概率与期望

zhou888

雅礼中学

December 22, 2018

离散概率

取值范围为有限或无限可数个实数的随机变量称为离散型随机变量。设离散型随机变量X取值x_i时的概率为$p(k)(k = 1, 2, \dots)$,则称X的所有取值以及对应概率为X的概率分布,记做$P(X = x_k) = p(k)(k = 1, 2, \dots)$。

常见的离散型随机变量的概率分布有两点分布,二项分布,几何分布,超几何分布,泊松分布。

离散概率的期望

设离散型随机变量X的分布律为$PX = xk = pk(k = 1, 2, ...)$,若$\sum_{i=1}^{\infty} |p_i * x_i|$存在,则称$\sum_{i=1}^{\infty} |p_i * x_i|$为$X$的数学期望,简称期望,记为$E(X)$。

连续概率

如果X是在实数域或区间上取连续值的随机变量,设X的概率分布函数为$F(x)=P(X\le x)$,若存在非负可积函数$f(x)$,使得对任意的x,有

$$F(x)=\int_{-\infty}^{x} f(t)\mathrm{d}t$$

,则称X为连续型随机变量,称$f(x)$为X的概率密度函数。要注意,概率密度不是概率。

常见的连续型随机变量分布有均匀分布,正态分布,指数分布。

连续概率的期望

设连续型随机变量X的概率密度函数为$f(x)$,若广义积分$\int_{-\infty}^{\infty}|xf(x)|\mathrm{d}x$收敛,则称$\int_{-\infty}^{\infty}|xf(x)|\mathrm{d}x$为连续型随机变量$X$的数学期望,记为 $E(X)$。

条件概率

定义：设A,B是两个事件，且$P(A)>0$,则称$P(B|A)=\frac{P(AB)}{P(A)}$为在事件$A$发生的条件下事件$B$发生的条件概率。

非负性:对于每一事件B,有$P(B|A)\geq 0$

规范性:对于必然事件S，有$P(S|A)=1$

可列可加性:设$B_1,B_2,\ldots,B_n$是两两互不相容的事件，则

$$P(\bigcup_{i=1}^{n} B_i|A)=\sum_{i=1}^{n} P(B_i|A)$$

条件概率：理解

统计2018年某地的天气状况，得到某一天下雨的概率为$\frac{5}{8}$,连续两天都下雨的概率为$\frac{1}{4}$

现在从2018年中抽取连续两天，求：

(1)若第一天下雨，那么第二天下雨的概率.

(2)若连续两天中某一天下雨，那么另一天下雨的概率.

全概率公式与贝叶斯公式

乘法原理:由定义可得$P(AB) = P(A|B) \times P(B) = P(B|A) \times P(A)$

全概率公式:设S为实验E的样本空间,$B_1, B_2, \ldots, B_n$为E的一组划分,且$\forall P(B_i) > 0$,则$P(A) = \sum_{i=1}^{n} P(A|B_i) * P(B_i)$

贝叶斯公式:设S为实验E的样本空间,$B_1, B_2, \ldots, B_n$为S的一组划分,A为E的一个事件,且$\forall P(B_i) > 0$,则

$$P(B_k|A) = \frac{P(B_kA)}{P(A)} = \frac{P(A|B_k) * P(B_k)}{\sum_{i=1}^{n} P(A|B_i) * P(B_i)}$$

全概率公式与贝叶斯公式

特别的对于划分的$n = 2$时：

全概率公式:$P(A) = P(A|B) * P(B) + P(A|\overline{B}) * P(\overline{B})$

贝叶斯公式:

$$P(B|A) = \frac{P(BA)}{P(A)} = \frac{P(A|B) * P(B)}{P(A|B) * P(B) + P(A|\overline{B}) * P(\overline{B})}$$

全期望公式

$$E(X)=E(E(X|Y))$$

对于离散概率：

$$E(X)=E(E(X|Y))=\sum_{y_i}E(X|(Y=y_i))\times P(Y=y_i)$$

对于概率密度为$f(y)$的连续概率：

$$E(X)=E(E(X|Y))=\int_{-\infty}^{\infty}E(X|(Y=y))f(y)\mathrm{d}y$$

独立性

设A, B为两事件,如果满足$P(AB) = P(A)P(B)$则称事件A, B相互独立

若$P(A) > 0$且$P(B) > 0$,则A, B相互独立与A, B互不相容不能同时成立

若A, B相互独立,则:

$$P(A|B) = P(A), P(B|A) = P(B)$$

$$P(A_1 A_2 \dots A_n) = \prod_{i=1}^{n} P(A_i)$$

$$P(A_1 \bigcup A_2) = P(A_1) + P(A_2) - P(A_1 A_2)$$

$$E(A \times B) = E(A) \times E(B)$$

期望的线性性

无论两个变量X_1, X_2是否独立,总有:

$$E[\alpha X_1 + \beta X_2] = \alpha E[X_1] + \beta E[X_2]$$

方差

方差代表随机变量的取值对于其数学期望的离散程度

$$Var(X) = E((X - E(X))^2) = E(X^2) - (E(x))^2$$

概率生成函数

对于一个实数序列a满足$\{a_0, a_1, \ldots\}$，如果存在一个离散性随机变量使得$P(X = k) = a_k$，则称a的生成函数为X的概率生成函数

$F(x) = \sum_{i=0}^{\infty} P(X = i)x^i$

概率生成函数性质

$$F(1) = \sum_{i=0}^{\infty} P(X = i) = 1$$

$$E(x^k) = \sum_{i=0}^{\infty} P(X = i) \times i^{\underline{k}} = F^{(k)}(1)$$

$$Var(X) = E(X^2) - (E(x))^2 = F''(1) - (F'(1))^2$$

perface

有一些概率题与计算几何相关

很多期望题，实际上都是计数题

期望一般要倒推

一些路径上dp的概率期望高斯消元套路题就不讲了(参见HNOI2011 XOR路径和,HNOI2013游走)

DFN

给定一棵n个结点的树，你从点S出发，走到一个叶子结点T时结束。

每次dfs到一个点会以$random_shuffle$的顺序dfs遍历它相邻未到达的结点。

每新到一个点会使计数器加1，回溯时也会加1，求计数器值的期望

$1 \le n \le 10^6$

PKUWC 2018随机游走

给定一棵n个结点的树，你从点x出发，每次等概率随机选择一条与所在点相邻的边走过去。

有Q次询问，每次询问给定一个集合S，求如果从x出发一直随机游走，直到点集S中所有点都至少经过一次的话，期望游走几步。

$1 \leq n \leq 18, 1 \leq Q \leq 5000$

PKUWC 2018猎人杀

一开始有n个猎人，第i个猎人有仇恨度w_i，每个猎人只有一个固定的技能：死亡后必须开一枪，且被射中的人也会死亡。假设当前还活着的猎人有$[i_1 \dots i_m]$，那么有$\frac{w_{i_k}}{\sum\limits_{j=1}^{m} w_{i_j}}$的概率是向猎人$i_k$开枪。

一开始第一枪由你打响，目标的选择方法和猎人一样（即有$\frac{w_i}{\sum\limits_{j=1}^{n} w_j}$的概率射中第$i$个猎人）。由于开枪导致的连锁反应，所有猎人最终都会死亡，现在1号猎人想知道它是最后一个死的的概率。

答案对998244353取模。

$w_i > 0, \sum_{i=1}^{n} w_i \leq 10^5$

Pushing Balls

在一条直线上有n个球和$n+1$个坑共$2n+1$个“物品”,i个球在第i和第$i+1$个坑之间。

球与坑之间是有距离的,这些距离组成了首项为a,公差为x的等差数列。即第i个物品和第$i+1$个物品之间的距离是 $a+(i-1)x$。

肥克会进行n轮操作,每轮操作中,先从剩下的球中等概率地选择一个,然后等概率地选择一个方向,这个球将会朝这个方向滚,直到遇到一个里面没有球的坑并落进去留在里面。然后这一轮的收益为球滚的距离。

请求出期望收益。

$n \le 2*10^5, 0 \le x \le 100, 1 \le d \le 100$

Rainbow Balls

袋子里有n种球，一开始第i种颜色有a_i个。每次操作随机选两个球，将第一个球染成第二个球的颜色。求全部颜色变成相同的期望次数。

$1 \le n \le 2500, 1 \le a_i \le 10^5$

Wearry's duliu problem

袋子里有r个红球,g个绿球和b个蓝球。每回合你等概率随机取出袋子中的一个球。

如果是红球,就扔掉。

如果是绿球或者蓝球,就把球放回袋子里。

问到当进行到刚好拿出过k次蓝球的时候,期望经过了多少回合?

$1 \leq r, g, b, k \leq 10^9$

新年的五维几何

设$x_1, x_2, \ldots, x_n$是n个实数变量，其中第i个变量x_i在区间$[l_i, r_i]$内均匀随机生成，所有l_i和r_i均为给定的整数且$l_i \le r_i$（约定$l_i = r_i$时，$[l_i, r_i]$表示单元素集合$\{l_i\}$）。

给定$n \times n$的整数矩阵，矩阵的每个元素代表一个约束，其中第i行第j列的元素$a_{i,j}$代表约束$x_i - x_j \ge a_{i,j}$求这$n \times n$个约束同时被满足的概率。

$$1 \le n \le 5, 0 \le l_i \le r_i \le 10, -10 \le a_{i,j} \le 10$$

ClosestRabbit

有一个$n \times m$的矩阵,有r个doe要走进来。它们一个一个走进来,每个doe进来的时候都等概率随机选择一个矩阵中的空格子然后站在里面,当然有doe了就不是空格子了。等它们都选好格子以后,我们对第i个doe,定义$f(i)$为另外一个离它欧几里得距离最近的doe的编号,如果有距离一样的,那么优先选择所在行编号最小的,如果还有一样的,优先选择所在列编号最小的。然后我们把i和$f(i)$连在一起,变成一个图,求这个图中联通块期望的数量。

$n, m \le 20$

Yes or No

一共有$n+m$个询问，有n个询问的答案是Yes，其余m个是No。

你依次回答这些询问，每回答一个询问后会告诉你答案。

求最优策略下你期望答对的询问个数。

$n, m \leq 5 \times 10^5$

密码锁

给出一张n个点的完全图,现在要给这个完全图的每一条边随机定向成一个有向图。对于一条边$(i,j)(i<j)$,这条边的方向是i到j的概率是$\frac{num_{i,j}}{10000}$($num_{i,j}$ 指这条边旁边的数字,只有m条边对应的数字不是5000),否则就是j到i。在随机定向后,设这张有向图的强连通分量数目为x,求$x\times 10000^{n\times(n-1)}$的期望,可以证明该期望值一定是一个整数。

$1\le n\le 38, 0\le m\le 19, 0\le w_i\le 10000$

CTSC2017游戏

题面

ZJOI2015地震后的幻想乡

一个n个点m条边的无向简单图,每条边的边权是一个$[0,1]$之间的随机数,并且各个边边权都是独立的。求这个图的MST的边权最大值的期望

$n \leq 10$

MST

一个n个点m条边的无向简单图,每条边的边权是一个$[0,1]$之间的随机数,并且各个边边权都是独立的。求这个图的MST的边权和的期望

$n \leq 10$

歌唱王国

给定一个长度为L的序列A。然后每次掷一个标有1到m的公平骰子并将其上的数字加入到初始为空的序列B的末尾,如果序列B中已经出现了给定序列A,即A是B的子串,则停止,求序列B的期望长度。

$L \leq 10^5$。

概率型生成函数解题的一般方法

使用生成函数来解决这类问题的方法通常是先定义一个概率生成函数$F(x)$和一个辅助生成函数$G(x)$,然后是在未结束的情况后加入一个数或一个给定序列,并根据实际情况来列出方程。最后通过代值和求导来解出所需要的$F'(1)$

练习题

一个m面的公平骰子,求最后n次结果相同就结束的期望次数或者求最后n次结果全不同就结束的期望次数。保证$n, m \le 10^6$,且对第二问保证$n \le m$。

Thanks