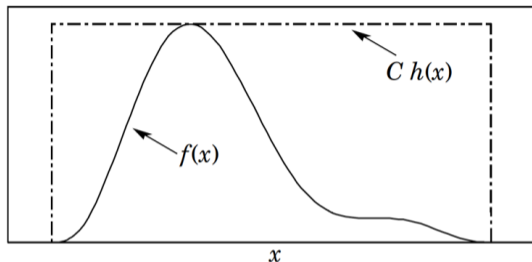
马尔可夫链法

逆函数法

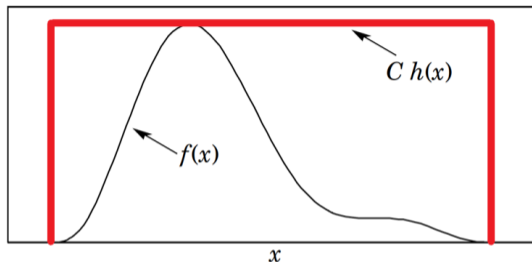
总结

| 试验者 | 时间 | 针长 | 投掷次数 | 相交次数 | π 的近似值 |
|-----------|------|--------|------|------|------------|
| Wolf | 1850 | 0.8 | 5000 | 2532 | 3.1596 |
| Smith | 1855 | 0.6 | 3204 | 1218 | 3.1554 |
| De Morgan | 1860 | 1.0 | 600 | 382 | 3.137 |
| Fox | 1884 | 0.75 | 1030 | 489 | 3.1595 |
| Lazzerini | 1901 | 0.83 | 3408 | 1808 | 3.1415929 |
| Reina | 1925 | 0.5419 | 2520 | 859 | 3.1795 |

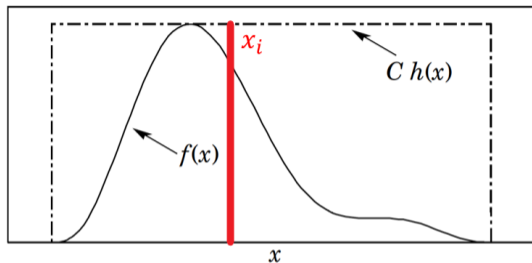
- 对比：祖冲之（429-500，南北朝），割圆术，3.1415927
 - 在圆周率的问题中，系统性取值比随机数效率高



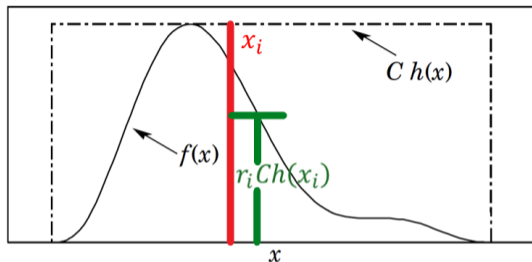
- ① 选取容易抽样的分布 $h(x)$ ，使得曲线 $Ch(x)$ 可以覆盖待抽样的分布 $f(x)$ ；
- ② 从 $h(x)$ 抽样得 x_i ；
- ③ 从 $[0, 1]$ 均匀分布抽样得 r_i ；
- ④ 如果 $r_i Ch(x_i) \leq f(x_i)$ 则保留 x_i 否则舍弃 (比大小)；
- ⑤ 保留的 x_i 序列服从 $f(x)$ 分布，可用序列的算术平均估算数学期望。



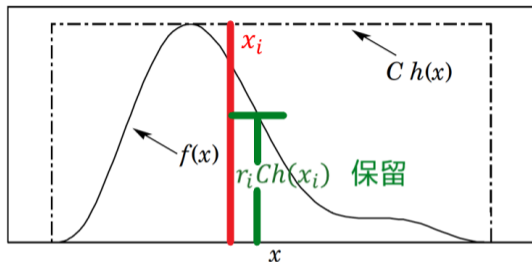
- ① 选取容易抽样的分布 $h(x)$ ，使得曲线 $Ch(x)$ 可以覆盖待抽样的分布 $f(x)$ ；
- ② 从 $h(x)$ 抽样得 x_i ；
- ③ 从 $[0, 1]$ 均匀分布抽样得 r_i ；
- ④ 如果 $r_i Ch(x_i) \leq f(x_i)$ 则保留 x_i 否则舍弃 (比大小)；
- ⑤ 保留的 x_i 序列服从 $f(x)$ 分布，可用序列的算术平均估算数学期望。



- ① 选取容易抽样的分布 $h(x)$ ，使得曲线 $Ch(x)$ 可以覆盖待抽样的分布 $f(x)$ ；
- ② 从 $h(x)$ 抽样得 x_i ；
- ③ 从 $[0, 1]$ 均匀分布抽样得 r_i ；
- ④ 如果 $r_i Ch(x_i) \leq f(x_i)$ 则保留 x_i 否则舍弃 (比大小)；
- ⑤ 保留的 x_i 序列服从 $f(x)$ 分布，可用序列的算术平均估算数学期望。



- ① 选取容易抽样的分布 $h(x)$ ，使得曲线 $Ch(x)$ 可以覆盖待抽样的分布 $f(x)$ ；
- ② 从 $h(x)$ 抽样得 x_i ；
- ③ 从 $[0, 1]$ 均匀分布抽样得 r_i ；
- ④ 如果 $r_i Ch(x_i) \leq f(x_i)$ 则保留 x_i 否则舍弃 (比大小)；
- ⑤ 保留的 x_i 序列服从 $f(x)$ 分布，可用序列的算术平均估算数学期望。

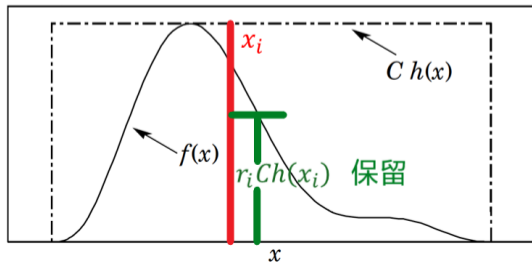


- ① 选取容易抽样的分布 $h(x)$ ，使得曲线 $Ch(x)$ 可以覆盖待抽样的分布 $f(x)$ ；
- ② 从 $h(x)$ 抽样得 x_i ；
- ③ 从 $[0, 1]$ 均匀分布抽样得 r_i ；
- ④ 如果 $r_i Ch(x_i) \leq f(x_i)$ 则保留 x_i 否则舍弃 (比大小)；
- ⑤ 保留的 x_i 序列服从 $f(x)$ 分布，可用序列的算术平均估算数学期望。

欲证明保留下来的 x_i 服从 $f(y)$ 分布, 即 $P(x_i \leq y) = \int_{-\infty}^y f(x)dx$ 。

舍选法的操作步骤恰是条件概率。证明如下:

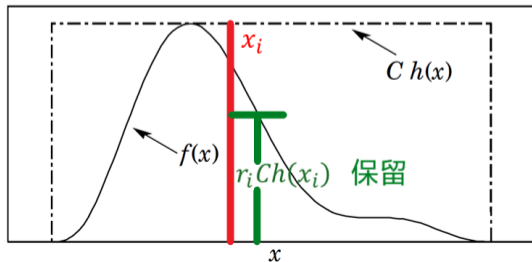
$$\begin{aligned}
 P(x_i \leq y) &= P[x_i \leq y | r_i C h(x_i) \leq f(x_i)] = \frac{P[x_i \leq y, r_i C h(x_i) \leq f(x_i)]}{P[r_i C h(x_i) \leq f(x_i)]} \\
 &= \frac{\int_{-\infty}^y \left[\int_0^{f(x)/Ch(x)} dr \right] h(x) dx}{\int_{-\infty}^{\infty} \left[\int_0^{f(x)/Ch(x)} dr \right] h(x) dx} = \frac{\frac{1}{C} \int_{-\infty}^y f(x) dx}{\frac{1}{C} \int_{-\infty}^{\infty} f(x) dx} = \int_{-\infty}^y f(x) dx
 \end{aligned}$$



欲证明保留下来的 x_i 服从 $f(y)$ 分布, 即 $P(x_i \leq y) = \int_{-\infty}^y f(x)dx$ 。

舍选法的操作步骤恰是条件概率。证明如下:

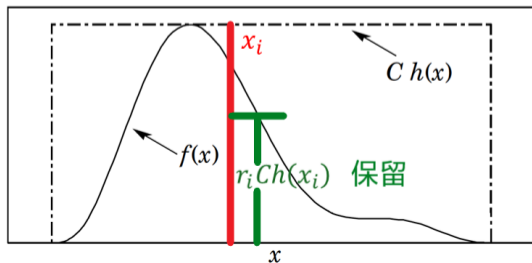
$$\begin{aligned}
 P(x_i \leq y) &= P[x_i \leq y | r_i C h(x_i) \leq f(x_i)] = \frac{P[x_i \leq y, r_i C h(x_i) \leq f(x_i)]}{P[r_i C h(x_i) \leq f(x_i)]} \\
 &= \frac{\int_{-\infty}^y \left[\int_0^{f(x)/Ch(x)} dr \right] h(x) dx}{\int_{-\infty}^{\infty} \left[\int_0^{f(x)/Ch(x)} dr \right] h(x) dx} = \frac{\frac{1}{C} \int_{-\infty}^y f(x) dx}{\frac{1}{C} \int_{-\infty}^{\infty} f(x) dx} = \int_{-\infty}^y f(x) dx
 \end{aligned}$$



欲证明保留下来的 x_i 服从 $f(y)$ 分布, 即 $P(x_i \leq y) = \int_{-\infty}^y f(x)dx$ 。

舍选法的操作步骤恰是条件概率。证明如下:

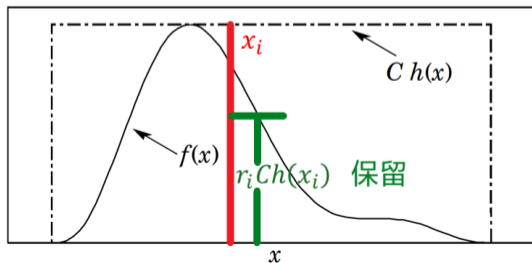
$$\begin{aligned}
 P(x_i \leq y) &= P[x_i \leq y | r_i C h(x_i) \leq f(x_i)] = \frac{P[x_i \leq y, r_i C h(x_i) \leq f(x_i)]}{P[r_i C h(x_i) \leq f(x_i)]} \\
 &= \frac{\int_{-\infty}^y \left[\int_0^{f(x)/Ch(x)} dr \right] h(x) dx}{\int_{-\infty}^{\infty} \left[\int_0^{f(x)/Ch(x)} dr \right] h(x) dx} = \frac{\frac{1}{C} \int_{-\infty}^y f(x) dx}{\frac{1}{C} \int_{-\infty}^{\infty} f(x) dx} = \int_{-\infty}^y f(x) dx
 \end{aligned}$$



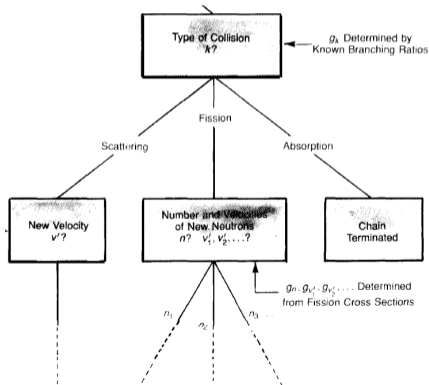
欲证明保留下来的 x_i 服从 $f(y)$ 分布, 即 $P(x_i \leq y) = \int_{-\infty}^y f(x)dx$ 。

舍选法的操作步骤恰是条件概率。证明如下:

$$\begin{aligned}
 P(x_i \leq y) &= P[x_i \leq y | r_i C h(x_i) \leq f(x_i)] = \frac{P[x_i \leq y, r_i C h(x_i) \leq f(x_i)]}{P[r_i C h(x_i) \leq f(x_i)]} \\
 &= \frac{\int_{-\infty}^y \left[\int_0^{f(x)/Ch(x)} dr \right] h(x) dx}{\int_{-\infty}^{\infty} \left[\int_0^{f(x)/Ch(x)} dr \right] h(x) dx} = \frac{\frac{1}{C} \int_{-\infty}^y f(x) dx}{\frac{1}{C} \int_{-\infty}^{\infty} f(x) dx} = \int_{-\infty}^y f(x) dx
 \end{aligned}$$



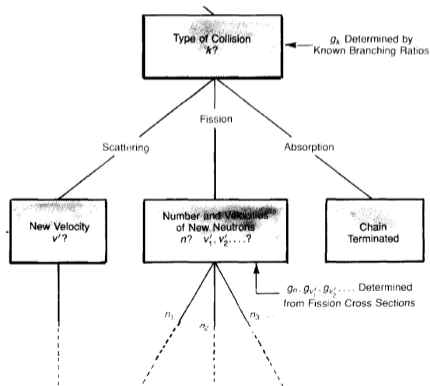
- 1940 年代由 Stan Ulam 和 von Neumann 计算中子扩散与核弹引爆机制，在最早的电子计算机上运行。
- “蒙特卡罗”是 Stan 叔叔常去的赌场区，被两位作者借用来代表“随机性”。



Stan Ulam, John von Neumann, and the Monte Carlo Method, 1987

- 我国科学团队在当时刚研制成功的“104 电子计算机”上进行了第十次蒙特卡罗计算，确认无误。
- 在计算机上复现了苏联专家的错误原因。
- 电子计算机的出现，让蒙特卡罗方法蓬勃发展。

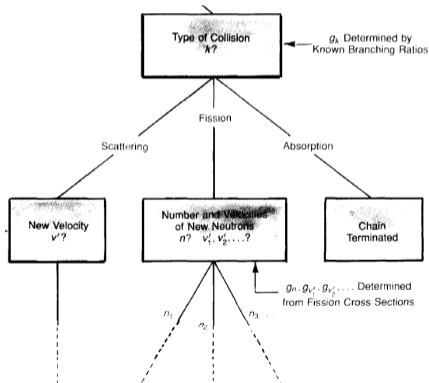
- 1940 年代由 Stan Ulam 和 von Neumann 计算中子扩散与核弹引爆机制，在最早的电子计算机上运行。
- “蒙特卡罗”是 Stan 叔叔常去的赌场区，被两位作者借用来代表“随机性”。



Stan Ulam, John von Neumann, and the Monte Carlo Method, 1987

- 我国科学团队在当时刚研制成功的“104 电子计算机”上进行了第十次蒙特卡罗计算，确认无误。
- 在计算机上复现了苏联专家的错误原因。
- 电子计算机的出现，让蒙特卡罗方法蓬勃发展。

- 1940 年代由 Stan Ulam 和 von Neumann 计算中子扩散与核弹引爆机制，在最早的电子计算机上运行。
- “蒙特卡罗”是 Stan 叔叔常去的赌场区，被两位作者借用来代表“随机性”。



Stan Ulam, John von Neumann, and the Monte Carlo Method, 1987

- 我国科学团队在当时刚研制成功的“104 电子计算机”上进行了第十次蒙特卡罗计算，确认无误。
- 在计算机上复现了苏联专家的错误原因。
- 电子计算机的出现，让蒙特卡罗方法蓬勃发展。

蒙特卡罗方法

续本达

引言

复习与概念
引入

定义

舍选法

马尔可夫链法

逆函数法

总结

马尔可夫链法

舍选法的关键步骤

如果 $r_i \text{Ch}(x_i) \leq f(x_i)$ 则保留 x_i , 否则舍弃。

- 困难：在高维问题中，无法知道 $f(x)$ 的绝对数值，只知道相对值 $\tilde{f}(x)$

$$f(x_i) = \frac{\tilde{f}(x_i)}{\int \tilde{f}(x) dx}$$

分母上的积分无法有效算出。

- 把 $f(x_i)$ 与上一个采样点 $f(x_{i-1})$ 比较，决定是否保留。
- 以概率 $\min \left[1, \frac{f(x_i)}{f(x_{i-1})} \right]$ 接受 x_i ，否则令 $x_i = x_{i-1}$

$$x_1 \xrightarrow{\text{接受}} x_2 \xrightarrow{\text{拒绝}} x_2 \xrightarrow{\text{拒绝}} x_2 \xrightarrow{\text{接受}} x_3 \rightarrow \dots$$

- 每个 x_i 都只依赖于 x_{i-1} ，不依赖于更早的历史，形成马尔可夫链。
- 规避了计算 $f(x)$ 归一化常数的困难。

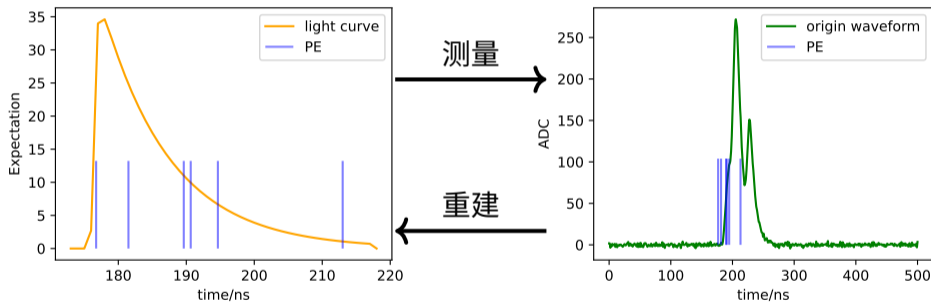
- 把 $f(x_i)$ 与上一个采样点 $f(x_{i-1})$ 比较，决定是否保留。
- 以概率 $\min \left[1, \frac{f(x_i)}{f(x_{i-1})} \right]$ 接受 x_i ，否则令 $x_i = x_{i-1}$

$$x_1 \xrightarrow{\text{接受}} x_2 \xrightarrow{\text{拒绝}} x_2 \xrightarrow{\text{拒绝}} x_2 \xrightarrow{\text{接受}} x_3 \rightarrow \dots$$

- 每个 x_i 都只依赖于 x_{i-1} ，不依赖于更早的历史，形成**马尔可夫链**。
- 规避了计算 $f(x)$ 归一化常数的困难。

设光敏仪器按一定的**光强曲线** $\mu\phi(t - t_0)$ 接收**光电子** (photoelectron, PE)，各光电子时间为 $\vec{z} := (t_1, t_2, \dots, t_N)$ 。 **波形** \vec{w} 是时间序列向量。

$$\mu, t_0 \rightarrow \vec{z} \rightarrow \vec{w}$$

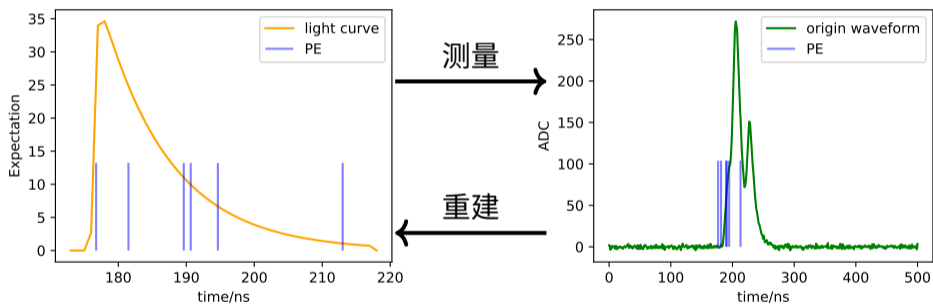


$$\mu, t_0 \xleftarrow{\text{拟合}} \vec{z} \xleftarrow{\text{分析计数}} \vec{w}$$

- 但脉冲可能有重叠，无法确切测量 \vec{z} ，成为粒子物理实验领域的难题。

设光敏仪器按一定的**光强曲线** $\mu\phi(t - t_0)$ 接收**光电子** (photoelectron, PE)，各光电子时间为 $\vec{z} := (t_1, t_2, \dots, t_N)$ 。 **波形** \vec{w} 是时间序列向量。

$$\mu, t_0 \rightarrow \vec{z} \rightarrow \vec{w}$$

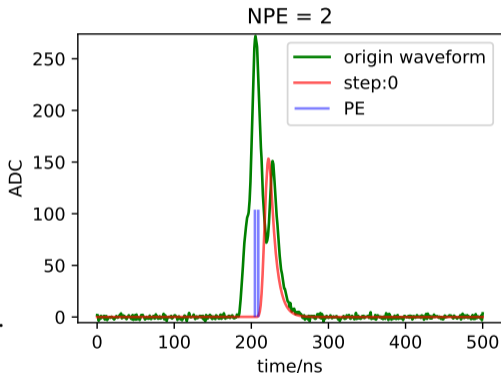


$$\mu, t_0 \xleftarrow{\text{拟合}} \vec{z} \xleftarrow{\text{分析计数}} \vec{w}$$

- 但脉冲可能有重叠，无法确切测量 \vec{z} ，成为粒子物理实验领域的难题。

$$\begin{aligned}
 p(\vec{w}|t_0, \mu) &= \underbrace{p(\vec{w}|\emptyset)p(\emptyset|t_0, \mu)}_{\text{空白解0重积分}} \\
 &+ \underbrace{\int_{\vec{z} \in T} p(\vec{w}|\vec{z})p(\vec{z}|t_0, \mu)d\vec{z}}_{\text{1重积分}} \\
 &+ \underbrace{\iint_{\vec{z} \in T^2} p(\vec{w}|\vec{z})p(\vec{z}|t_0, \mu)d\vec{z}}_{\text{2重积分}} \\
 &+ \underbrace{\iiint_{\vec{z} \in T^3} p(\vec{w}|\vec{z})p(\vec{z}|t_0, \mu)d\vec{z} \dots}_{\text{3重积分}}
 \end{aligned}$$

NPE 代表 \vec{z} 的元素个数，它不确定。



- 使用马尔可夫链，令 \vec{z} 遍历各重积分，问题得到完满解决。

$$\begin{aligned}
 p(\vec{w}|t_0, \mu) &= \underbrace{p(\vec{w}|\emptyset)p(\emptyset|t_0, \mu)}_{\text{空白解0重积分}} \\
 &+ \underbrace{\int_{\vec{z} \in T} p(\vec{w}|\vec{z})p(\vec{z}|t_0, \mu)d\vec{z}}_{\text{1重积分}} \\
 &+ \underbrace{\iint_{\vec{z} \in T^2} p(\vec{w}|\vec{z})p(\vec{z}|t_0, \mu)d\vec{z}}_{\text{2重积分}} \\
 &+ \underbrace{\iiint_{\vec{z} \in T^3} p(\vec{w}|\vec{z})p(\vec{z}|t_0, \mu)d\vec{z} \cdots}_{\text{3重积分}}
 \end{aligned}$$

NPE 代表 \vec{z} 的元素个数，它不确定。

- 使用马尔可夫链，令 \vec{z} 游历各重积分，问题得到完满解决。

蒙特卡罗方法

续本达

引言

复习与概念
引入

定义

舍选法

马尔可夫链法

逆函数法

总结

逆函数法

考虑均匀分布的概率密度函数

$$u(x) = \begin{cases} 1 & x \in [0, 1] \\ 0 & \text{其它} \end{cases}$$

改写

$$\frac{1}{n} \sum_{i=1}^n f(X_i) \xrightarrow{P} \int_0^1 f(x) dx \implies \frac{1}{n} \sum_{i=1}^n f(X_i) \xrightarrow{P} \int_0^1 f(x) u(x) dx$$

困难

- 被积函数 $f(x)$ 的值很小的地方，加和的效率低下
- 如果能高效取到使 $f(X_i)$ 值更大的 X_i ，效率就会提升。

复习：变量替换

设随机变量 Y 的分布函数是 $F_Y(y) = P(Y < y)$ ，那么 $X \equiv F_Y(Y)$ 服从 $[0, 1]$ 区间上的均匀分布。

设 Y 概率密度是 $f_Y(y) = F'_Y(y)$ ，把 x 替换为 $F_Y(y)$

$$\frac{1}{n} \sum_{i=1}^n f(X_i) \xrightarrow{P} \int_0^1 f(x) dx = \int_{-\infty}^{+\infty} f[F_Y(y)] f_Y(y) dy$$

$$(\text{转移 } f_Y) \implies \frac{1}{n} \sum_{i=1}^n \frac{f(X_i)}{f_Y(Y_i)} \xrightarrow{P} \int_{-\infty}^{+\infty} f[F_Y(y)] dy = \int_0^1 f(x) \frac{1}{f_Y(F_Y^{-1}(x))} dx$$

$$(Y_i = F_Y^{-1}(X_i)) \implies \frac{1}{n} \sum_{i=1}^n \frac{f(X_i)}{f_Y[F_Y^{-1}(X_i)]} \xrightarrow{P} \int_0^1 f(x) dx$$

蒙特卡罗方法

续本达

引言

复习与概念
引入

定义

舍选法

马尔可夫链法

逆函数法

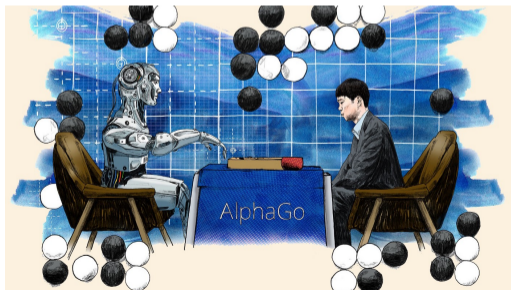
总结

总结

- ↓ 已知的信息越来越少，适用范围越来越广

| | 方法 | 已知的分布信息 | 维度 |
|--------|----------|--------------|----|
| ✓ | 逆函数法 | 完全已知 | 低 |
| ✓ | 舍选法 | 与标准分布的比值 | 低 |
| ✓ | 马尔可夫链法 | 与自身的比值 | 高 |
| 参考文献 5 | 对抗生成网络法* | 无解析形式，无法计算比值 | 高 |

- 蒙特卡罗方法在物理之外的几乎所有学科中都有应用（延伸阅读 1）
The results (马尔可夫链法) are really used in every aspect of scientific inquiry... I believe you can take any area of science, from hard to social, and find a burgeoning MCMC (马尔可夫链法) literature specifically tailored to that area.



- ① Diaconis, P., 2009. The Markov chain Monte Carlo revolution. Bull. Amer. Math. Soc. 46, 179 – 205.
<https://doi.org/10.1090/S0273-0979-08-01238-X>
- ② Dellaportas, P., Roberts, G.O., 2003. An Introduction to MCMC, in: Møller, J. (Ed.), Spatial Statistics and Computational Methods, Lecture Notes in Statistics. Springer, New York, NY, pp. 1 – 41.
https://doi.org/10.1007/978-0-387-21811-3_1
- ③ MacKay, D.J.C., MacKay, vid J.C., Kay, D.J.C.M., 2003. Information Theory, Inference and Learning Algorithms. Cambridge University Press. 第 29 章, Monte Carlo Methods
- ④ Galison, P.L., 1997. Image and Logic: A Material Culture of Microphysics, 1st edition. ed. University of Chicago Press, Chicago. 第 6 章, The Electronic Image: Iconoclasm and the New Icons

- ① Xu, D.C., Xu, B.D., Bao, E.J., Wu, Y.Y., Zhang, A.Q., Wang, Y.Y., Zhang, G.L., Xu, Y., Guo, Z.Y., Pei, J.H., Mao, H.Y., Liu, J.S., Wang, Z., Chen, S.M., 2022. Towards the ultimate PMT waveform analysis for neutrino and dark matter experiments. *J. Inst.* 17, P06040.
<https://doi.org/10.1088/1748-0221/17/06/P06040>
- ② Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E., 1953. Equation of State Calculations by Fast Computing Machines. *J. Chem. Phys.* 21, 1087 – 1092. <https://doi.org/10.1063/1.1699114>
- ③ Hastings, W.K., 1970. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* 57, 97 – 109.
<https://doi.org/10.1093/biomet/57.1.97>
- ④ Yao, A.C.-C., 1977. Probabilistic computations: Toward a unified measure of complexity, in: 18th Annual Symposium on Foundations of Computer Science (Sfcs 1977). Presented at the 18th Annual Symposium on Foundations of Computer Science (sfcs 1977), pp. 222 – 227.
<https://doi.org/10.1109/SFCS.1977.24>

- ① Paganini, M., de Oliveira, L., Nachman, B., 2018. Accelerating Science with Generative Adversarial Networks: An Application to 3D Particle Showers in Multilayer Calorimeters. *Phys. Rev. Lett.* 120, 042003.
<https://doi.org/10.1103/PhysRevLett.120.042003>