本课件里的题不按难度排序。

# 概率 (Probability)

## 样本点，事件

随机试验中的每一个可能出现的试验结果称为这个试验的一个样本点，记作 $\omega_i$。全体样本点组成的集合称为这个试验的样本空间，记作 $\Omega$。即 $\Omega=\{\omega_1,...,\omega_n\}$。仅含一个样本点的随机事件称为基本事件，含有多个样本点的随机事件称为复合事件。

概率，亦称"或然率"，它是反映随机事件出现的可能性大小。我们约定 $P(A)$ 为事件 $A$ 的发生概率。

比如在上帝之手扔骰子这个随机试验中，"1 所在的面朝上" 就是这个试验的一个样本点。而 "1 所在的面朝上，2 所在的面朝北" 就是一个复合事件，而这个事件的概率是 $\frac{1}{24}$。

## 条件概率

我们称在事件 $B$ 已经发生的情况下，事件 $A$ 发生的概率为在 $B$ 的条件下 $A$ 的概率，记作 $P(A|B)$。根据定义，我们显然有条件概率的算数定义式：

$$P(A|B)=\frac{P(A\cap B)}{P(B)}$$

## 独立事件

当两个事件发生的概率互不影响时，我们称这两个事件互为独立事件。形式化的，我们称事件 $A$ 和事件 $B$ 互为独立事件，当且仅当 $P(A|B)=P(A)$。此时，我们获得了概率的乘法原理 $P(A\cap B)=P(A)P(B)$，其中，事件 $A$ 和事件 $B$ 独立。

# 例题

## LG 2911

此时你拥有三个面数分别为 $S_1,S_2,S_3$ 的骰子（$2\le S_i\le 40$）。你将使用上帝之手投掷它们。请你求出一个整数 $Q$，使得三个骰子向上的面的数字之和为 $Q$ 的概率最大。

### Sol

容易发现三个骰子是独立的。暴力枚举三个骰子向上的面，找到数字之和的众数，复杂度 $O(\prod S_i)$。

## 贝叶斯公式

根据条件概率的算数定义式：

$$P(A|B)=\frac{P(A\cap B)}{P(B)}\Rightarrow P(A|B)P(B)=P(A\cap B)$$
$$\therefore\frac{P(A|B)}{P(A)}=\frac{P(B|A)}{P(B)}$$

# 期望(Expect)

在概率论和统计学中，一个离散性随机变量的**期望值**是试验中每次可能的结果乘以其结果概率的总和。我们约定 $E(x)$ 为随机变量 $x$ 的期望值。

根据期望的定义，我们可以写出两个期望值的算数定义式。

$$E(x) = \sum_{i} P(x=i) \times i$$
$$E(x) = \sum_{i \geq 0} P(x>i) + \sum_{i \leq 0} -P(x<i)$$

比如在上帝之手扔骰子这个随机试验中，我们把朝上的面的数字看作随机变量 $x$，那么 $E(x)=\frac{7}{2}$

# 例题

## CF1194F

游戏有 $n$ 个关卡，第 $i$ 个关卡有 $\frac{1}{2}$ 的概率需要 $a_i+1$ 的能量通过，有 $\frac{1}{2}$ 的概率需要 $a_i$ 的能量通过。只能按顺序通关。你初始有 $T$ 的能量，求通过关数的期望。 $(1 \leq n \leq 2 \times 10^5, 1 \leq T \leq 2 \times 10^{14})$

### Sol

拿出期望的定义式 $E(x)=\sum\limits_{i \geq 0} P(x>i)=\sum\limits_{i \geq 1} P(x \geq i)$

设 $sum_i$ 表示 $\sum\limits_{j=1}^{i} a_j$，$D_i$ 表示前 $i$ 关加了多少个 1，那么

$P(x \geq i)=P(T-sum_i \geq D_i)=(\frac{1}{2})^i \sum\limits_{j=0}^{T-sum_i} \binom{i}{j}$

考虑组合数前缀和为 $S(n,m)=\sum\limits_{i=0}^{m} \binom{n}{i}$，虽然 $T$ 很大，但是 $m>n$ 时，我们可以让 $m=min(n,m)$

假设我们已经知道了 $S(x,y)$，那么我们可以 $O(1)$ 获得 $S(x+1,y), S(x,y+1), S(x-1,y), S(x,y-1)$，我们发现我们要计算 $S(i, min(T-sum_i, i))$ 时，第一个参数单调递增，第二个参数先增后减，且都在 $O(n)$ 范围内，如果预处理了阶乘和阶乘的逆元就可以 $O(n)$ 求得答案。

# 期望的线性性

对于任意两个随机变量 $x,y$，$E(ax+by)=aE(x)+bE(y)$。

证明可以考虑使用全概率公式暴力展开。

# 例题

## ARC114E

你现在有一个 $w \times h$ 的网格，其中 $(x_1,y_1)$，$(x_2,y_2)$ 的格子是黑的，其他都是白的。你每次会随机选择一条平行于坐标轴且不经过格子内部且不在网格边缘的一条直线，把网格沿直线割开。如果割出来两个部分都有黑的格子，则游戏结束。否则舍弃掉没有黑格子的那部分，对剩下部分继续游戏。求你割的刀数的期望。 $(1 \leq w,h \leq 10^5, (x_1,y_1) \neq (x_2,y_2))$

### Sol

我们可以把不分割黑格矩形的直线分成四个部分，分别是黑格矩形上方平行于 $x$ 轴的，黑格矩形下方平行于 $x$ 轴的，黑格矩形左方平行于 $y$ 轴的，黑格矩形右方平行于 $y$ 轴的。根据期望的线性性，容易发现答案就是每个部分期望被割次数之和 $+1$，那么做一个 $O(w+h)$ 的 `dp` 即可。

## LOJ 2263

$A$ 和 $B$ 在玩游戏，每一局不是 $A$ 赢就是 $B$ 赢，对于第一局 $A$ 有 $p_i$ 的概率赢。对于第 $i$ 局（$i \geq 2$），如果上一局是 $A$ 赢，那么 $A$ 有 $p_i$ 的概率赢，如果上一局是 $B$ 赢，那么 $A$ 有 $q_i$ 的概率赢。求 $A$ 赢的期望局数。

然后现在有 $q$ 次修改，每次形如：

- `add i c`，表示固定第 $i$ 局的输赢情况，如果 $c=1$ 则 $A$ 赢
- `del i`，表示第 $i$ 局的输赢情况不确定了。

每次修改后求答案。

（注意，根据贝叶斯公式，第 $i$ 局的输赢情况与第 $i-1$ 局有关时，事实上第 $i-1$ 局的输赢情况也和第 $i$ 局有关）

### Sol

假设第 $i$ 局的输赢情况为 $a_i$，$a_i=1$ 表示 $A$ 在第 $i$ 局赢了。

发现对于被已经固定的局分开的每一段而言，段与段之间是独立的，考虑分别求和。

对于段 $[l,r]$ 来说，第 $l$ 局和第 $r$ 局是固定的。

考虑期望的线性性，问题变成求解 $\sum P(a_m=1|a_l\cap a_r)$，根据贝叶斯公式：

$$\begin{aligned}P(a_m=1|a_l\cap a_r) &= \frac{P(a_m=1)P(a_l\cap a_r|a_m=1)}{P(a_l\cap a_r)}\\ &= \frac{P(a_l)P(a_r\cap a_m=1|a_l)}{P(a_l)P(a_r|a_l)}\\ &= \frac{P(a_m=1|a_l)P(a_r|a_l,a_m=1)}{P(a_r|a_l)}\\ &= \frac{P(a_m=1|a_l)P(a_r|a_m=1)}{P(a_r|a_l)}\end{aligned}$$

现在就从确定两边变成了确定了一端，已经和不修改的版本一样可以做到 $O(n)$ 一次查询了。

使用线段树维护转移矩阵就可以做到 $O(w^3(n+q)\log n)$，其中 $w$ 是转移矩阵的大小，在此题中为 2。

# 独立变量

对于两个变量 $x,y$，我们称 $x$ 和 $y$ 是相互独立的当且仅当 $\forall a,b,P(x=a\cap y=b)=P(x=a)P(y=b)$，更进一步的，$E(xy)=E(x)E(y)$。

# 例题

## 不知道哪来的题

你在 $(0,0)$，每次你向随机一个方向走一个单位长度，求你走完 $n$ 步后距离 $(0,0)$ 的欧几里得距离（勾股距离）的平方的期望，对 $998244353$ 取模，你可以证明答案是有理数。 $(1\leq n\leq 1000)$

### Sol

输入 $n$，输出 $n$。

这个过程可以看作随机 $n$ 个角度 $\{\theta_i\}$，那么距离原点的距离的平方就是 $(\sum cos(\theta_i))^2+(\sum sin(\theta_i))^2=n+\sum\limits_{i\neq j} cos(\theta_i)cos(\theta_j)+sin(\theta_i)sin(\theta_j)$，由于 $\theta_i$ 和 $\theta_j$ 是相互独立的，所以 $E(cos(\theta_i)cos(\theta_j)+sin(\theta_i)sin(\theta_j))=E(cos(\theta_i))E(cos(\theta_j))+E(sin(\theta_i))E(sin(\theta_j))=0$，所以

$E((\sum cos(\theta_i))^2+(\sum sin(\theta_i))^2)=n$。

## LG 1654

有一个序列，序列上的第 $i$ 的位置有 $p_i$ 的概率为 $1$，$1-p_i$ 的概率为 $0$，一个序列的权值是他所有极长的全 $1$ 区间的长度的立方和。求序列的权值的期望。（$1\le n\le 10^6$）

### Sol

设 $dp_{i,k}$ 表示以第 $i$ 个位置为结尾的极长全 $1$ 区间长度的 $k$ 次方的期望（$k\in\{1,2\}$）。那么由于两个位置是独立的：

$dp_{x,k}=(1+\sum\limits_{i=1}^{k}\binom{k}{i}dp_{x-1,i})\times p_i$，那么我们再记 $ans_i$ 为前 $i$ 个位置的答案，那么类似的可以得到 $ans_i=ans_{i-1}\times(1-p_i)+(ans_{i-1}+3\times dp_{i-1,2}+3\times dp_{i-1,1}+1)\times p_i$，$O(n)$ `dp` 亦可。

## 均值与方差

对于一个随机变量 $x$ 而言，它的均值即为它的期望值 $E(x)$。而它的方差也可以类比统计学里的定义写成 $V(x)=E((x-E(x))^2)=E(x^2)-2E(xE(x))+E^2(x)$，由于 $E(x)$ 是常数，所以我们可以得到 $V(x)=E(x^2)-E^2(x)$。

# 随机游走问题

## 一道大家都会的例题

给你一个无向联通简单图，你初始在 $u$ 处，你想到达 $v$ 处，每次你会均匀随机选择你所在的点连出的边中的一条并走到边的另一个端点，求第一次到达 $v$ 时你走的期望步数，对于每一对 $(u,v)$ 都输出答案。（$1\le n\le 50,n-1\le m\le \frac{(n-1)n}{2}$）

### Sol

我们枚举 $v$ 来求答案。

设 $dp_i$ 表示初始在 $i$ 这个点，第一次走到 $v$ 的期望步数。那么显然 $dp_v=0,dp_i=\frac{\sum\limits_{(i,j)\in E}dp_j+1}{deg_i}(i\ne v)$

然后虽然 $dp$ 有后效性，但我们可以把每个转移式子看作方程，高斯消元得到答案。

复杂度 $O(n^4)$。

## 又一道大家都会的例题

给你一棵无根树，你初始在 $u$ 处，你想到达 $v$ 处，每次你会均匀随机选择你所在的点连出的边中的一条并走到边的另一个端点，求第一次到达 $v$ 时你走的期望步数，对于每一对 $(u,v)$ 都输出答案。（$1\le n\le 5000$）

### Sol

我们枚举 $v$ 来求答案。

和上面一样的 `dp`，但是把父亲的贡献单独拎出来，$dp_u=\frac{dp_{fa}+\sum\limits_{v\in son_u}dpv}{deg_u}+1$

我们尝试把式子写成 $dp_u=a_udp_{fa}+b_u$ 的形式，然后根据 $dp_v=0$ 的性质可以推出所有答案。对于叶子节点来说，初始有 $dp_u=dp_{fa}+1$，其他的节点可以从下往上推出 $a_u,b_u$，复杂度 $O(n^2)$

### 双一道大家都会的例题

给你一棵无根树，你初始在 $u$ 处，你想到达 $v$ 处，每次你会均匀随机选择你所在的点连出的边中的一条并走到边的另一个端点，求第一次到达 $v$ 时你走的步数的平方的期望，对于询问 $q$ 次 $(u,v)$ 都输出答案。 $(1 \le n, q \le 100000)$

## Sol

首先考虑步数的期望

设$up_u$为节点$u$向上走到的 $fa$的期望的步数，$D_u$为$fa$走到$u$的期望的步数

所以 $up_u = \frac{1}{deg_u} + \frac{\sum_{v \in son_u}(1+up_v+up_u)}{deg_u}$，$D_u$ 有一个类似的式子。

预处理出每个节点$u$向上走到的 $fa$的期望的步数，以及$fa$走到$u$的期望的步数，因为期望的线性性，一个$u$到$v$的路径可以由这些累加起来，直接倍增就行了

然后我们发现步数平方的期望没有线性性!

对于两个独立变量$x$,$y$，我们考虑怎么求$E((x+y)^2)$

容易发现$E((x+y)^2) = E(x^2+2xy+y^2) = E(x^2)+2E(x)E(y)+E(y^2)$

所以同时维护

1.向上走到的 $fa$ 步数的期望，以及 $fa$ 走到 $u$ 步数的期望

2.向上走到的 $fa$ 步数平方的期望，以及 $fa$ 走到 $u$ 步数平方的期望

平方的期望可以参考上面的式子使用期望的线性性和独立变量的期望的可乘性来维护。

然后倍增就做完了

### 叒一道大家都会的例题 [LOJ 2542]

给你一棵无根树，你初始在 $u$ 处，你想到达 $S$ 中所有点至少一次，每次你会均匀随机选择你所在的点连出的边中的一条并走到边的另一个端点，求第一次经过 $S$ 内所有至少一次的时刻的期望值，对于询问 $q$ 次 $S$ 都输出答案。 $(1 \le n \le 18, 1 \le q \le 5000)$

## Sol

假设有一个集合 $S$，我们认为 $max\{S\}$ 为 $S$ 中每一个点第一次被到达的时刻的最大值，类似的有 $min\{S\}$。容易发现 $E(max\{S\})$ 就是答案。根据 $min-max$ 容斥：

$$max\{S\} = \sum_{T \subseteq S}(-1)^{|T|+1}min\{T\}$$
$$\therefore E(max\{S\}) = \sum_{T \subseteq S}(-1)^{|T|+1}E(min\{T\})$$

现在我们的问题变成了求 $E(min\{S\})$，我们可以枚举 $S$，使用与之前相类似的式子求解，此时复杂度为 $O(n2^n)$。

对于询问，由于数据足够水，可以通过 $O(2^n)$ 暴力枚举 $T$ 来求解，或者可以通过高维前缀和 $O(n2^n)$ 预处理，对于每组询问 $O(1)$ 输出答案。

# 一些杂题

## 一道大家可能都做过的题 CF1295F

给定 $n$ 个区间 $[l_i,r_i]$，$a_i$ 是在 $[l_i,r_i]$ 中随机出的一个整数，求 $a_i$ 不降的概率，输出对 $998244353$ 取模后的结果。（$1\le n\le 50, 0\le l_i\le r_i\le 998244351$）

### Sol

值域很大，考虑离散化成 $O(n)$ 个区间 $[L_i,R_i)$。

众所周知概率期望题就是组合计数题，考虑计算合法的方案数，再除以总方案 $\prod(r_i-l_i+1)$

设 $dp_{u,j}$ 表示第 $u$ 个数在第 $j$ 个区间的方案数，有式子 $dp_{u,j}=\sum_{v<u,l<j}dp_{v,l}\times\binom{R_j-L_j}{u-v}$，然后前缀和优化可以做到 $O(n^3)$

## 又一道大家可能都做过的题 ARC113F

给定 $n+1$ 个整数 $0=X_0<X_1<\cdots<X_n\le 10^6$，有 $n$ 个人，第 $i$ 个人在 $[X_{i-1},X_i]$ 中随机一个实数并站在数轴上相对应的位置。求人两两距离的最小值的期望，对 $998244353$ 取模。（$2\le n\le 20$）

### Sol

根据期望的算数定义式，我们考虑对于一个值 $x$，计算最小距离 $>x$ 的概率。

考虑第 $i$ 个人的位置是 $a_i$，那么 $a_i-a_{i-1}>x\Rightarrow(a_i-xi)-(a_{i-1}-x(i-1))>0$

所以问题变成了每个人在 $[X_{i-1}-xi,X_i-xi]$ 中随机，最后序列单增的概率。

考虑和上一题类似的式子，如果区间端点的大小关系没有改变，那么我们可以把 $dp_{u,j}$ 看作一个关于 $x$ 的不超过 $n$ 次的多项式。我们并不需要维护多项式，只需要维护多项式在 $0,\dots,n$ 处的点值，最后把多项式暴力 $O(n^2)$ 或 $O(n^3)$ 插值出来对 $x$ 积分即可。此时复杂度是 $O(n^4)$。

考虑到区间端点大小关系只会改变 $O(n^2)$ 次，所以对每次都做一遍 `dp` 即可。复杂度 $O(n^6)$。

## 双一道大家可能都做过的题 LOJ 2541

有 $n$ 个人，每个人仇恨值为 $w_i$，每一轮随机杀死一个还活着的人，一个人被杀死的概率与仇恨值成正比。求第 1 个人最后一个被杀死的概率。（$1\le w_i,\sum w_i\le 10^5$）

### Sol

一个人在一轮死掉的概率为 $\frac{w_i}{\sum\limits_{j\ is\ alive}w_j}$，如果一个人死后就消失了，下面的分母会改变，很烦。所以我们考虑每一轮都是给某个人打上一个死亡标记，即使已经被打过标记的人也可以被打标记，求若干轮后除了 1 之外都被打了标记的概率，根据条件概率的式子容易发现这个和原题答案相等。

考虑容斥，假设**至少** $S$ 中的人比 1 死得晚的概率为 $P(S)$，那么答案就是 $\sum(-1)^S P(S)$。

那么就是若干轮只给 $S\cup\{1\}$ 之外的人打标记，然后某一轮直接给 1 打了标记，写成式子就是：

$$
\begin{aligned}
P(S)&=\sum_{i\ge 0}\left(\frac{sum(U)-sum(S\cup\{1\})}{sum(U)}\right)^i\left(\frac{w_1}{sum(U)}\right)\\
&=\frac{w_1}{sum(S\cup\{1\})}
\end{aligned}
$$

考虑到，$sum(U)$ 很小，那么假设 $cnt_i=\sum\limits_{sum(S)=i}(-1)^{|S|}$，那么答案变成了 $\sum cnt_i\frac{w_1}{w_1+i}$

此时只需要考虑求 $cnt_i$，我们考虑 $cnt_i$ 的生成函数 $F=\sum cnt_i x^i$，发现 $F$ 就是 $\prod\limits_{i\ge 2}(1-x^{w_i})$，可以分治加 `NTT` 求解。

复杂度 $O(\sum w_i \log n \log \sum w_i)$

## 一道差不多的 gym102978H

有 $n$ 个特殊物品和 $m$ 个普通物品，每个物品有一个权值 $w_i$，特殊物品的权值和不超过 $10000$。每一轮随机删掉一个还存在的物品，一个物品被删除的概率与 $w_i$ 成正比。求删掉第一次删掉所有特殊物品的概率。 $(1 \le n, m \le 100)$

## Sol

可以观察到，答案就是期望有多少个普通物品在结束前被删掉加上 $n$，根据期望的线性性，我们只需要把每一个普通物品在结束前被删掉的概率相加亦可。

那么就是有 $n+1$ 个物品，求最后一个物品不是最后一个被删除的概率，那么就是上一题了。

这里特殊物品的权值不大，可以暴力做卷积，复杂度 $O(10000^2 + 10000m)$。

## 叒一道大家可能都做过的题 CF1349D

有 $n$ 个人，第 $i$ 个有 $a_i$ 个饼干，每一轮随机选择一个有饼干的人 $A$，一个人被选中的概率与持有饼干的数量成正比。然后再等概率随机另一个人 $B$，$A$ 将一块饼干给 $B$。当所有饼干都在同一个人手上游戏结束，求游戏轮数的期望，对 $998244353$ 取模。 $(2 \le n \le 10^5, 1 \le \sum a_i \le 3 \times 10^5)$

## Sol

设 $P_i$ 为第 $i$ 个人胜出的概率，$E_i$ 表示第 $i$ 个人胜出时，局数的条件期望，那么答案显然为 $\sum E_i$。

我们修改游戏规则，使得必须让饼干都到达 $i$ 手上时游戏才结束，把此时的期望轮数看作 $F_i$。

我们假设饼干全都在某人手上，然后全部转移到一个被钦定的人手上的期望轮数为 $C$。

那么显然有：

$$
\begin{gathered}
E_x = F_x - \sum_{i \ne x} E_i + P_i C \\
\sum E_i = F_x - C \sum_{i \ne x} P_i \\
n \sum E_i = \sum F_i - (n-1)C
\end{gathered}
$$

考虑 $g_i$ 为某个人手上有 $i$ 个饼干，下一次他手上拥有 $i+1$ 个饼干的期望轮数。

根据期望的线性性，那么 $F_i = \sum\limits_{j=a_i}^{sum-1} g_j, C = \sum\limits_{j=0}^{sum-1} g_j$。

考虑求 $g_i$，那么有：

$$
g_i = \frac{i}{sum}(g_{i-1} + g_i + 1) + \frac{sum - i}{sum}(\frac{1}{n-1} + \frac{n-2}{n-1}(g_i + 1))
$$

考虑 $g_0 = n - 1$，那么 $g_i$ 不难 $O(\sum a_i)$ 递推出。

## 来一道大家可能猜不出结论的题 CF1438F

#### 这是一道交互题

有一个高度为 $h$ 的满二叉树，但编号是未知的。你每次可以询问一个三元组 $(u, v, w)$，其中 $u, v, w$ 两两不同，交互器会返回在以编号为 $w$ 的节点为根时，编号为 $u$ 的节点和编号为 $v$ 的节点的 `lca` 的编号。你可以询问 $n + 420$ 次，求二叉树根节点的编号。 $(3 \le h \le 18)$

## Sol

我们按照如下的方式询问：

- 随机 $420$ 个三元组 $(u,v,w)$ 询问
- 找出交互器返回的数中出现次数最多的两个，设为 $c_1,c_2$
- 我们可以认为 $c_1,c_2$ 有极大概率就是根的两个儿子，所以对每一个 $i$，询问 $(c_1,c_2,i)$，如果返回的数字就是 $i$，那么 $i$ 是根。

可以发现，当我随机一个三元组时，某个点被返回的概率就是 $\frac{6(2^{h-dep_u-1}-1)^2(2^h-2^{h-dep_u})+2((2^{h-dep_u-1}-1)^2+2(2^{h-dep_u-1}-1)(2^h-2^{h-dep_u}))}{\binom{2^h-1}{3}}$，观察式子可以发现根的两个儿子是被返回概率最大的点，容易发现即使是 $h=18$ 时，根的儿子被返回的概率大约为 $18\%$，在更小的 $h$ 下，这个概率会更大。$420$ 次询问已经足以保证两个儿子在被返回的序列中成为众数。

### pbds 给的神仙题 AGC032F

你在圆周上随机选择 $n$ 个点，求任意两个不同的点之间弧度和 $\frac{2\pi}{3}$ 的差值的最小值的期望除以 $2\pi$ 的值，对 $10^9+7$ 取模，你可以证明答案是有理数。 $(2\le n\le 10^6)$

## Sol

我们考虑，发现我们可以钦定第一个点位置，答案不变，我们从这个位置开始沿顺时针每转 $\frac{2\pi}{3}$ 划分一个区域，把这三个区域里的点分别染上红蓝绿三色。然后把第二个区域的点逆时针转 $\frac{2\pi}{3}$，第三个区域顺时针转 $\frac{2\pi}{3}$，发现两个不同区域里的点的弧度与 $\frac{2\pi}{3}$ 的差值变成了两个不同颜色点之间的弧度。然后可以发现在同一个区域内的点连出来的劣弧一定不优（考虑在区域边界上的点可以被划分到任一个区间）。

现在考虑在 $[0,\frac{2\pi}{3})$ 内随机 $n-1$ 个点（固定第 $1$ 个点在 $0$），然后每个颜色随机染色，求两个不同颜色点之间的弧度的最小值的期望。

我们可以知道 $[0,1)$ 中随机 $n-1$ 个点，把线段分成 $n$ 段，线段的最小值的期望为 $F_1=\frac{1}{n^2}$，那么次小值 $F_2=\frac{1-nF_1}{(n-1)^2}+F_1=\frac{1}{n}(\frac{1}{n}+\frac{1}{n-1})$，三小值 $F_3=\frac{1-F_1-(n-1)F_2}{(n-2)^2}+F_2=\frac{1}{n}(\frac{1}{n}+\frac{1}{n-1}+\frac{1}{n-2})$，归纳得 $F_k=\frac{1}{n}\sum_{i=n-k+1}^{n}\frac{1}{i}$。

那么我们假设答案是划分出来的第 $k$ 小线段，那么发现必然有前 $k-1$ 小的线段两端同色。那么答案为：

$$
\begin{aligned}
&\sum_{i=1}^{n}(\frac{1}{3^{i-1}}-\frac{1}{3^i})\frac{F_i}{3}\\
=&\frac{1}{n}\sum_{i=1}^{n}\frac{1}{3^i(n-i+1)}
\end{aligned}
$$

$O(n)$ 计算亦可。

# 概率生成函数(PGF)

如果 $y$ 是一个仅取非负整数值的随机变量，我们可以写出它的概率生成函数：

$$F(x)=\sum P(y=i)x^i=E(x^y)$$

容易发现 $F(1)=1,F'(1)=E(y),V(x)=F''(1)+F'(1)(1-F'(1))$

对于随机变量 $x,y$ 和它们的 `PGF` $F_x,F_y$，可以知道 $F_{x+y}=F_x\times F_y$。

# 例题

## LG 4548

给定字符集大小 $n$ 和字符串 $S$。

你初始有一个空串 $T$，你每次从字符集里等概率随机出一个字符加到 $T$ 的末尾，直到 $S$ 成为 $T$ 的子串。

求 $T$ 的期望长度。 $(1 \le n \le 10^5, 1 \le |S| \le 10^5$，多组询问$)$

### Sol

设结束时 $|T|$ 的 `PGF` 为 $F(x)$，那么 $E(|T|) = F'(1)$。

设 $g_i$ 表示 $|T| = i$ 时 $S$ 还不是 $T$ 的子串的概率，设 $g$ 的生成函数为 $G(x)$

显然有式子 $xG(x) + 1 = F(x) + G(x)$，表示加 $1$ 个字符后要么结束了，要么没结束。

那么对两边求导，得 $xG'(x) + G(x) = F'(x) + G'(x)$，代入 $x = 1$，得 $F'(1) = G(1)$。

考虑到往 $T$ 后面接上一个 $S$，那么肯定会结束。但是我们可能在加入 $S$ 的过程的中间就结束了，此时一定是因为我们加入了 $S$ 的某个 `border`。我们设 $op_i$ 表示 $S_{[1,i]}$ 是不是 $S$ 的 `border`，那么我们可以写出式子：

$$(\frac{x}{n})^{|S|} G(x) = \sum_{i=1}^{|S|} op_i F(x) (\frac{x}{n})^{|S|-i}$$
$$G(x) = \sum_{i=1}^{|S|} op_i F(x) (\frac{n}{x})^{i}$$

代入 $x = 1$，得 $G(1) = \sum_{i=1}^{|S|} op_i F(1) n^i$，注意到 $F(1) = 1$，所以答案就是 $\sum_{i=1}^{|S|} op_i n^i$。

## LG 5326

你面前有 $n$ 个开关，每个开关有一个 `01` 状态，初始都为 `0`。

你每次按照 $p_i$ 的概率选择一个开关，使这个开关状态翻转。

给出目标状态，求出第一次得到这个目标状态的期望步数。 $(1 \le n \le 100)$

### Sol

我们对于每个开关分别考虑。

设 $F_i(x)$ 为第$i$个开关达到目标状态的期望的 `EGF`。

对于目标为 0 的开关：

$$F_i(x) = \sum_{k=0}^{+\infty} \frac{p_i^{2k} x^{2k}}{(2k)!} = \frac{e^{p_i x} + e^{-p_i x}}{2}$$

对于目标为 1 的开关：

$$F_i(x) = \sum_{k=0}^{+\infty} \frac{p^{2k+1} x^{2k+1}}{(2k+1)!} = \frac{e^{p_i x} - e^{-p_i x}}{2}$$

所以到达目标方案的期望的 `EGF` 为：

$$F(x) = \prod_{i=1}^{n} F_i(x)$$

但是我们发现我们要求的是第一次到达目标状态的期望，于是我们设 $G(x)$ 为做 $n$ 次变回原样的 `EGF`，设 $H(x)$ 为第一次到达目标状态的 `EGF`。

$$\therefore G(x)=\prod_{i=1}^{n} G_i(x)=\prod_{i=1}^{n}\frac{e^{p_i x}+e^{-p_i x}}{2}$$

又设 $f(x)$ 为 $F(x)$ 的 `OGF`，$g(x)$ 为 $G(x)$ 的 `OGF`，$h(x)$ 为 $H(x)$ 的 `OGF`。

所以我们知道 $h(x)g(x)=f(x)\Rightarrow h(x)=f(x)g(x)^{-1}$。

现在我们只需要求 $f(x)$ 和 $g(x)$ 就好了。

我们发现：如果 $F(x)$ 可以表示成 $\sum a_i e^{b_i x}$ 的形式，$f(x)$ 就为 $\sum \frac{a_i}{1-b_i x}$。

然后暴力用背包即可。