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$p(k)(k=1,2,\ldots)$,则称X的所有取值以及对应概率为X的概率分布,记做$P(X=x_k)=p(k)(k=1,2,\ldots)$。

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$F(x)=P(X\leq x)$,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。

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。

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.

Solution

(1)

设事件A为“第一天下雨”，事件B为“第二天下雨”。

由题意可知$P(A) = \frac{5}{8}, P(AB) = \frac{1}{4}$

所以有：

$$P(B|A) = \frac{P(AB)}{P(A)} = \frac{\frac{1}{4}}{\frac{5}{8}} = \frac{2}{5}$$

Solution

(2)设事件C为“某一天下雨”。由题意可知$P(B)=\frac{5}{8}$

所以有：

$$
\begin{aligned}
P(C) &= P(A \cup B) \\
&= P(A) + P(B) - P(AB) \\
&= \frac{5}{8} + \frac{5}{8} - \frac{1}{4} = 1
\end{aligned}
$$

所以$P((AB)|C) = \frac{P(ABC)}{P(C)} = \frac{\frac{1}{4}}{1} = \frac{1}{4}$

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：

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

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:

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

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

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

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

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

$1 \le n \le 10^6$

Solution

我们考虑最后的期望其实等于经过每条边的期望之和。

Solution

我们考虑最后的期望其实等于经过每条边的期望之和。

我们将T到S的链提出来，并将这条链上的边称为关键边，在这条链上的点称为关键点。

Solution

我们考虑最后的期望其实等于经过每条边的期望之和。

我们将T到S的链提出来，并将这条链上的边称为关键边，在这条链上的点称为关键点。

对于每条关键边，它一定并且只会经过一次，故它的贡献就为1。

Solution

我们考虑最后的期望其实等于经过每条边的期望之和。

我们将T到S的链提出来，并将这条链上的边称为关键边，在这条链上的点称为关键点。

对于每条关键边，它一定并且只会经过一次，故它的贡献就为1。

对于每条非关键边，它只可能经过0次或2次。找到它最近的关键点祖先，那么，它为0还是为2取决于这个关键点是先遍历关键边还是先遍历它所对应的那个关键点下的儿子，而这两种情况的概率均为$\frac{1}{2}$，故这条边的贡献仍为1。

Solution

我们考虑最后的期望其实等于经过每条边的期望之和。

我们将T到S的链提出来，并将这条链上的边称为关键边，在这条链上的点称为关键点。

对于每条关键边，它一定并且只会经过一次，故它的贡献就为1。

对于每条非关键边，它只可能经过0次或2次。找到它最近的关键点祖先，那么，它为0还是为2取决于这个关键点是先遍历关键边还是先遍历它所对应的那个关键点下的儿子，而这两种情况的概率均为$\frac{1}{2}$，故这条边的贡献仍为1。

所以答案就是$n-1$。

Solution

我们考虑最后的期望其实等于经过每条边的期望之和。

我们将T到S的链提出来，并将这条链上的边称为关键边，在这条链上的点称为关键点。

对于每条关键边，它一定并且只会经过一次，故它的贡献就为1。

对于每条非关键边，它只可能经过0次或2次。找到它最近的关键点祖先，那么，它为0还是为2取决于这个关键点是先遍历关键边还是先遍历它所对应的那个关键点下的儿子，而这两种情况的概率均为$\frac{1}{2}$，故这条边的贡献仍为1。

所以答案就是$n-1$。惊不惊喜意不意外

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。

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

Solution

大家来看他的博客

Solution

大家来看他的博客

我们来请他讲一下吧

Solution

首先我们可以通过$min-max$反演，将经过这个集合每一次的时间转化为第一次经过这个集合里的点的时间。

Solution

首先我们可以通过$min-max$反演，将经过这个集合每一次的时间转化为第一次经过这个集合里的点的时间。

求第一次经过某个集合的点的时间的话可以设$f(i)$为从i出发经过某个集合内的点的期望次数。

Solution

首先我们可以通过$min-max$反演，将经过这个集合每一次的时间转化为第一次经过这个集合里的点的时间。

求第一次经过某个集合的点的时间的话可以设$f(i)$为从i出发经过某个集合内的点的期望次数。

则$f(i)=\frac{1}{d[i]}\times(\sum_{v\in ch[i]}f(v)+f(fa[i]))$，集合内的元素的$f$值为0

Solution

首先我们可以通过$min-max$反演，将经过这个集合每一次的时间转化为第一次经过这个集合里的点的时间。

求第一次经过某个集合的点的时间的话可以设$f(i)$为从i出发经过某个集合内的点的期望次数。

则$f(i)=\frac{1}{d[i]}\times(\sum_{v\in ch[i]}f(v)+f(fa[i])$，集合内的元素的$f$值为0

对于每个点我们可以把$f(x)$表示成$A\times f(fa[x])+B$的形式，这样一路递推至根结点。由于根结点没有父亲，f值可以直接求，再回带至每个结点即可。

Solution

首先我们可以通过$min - max$反演，将经过这个集合每一次的时间转化为第一次经过这个集合里的点的时间。

求第一次经过某个集合的点的时间的话可以设$f(i)$为从i出发经过某个集合内的点的期望次数。

则$f(i) = \frac{1}{d[i]} \times (\sum_{v \in ch[i]} f(v) + f(fa[i]))$，集合内的元素的$f$值为0

对于每个点我们可以把$f(x)$表示成$A \times f(fa[x]) + B$的形式，这样一路递推至根结点。由于根结点没有父亲，f值可以直接求，再回带至每个结点即可。

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$[i_1 \ldots i_m]$，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。

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

Solution

我们考虑对于已经死亡的猎人不将它去掉，而是将它打上一个标记，如果某一次随机到有标记的猎人就重新随机。

Solution

我们考虑对于已经死亡的猎人不将它去掉，而是将它打上一个标记，如果某一次随机到有标记的猎人就重新随机。

考虑容斥，枚举每一个集合S，计算至少有这个集合在1之后再死亡的概率。设这个集合的元素之和为W,所有元素的总和为sum,则答案为 $\sum_S (-1)^{|S|} \sum_{i=0}^{\infty} (1-\frac{w_1+W}{sum})^i \times \frac{w_1}{sum}$

Solution

我们考虑对于已经死亡的猎人不将它去掉，而是将它打上一个标记，如果某一次随机到有标记的猎人就重新随机。

考虑容斥，枚举每一个集合S，计算至少有这个集合在1之后再死亡的概率。设这个集合的元素之和为W,所有元素的总和为sum,则答案为 $\sum_S(-1)^{|S|}\sum_{i=0}^{\infty}(1-\frac{w_1+W}{sum})^i\times\frac{w_1}{sum}$

而$\sum_{i=0}^{\infty}(1-\frac{w_1+W}{sum})^i\times\frac{w_1}{sum}=\frac{sum}{w_1+W}\times\frac{w_1}{sum}=\frac{w_1}{w_1+W}$

Solution

这个式子直接做是没法做的，但我们发现$W \leq 10^5$，我们考虑每一种W的系数。

Solution

这个式子直接做是没法做的，但我们发现$W \le 10^5$，我们考虑每一种W的系数。

这个过程类似于一个背包，系数就是$\prod_{i=1}^{n}(1-x^{w_i})$中$x^W$项的系数

Solution

这个式子直接做是没法做的，但我们发现$W \le 10^5$，我们考虑每一种W的系数。

这个过程类似于一个背包，系数就是$\prod_{i=1}^{n}(1-x^{w_i})$中$x^W$项的系数

分治fft即可

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。

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

Solution

首先把“操作”换一种表述:等概率选择相邻的两个物品,并拿走他们,收益为两物品之间的距离,这样就不必区分球和坑了。

Solution

首先把“操作”换一种表述:等概率选择相邻的两个物品,并拿走他们,收益为两物品之间的距离,这样就不必区分球和坑了。

我们会发现,每次操作之后新局面的结构相同,能进行的操作也相同。我们可以直接把新局面取平均值来计算,而新局面的“期望局面”仍然是一个等差数列。

Solution

首先把“操作”换一种表述:等概率选择相邻的两个物品,并拿走他们,收益为两物品之间的距离,这样就不必区分球和坑了。

我们会发现,每次操作之后新局面的结构相同，能进行的操作也相同。我们可以直接把新局面取平均值来计算，而新局面的“期望局面”仍然是一个等差数列。

记m为现在剩下的区间数,一开始$m = 2 \times n$。计算得到

$$a' = \frac{(m+2)a + 5x}{m}$$

$$x' = \frac{m+4}{m}x$$

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。

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

Solution

我们考虑对于每种颜色分开考虑。设$f(i)$为最终要变成的颜色当前只有i个球，变到所有球相同的期望。s为所有$\sum_{i=1}^{n} a_i$

Solution

我们考虑对于每种颜色分开考虑。设$f(i)$为最终要变成的颜色当前只有i个球，变到所有球相同的期望。s为所有$\sum_{i=1}^{n} a_i$

则我们可以列出方程，令$p = \frac{i(s-i)}{s(s-1)}$，我们可以发现变多和变少的概率均为$p$,则我们可以列出转移

$$f(i) = p \times f(i+1) + p \times f(i-1) + (1-2p) \times f(i) + \frac{i}{s}$$

$f(n) = 0, f(0)$无定义

Solution

我们考虑对于每种颜色分开考虑。设$f(i)$为最终要变成的颜色当前只有i个球，变到所有球相同的期望。s为所有$\sum_{i=1}^{n} a_i$

则我们可以列出方程，令$p=\frac{i(s-i)}{s(s-1)}$，我们可以发现变多和变少的概率均为p,则我们可以列出转移

$$f(i)=p\times f(i+1)+p\times f(i-1)+(1-2p)\times f(i)+\frac{i}{s}$$

$f(n)=0,f(0)$无定义

为什么是$\frac{i}{s}$？因为后面加上的这个期望应该是向左和向右的概率均等，到达n之前没有到达过0的期望。

Solution

我们发现转移式可以变为$2f(i) = f(i-1) + f(i+1) + \frac{s-1}{s-i}$

对于$f(1)$:$2f(1) = f(2) + 1; f(i-1) - f(i) = f(i) - f(i+1) - \frac{s-1}{s-i}$

Solution

我们发现转移式可以变为$2f(i)=f(i-1)+f(i+1)+\frac{s-1}{s-i}$

对于$f(1)$:$2f(1)=f(2)+1$;$f(i-1)-f(i)=f(i)-f(i+1)-\frac{s-1}{s-i}$

$$
\begin{aligned}
f(1)=f(1)-f(n) &= \sum_{i=2}^{n} f(i-1)-f(i) \\
&= (s-1)(f(1)-f(2))+\sum_{i=2}^{n}\frac{s-1}{s-i}\times(s-i) \\
&= (s-1)(1-f(1))+(s-1)\times(s-2)
\end{aligned}
$$

所以$f(1)=\frac{(s-1)^2}{s}$，直接递推即可

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?

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

Solution

不妨分别计算结束之前,期望拿出多少个各种颜色的球。

Solution

不妨分别计算结束之前,期望拿出多少个各种颜色的球。

显然期望拿出正好k个蓝球。

Solution

不妨分别计算结束之前,期望拿出多少个各种颜色的球。

显然期望拿出正好k个蓝球。

考虑期望拿出多少个绿球,注意到蓝球和绿球可以当作相同的,所以期望拿出的绿球个数为蓝球个数的$\frac{g}{b}$倍。

Solution

不妨分别计算结束之前,期望拿出多少个各种颜色的球。

显然期望拿出正好k个蓝球。

考虑期望拿出多少个绿球,注意到蓝球和绿球可以当作相同的,所以期望拿出的绿球个数为蓝球个数的$\frac{g}{b}$倍。

考虑期望拿出多少个红球,不妨单独考虑每个红球,计算它被拿出来的概率,注意到这个时候我们可以无视其它所有的红色球和绿色球的存在,只考虑单个红色球与蓝色球。每两次拿蓝球之间的段结束当且仅当抽到一个蓝球或抽到这个红球，而这个被拿出来的概率是1−连续k次拿蓝球的概率$= 1 - (\frac{b}{b+1})^k$。

Solution

不妨分别计算结束之前,期望拿出多少个各种颜色的球。

显然期望拿出正好k个蓝球。

考虑期望拿出多少个绿球,注意到蓝球和绿球可以当作相同的,所以期望拿出的绿球个数为蓝球个数的$\frac{g}{b}$倍。

考虑期望拿出多少个红球,不妨单独考虑每个红球,计算它被拿出来的概率,注意到这个时候我们可以无视其它所有的红色球和绿色球的存在,只考虑单个红色球与蓝色球。每两次拿蓝球之间的段结束当且仅当抽到一个蓝球或抽到这个红球，而这个被拿出来的概率是1−连续k次拿蓝球的概率$= 1 - (\frac{b}{b+1})^k$。

简单求和就可以计算答案了。

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。

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

Solution

枚举每个x_i的整数部分.

Solution

枚举每个x_i的整数部分.

当$l_i = r_i$时，x_i是确定的，无需枚举；否则，记$x_i = p_i + q_i$，其中p_i在 $[l_i, r_i)$的整数中均匀随机,q_i在$[0, 1)$的实数中均匀随机.

Solution

枚举每个x_i的整数部分.

当$l_i = r_i$时，x_i是确定的，无需枚举；否则，记$x_i = p_i + q_i$，其中p_i在 $[l_i, r_i)$的整数中均匀随机,q_i在$[0, 1)$的实数中均匀随机.

我们枚举每个p_i，计算随机生成q_i有多少的概率满足所有约束，每组p_i下概率的平均值就是答案。

Solution

$x_i - x_j \geq a_{i,j}$等价于$q_i - q_j \geq a_{i,j} + p_j - p_i$

Solution

$x_i - x_j \geq a_{i,j}$等价于$q_i - q_j \geq a_{i,j} + p_j - p_i$

当$a_{i,j} + p_j - p_i \leq -1$时，约束总能满足，不考虑该约束

Solution

$x_i - x_j \geq a_{i,j}$等价于$q_i - q_j \geq a_{i,j} + p_j - p_i$

当$a_{i,j} + p_j - p_i \leq -1$时，约束总能满足，不考虑该约束

当$a_{i,j} + p_j - p_i \geq 1$时，约束不可能满足，该情况的概率直接返回0

Solution

$x_i - x_j \geq a_{i,j}$等价于$q_i - q_j \geq a_{i,j} + p_j - p_i$

当$a_{i,j} + p_j - p_i \leq -1$时，约束总能满足，不考虑该约束

当$a_{i,j} + p_j - p_i \geq 1$时，约束不可能满足，该情况的概率直接返回0

当$a_{i,j} + p_j - p_i = 0$ 时，约束等价于$q_i \geq q_j$。

Solution

现在只要考虑若干个形如$q_i \geq q_j$的约束同时被满足的条件。

Solution

现在只要考虑若干个形如$q_i \geq q_j$的约束同时被满足的条件。

注意到如果所有的q_i均匀随机生成，那么有两个q_i相等的概率为0,q_i可能出现的$n!$大小顺序是等概率的，每种的概率均为$\frac{1}{n!}$。因此对每个$l_i < r_i$的变量建一个点i，每个约束$qi \geq qj$ 连一条边$j \to i$，设该图的拓扑序有T个，则满足所有约束的概率为$\frac{T}{n!}$。

Solution

现在只要考虑若干个形如$q_i \geq q_j$的约束同时被满足的条件。

注意到如果所有的q_i均匀随机生成，那么有两个q_i相等的概率为0,q_i可能出现的$n!$大小顺序是等概率的，每种的概率均为$\frac{1}{n!}$。因此对每个$l_i < r_i$的变量建一个点i，每个约束$qi \geq qj$ 连一条边$j \to i$，设该图的拓扑序有T个，则满足所有约束的概率为$\frac{T}{n!}$。

直接状压DP即可计算拓扑序个数。

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。

$n, m \le 20$

solution

我们发现如果把$i \to f(i)$当作一条有向边，则这张图构成了一个基环内向树，每个联通块以环为根。则计算联通块数量的期望可以转化为查询环的数量的期望。

solution

我们发现如果把$i \to f(i)$当作一条有向边，则这张图构成了一个基环内向树，每个联通块以环为根。则计算联通块数量的期望可以转化为查询环的数量的期望。

我们发现对于一个长度为n的环,环上的点为$a_1, a_2, \ldots, a_n$,那么$\forall_i dis(a_i, a_{i+1}) \le dis(a_{i+1}, a_{i+2})$，所以环上的所有$dis(i, f(i))$应为相等的。

solution

我们发现如果把$i \to f(i)$当作一条有向边，则这张图构成了一个基环内向树，每个联通块以环为根。则计算联通块数量的期望可以转化为查询环的数量的期望。

我们发现对于一个长度为n的环,环上的点为$a_1, a_2, \ldots, a_n$,那么$\forall_i dis(a_i, a_{i+1}) \le dis(a_{i+1}, a_{i+2})$，所以环上的所有$dis(i, f(i))$应为相等的。

而我们发现a_1的字典序应小于a_3小于a_5，所以只要$n \ge 3$，就产生矛盾。

solution

我们发现如果把$i \to f(i)$当作一条有向边，则这张图构成了一个基环内向树，每个联通块以环为根。则计算联通块数量的期望可以转化为查询环的数量的期望。

我们发现对于一个长度为n的环,环上的点为$a_1, a_2, \ldots, a_n$,那么$\forall_i dis(a_i, a_{i+1}) \le dis(a_{i+1}, a_{i+2})$，所以环上的所有$dis(i, f(i))$应为相等的。

而我们发现a_1的字典序应小于a_3小于a_5，所以只要$n \ge 3$，就产生矛盾。

所以我们统计每一对(i_1, j_1),(i_2, j_2)的贡献，根据期望的线性性加起来即可。这个问题就是一个简单的组合计数问题了。

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。

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

solution

很容易发现每次的策略会选择当前剩下的n,m中较大的哪一个，只有在$n = m$的时候会以$\frac{1}{2}$的概率乱猜一个。

solution

很容易发现每次的策略会选择当前剩下的n,m中较大的哪一个，只有在$n = m$的时候会以$\frac{1}{2}$的概率乱猜一个。

考虑每一种做题方式就对应着一种从(n, m)到$(0, 0)$的路径，yes对应向左,no对应向下。

solution

很容易发现每次的策略会选择当前剩下的n,m中较大的哪一个，只有在$n = m$的时候会以$\frac{1}{2}$的概率乱猜一个。

考虑每一种做题方式就对应着一种从(n, m)到$(0, 0)$的路径，yes对应向左,no对应向下。

当你不经过$y = x$这条线时，$max(n, m)$是不会变的，所以你的策略永远是选择同一个方向走。

solution

很容易发现每次的策略会选择当前剩下的n,m中较大的哪一个，只有在$n = m$的时候会以$\frac{1}{2}$的概率乱猜一个。

考虑每一种做题方式就对应着一种从(n, m)到$(0, 0)$的路径，yes对应向左,no对应向下。

当你不经过$y = x$这条线时，$max(n, m)$是不会变的，所以你的策略永远是选择同一个方向走。

从上一次经过$y = x$的点(u, u)到这一次经过$y = x$的(v, v)的贡献永远是$u - v$。所以无论如何答案都有$max(n, m)$的贡献。

solution

很容易发现每次的策略会选择当前剩下的n,m中较大的哪一个，只有在$n=m$的时候会以$\frac{1}{2}$的概率乱猜一个。

考虑每一种做题方式就对应着一种从(n,m)到$(0,0)$的路径，yes对应向左,no对应向下。

当你不经过$y=x$这条线时，$max(n,m)$是不会变的，所以你的策略永远是选择同一个方向走。

从上一次经过$y=x$的点(u,u)到这一次经过$y=x$的(v,v)的贡献永远是$u-v$。所以无论如何答案都有$max(n,m)$的贡献。

只要计算每次经过$y=x$上的点的贡献,贡献为(n,m)到(x,x)的方案数，直接组合数。

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。

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

myy's solution

我们先统计出强连通分量数的期望,再乘上$10000^{n\times(n-1)}$即可。

myy's solution

我们先统计出强连通分量数的期望,再乘上$10000^{n\times(n-1)}$即可。

注意到定向后的图是个竞赛图，竞赛图的强连通分量缩点之后，一定是一条链，满足每个点和后面的点都有连边。

myy's solution

我们先统计出强连通分量数的期望,再乘上$10000^{n\times(n-1)}$即可。

注意到定向后的图是个竞赛图，竞赛图的强连通分量缩点之后，一定是一条链，满足每个点和后面的点都有连边。

答案即为链的长度。

myy's solution

我们先统计出强连通分量数的期望,再乘上$10000^{n\times(n-1)}$即可。

注意到定向后的图是个竞赛图，竞赛图的强连通分量缩点之后，一定是一条链，满足每个点和后面的点都有连边。

答案即为链的长度。

链的长度的统计可以转化为链的前缀的数量减一，这条链的前缀显然就是一个点集$S\in V$，满足所有的S与$V-S$之间的连边都是从S中的点连出去的，我们定义这样的点集为割集，那么我们统计割集数量即可。

myy's solution

对于一个点集，它是割集的概率就是这个点集中的点与点集外的所有点的所有连边都是从这个点集连出去的概率，也就是每条边是连出去的概率乘起来。

myy's solution

对于一个点集，它是割集的概率就是这个点集中的点与点集外的所有点的所有连边都是从这个点集连出去的概率，也就是每条边是连出去的概率乘起来。

对于概率非0.5的边我们称其为特殊边,对于大小为x的点集，如果其中没有任何特殊边则是割集的概率为$0.5^{x(n-x)}$，如果其中存在一条特殊边，设这条边从x连出的概率为P，那么这个概率要乘上$2P$，也就是这个特殊边的贡献。

myy's solution

我们不妨统计出每个大小的连通块，其中所有特殊边的贡献，最后对每个大小的连通块乘上$0.5^{x(n-x)}$再加起来即可。

myy's solution

我们不妨统计出每个大小的连通块，其中所有特殊边的贡献，最后对每个大小的连通块乘上$0.5^{x(n-x)}$再加起来即可。

如果一条边有贡献，那么它的两个端点必然有且仅有一个在S中，我们对于每个连通块直接枚举哪些点在S中。计算贡献，最后再做一个类似背包的暴力dp即可。

myy's solution

我们不妨统计出每个大小的连通块，其中所有特殊边的贡献，最后对每个大小的连通块乘上$0.5^{x(n-x)}$再加起来即可。

如果一条边有贡献，那么它的两个端点必然有且仅有一个在S中，我们对于每个连通块直接枚举哪些点在S中。计算贡献，最后再做一个类似背包的暴力dp即可。

复杂度$O(2^{m+1}n+n^2)$

CTSC2017游戏

题面

Solution

由贝叶斯公式：

$$
\begin{aligned}
P(X_m=1|X_l,X_r) &= \frac{P(X_m=1,X_l,X_r)}{P(X_l)\cdot P(X_r|X_l)} \\
&= \frac{P(X_m=1)\cdot P(X_l,X_r|X_m=1)}{P(X_l)\cdot P(X_r|X_l)} \\
&= \frac{P(X_m=1)\cdot P(X_l|X_m=1)\cdot P(X_r|X_m=1)}{P(X_l)\cdot P(X_r|X_l)} \\
&= \frac{P(X_m=1|X_l)\cdot P(X_r|X_m=1)}{P(X_r|X_l)}
\end{aligned}
$$

Solution

我们发现分母是常数，我们只需要计算$\sum_{i=l}^{r} P(X_i = 1|X_l) \cdot P(X_r|X_i = 1)$即可.

Solution

我们发现分母是常数，我们只需要计算$\sum_{i=l}^{r} P(X_i = 1|X_l) \cdot P(X_r|X_i = 1)$即可.

每次插入删除一个点时维护前驱后继，上面的式子只需要用线段树维护区间dp矩阵乘法。

Solution

我们发现分母是常数，我们只需要计算$\sum_{i=l}^{r} P(X_i = 1|X_l) \cdot P(X_r|X_i = 1)$即可.

每次插入删除一个点时维护前驱后继，上面的式子只需要用线段树维护区间dp矩阵乘法。

线段树上维护两个2×2的矩阵，一个表示概率，一个表示答案。每次合并左右区间时注意一下$pushup$

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

$n \leq 10$

Solution

设$P(x)$为答案大于x的概率，则$P(t) = \int_t^1 p(x)\mathrm{d}x$

则，答案要求的是:

Solution

设$P(x)$为答案大于x的概率，则$P(t)=\int_t^1 p(x)\mathrm{d}x$

则，答案要求的是:

$$\begin{aligned}
E(x) &= \int_0^1 xp(x)\mathrm{d}x \\
&= \int_0^1 \int_0^x p(x)\mathrm{d}x\mathrm{d}y \\
&= \int_0^1 \int_y^1 p(x)\mathrm{d}x\mathrm{d}y \\
&= \int_0^1 P(x)\mathrm{d}x
\end{aligned}$$

Solution

我们设$P_s(x)$为s这个集合联通所需要的边权最大值大于x的概率

Solution

我们设$P_s(x)$为s这个集合联通所需要的边权最大值大于x的概率

则我们对于一个集合s只需要求枚举将边权小于x的边全部连上之后一号点所在的联通块的集合u，并且保证$\overline{u}$与u之间所有的连边的权值都大于x即可。

$$P_s(x) = \sum_{\{1\} \subseteq u \subset s} (1 - P_u(x)) \times (1 - x)^{E(u,\overline{u})}$$

Solution

$$
\begin{aligned}
P_s &= \sum_{\{1\}\subseteq u\subset s}\int_0^1 (1-P_u(x))\times(1-x)^{E(u,\overline{u})}\mathrm{d}x \\
&= \sum_{\{1\}\subseteq u\subset s}\int_0^1 (1-x)^{E(u,\overline{u})}\mathrm{d}x + \sum_{\{1\}\subseteq u\subset s}\int_0^1 P_u(x)(1-x)^{E(u,\overline{u})}\mathrm{d}x \\
&= \sum_{\{1\}\subseteq u\subset s}\frac{1}{1+E(u,\overline{u})} + \sum_{\{1\}\subseteq u\subset s}\int_0^1 P_u(x)\times(1-x)^{E(u,\overline{u})}\mathrm{d}x
\end{aligned}
$$

Solution

设$dp(s,d)=\int_0^1 P_s(x)\times(1-x)^{d+E(u,\overline{u})}\mathrm{d}x$,则我们可以对这个式子直接$dp$

dp的初值为$dp(1,d)=0$

Solution

设$dp(s,d)=\int_0^1 P_s(x)\times(1-x)^{d+E(u,\overline{u})}\mathrm{d}x$,则我们可以对这个式子直接$dp$

dp的初值为$dp(1,d)=0$

复杂度$O(3^n\times m)$

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

$n \leq 10$

MST

由于存在两条边相同的概率为0，所以我们可以把每条边的权值都当作不同。那么对于任意一种情况，这张图的MST可以当作是唯一的。

MST

由于存在两条边相同的概率为0，所以我们可以把每条边的权值都当作不同。那么对于任意一种情况，这张图的MST可以当作是唯一的。

那么我们可以考虑对于每一条边计算期望，根据期望的线性性相加即可。

MST

由于存在两条边相同的概率为0，所以我们可以把每条边的权值都当作不同。那么对于任意一种情况，这张图的MST可以当作是唯一的。

那么我们可以考虑对于每一条边计算期望，根据期望的线性性相加即可。

若一条边连接(a,b)，则这条边的$E=\int_0^1 xp(x)\mathrm{d}x$，其中$p(x)$代表将小于$x$的所有边连接后$a$与$b$仍然不联通的概率。

MST

由于存在两条边相同的概率为0，所以我们可以把每条边的权值都当作不同。那么对于任意一种情况，这张图的MST可以当作是唯一的。

那么我们可以考虑对于每一条边计算期望，根据期望的线性性相加即可。

若一条边连接(a,b)，则这条边的$E=\int_0^1 xp(x)\mathrm{d}x$，其中$p(x)$代表将小于$x$的所有边连接后$a$与$b$仍然不联通的概率。

枚举每个联通块，每个块联通的概率仍然是一个多项式，采取和上题差不多的方法即可。

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。

$L \leq 10^5$。

solution

令a_i表示i是否为A的一个$border$,设$f(i)$为结束时长度为i的概率，$g(i)$为长度达到i还未结束的概率。设$f(i), g(i)$的概率生成函数分别为$F(x), G(x)$.则我们要求的即为$F'(1)$.我们有

solution

令a_i表示i是否为A的一个$border$,设$f(i)$为结束时长度为i的概率，$g(i)$为长度达到i还未结束的概率。设$f(i),g(i)$的概率生成函数分别为$F(x),G(x)$.则我们要求的即为$F'(1)$.我们有

$$F(x)+G(x)=G(x)\cdot x+1 \tag{1}$$

$$G(x)\cdot(\frac{1}{m}x)^L=\sum_{i=1}^{L}a_i\cdot F(x)\cdot(\frac{1}{m}x)^{(L-i)} \tag{2}$$

solution

(1)的意思为在一个未结束的串后加上一个串，可能结束也可能未结束

solution

(1)的意思为在一个未结束的串后加上一个串，可能结束也可能未结束

(2)的意思为在一个未结束的串后加上A,一定会结束，但是可能还没加完就结束。枚举它在放第几个时结束，可能结束的条件即为i为A的一个$border$

solution

(1)的意思为在一个未结束的串后加上一个串，可能结束也可能未结束

(2)的意思为在一个未结束的串后加上A,一定会结束，但是可能还没加完就结束。枚举它在放第几个时结束，可能结束的条件即为i为A的一个$border$

(1)两边求导得:$F'(x) + G'(x) = G'(x) \cdot x + G(x)$

solution

(1)的意思为在一个未结束的串后加上一个串，可能结束也可能未结束

(2)的意思为在一个未结束的串后加上A,一定会结束，但是可能还没加完就结束。枚举它在放第几个时结束，可能结束的条件即为i为A的一个$border$

(1)两边求导得:$F'(x) + G'(x) = G'(x) \cdot x + G(x)$

当$x = 1$时：$F'(1) + G'(1) = G'(1) + G(1)$,得到$F'(1) = G(1)$

solution

将$x=1$代入(2)得$G(1)=\sum_{i=1}^{L} a_i \cdot F(1) \cdot m^i$

solution

将$x=1$代入(2)得$G(1)=\sum_{i=1}^{L} a_i \cdot F(1) \cdot m^i$

由于$F(1)=1$，所以$F'(1)=G(1)=\sum_{i=1}^{L} a_i \cdot m^i$

solution

将$x=1$代入(2)得$G(1)=\sum_{i=1}^{L} a_i \cdot F(1) \cdot m^i$

由于$F(1)=1$，所以$F'(1)=G(1)=\sum_{i=1}^{L} a_i \cdot m^i$

故我们使用kmp或$hash$求$border$即可

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$n, m \leq 10^6$,且对第二问保证$n \leq m$。

solution

$F(x), G(x)$定义同上题

对于第一问

solution

$F(x), G(x)$定义同上题

对于第一问

$$F(x) + G(x) = G(x) \cdot x + 1$$

$$G(x) \cdot (\frac{1}{m}x)^{n-1} = \sum_{i=1}^{n} F(x) \cdot (\frac{1}{m}x)^{(n-i)}$$

solution

$F(x),G(x)$定义同上题

对于第一问

$$F(x)+G(x)=G(x)\cdot x+1$$
$$G(x)\cdot(\frac{1}{m}x)^{n-1}=\sum_{i=1}^{n}F(x)\cdot(\frac{1}{m}x)^{(n-i)}$$

解得：$F'(1)=\frac{m^n-1}{m-1}$

solution

对于第二问

solution

对于第二问

$$F(x) + G(x) = G(x) \cdot x + 1$$

$$G(x) \cdot (\frac{1}{m}x)^n * m^{\underline{n}} = \sum_{i=1}^{n} F(x) \cdot (\frac{1}{m}x)^{(n-i)} \cdot (m-i)^{\underline{n-i}}$$

solution

对于第二问

$$F(x) + G(x) = G(x) \cdot x + 1$$

$$G(x) \cdot (\frac{1}{m}x)^n * m^{\underline{n}} = \sum_{i=1}^{n} F(x) \cdot (\frac{1}{m}x)^{(n-i)} \cdot (m-i)^{\underline{n-i}}$$

解得：$F'(1) = \sum_{i=1}^{n} m^i \frac{(m-i)!}{m!}$

Thanks