实验舱
青少年编程
走近科学 走进名校

提高算法班

概率与期望DP

Mas

随机试验 & 样本空间

随机试验 (random test) 是试验结果呈现出不确定性的试验

随机试验满足以下条件:

- 试验可在相同条件下重复进行
- 试验的可能结果不止一个且所有可能结果可事先预知
- 每次试验的结果只有一个，但不能事先预知

在一次随机试验 E 中可能发生的不可再细分的结果被称为 **样本输出/样本点** (sample point)

在随机试验中可能发生的所有样本输出的集合称为 **样本空间** (sample space) 用 Ω 表示

进行一次随机试验 E，其结果一定符合 Ω 中的恰好一个元素，不可能是零个或多个

掷六面骰子的随机试验中用点数表示样本输出

可能出现 6 个样本输出，样本空间可表示为 $\Omega = \{1,2,3,4,5,6\}$

随机事件

当随机实验的结果为某个样本点 ω ，若 $\omega \in A$ 则称事件 A 发生了否则称事件 A 未发生

随机事件 (random event) 是样本空间 Ω 的子集

它由样本空间 Ω 中的元素构成，用大写字母 $A, B, C, \ldots$ 表示

在掷两个六面骰子的随机试验中

设随机事件 A 为 “点数和大于 10”

那么 A 是 3 个样本输出组成的集合 $A = \{ (5,6), (6,5), (6,6) \}$

仅含有一个样本点的随机事件称为 **基本事件** (elementary event)

整个样本空间也是事件，称为 **必然事件** (certain event)

空集也是事件，称为 **不可能事件** (impossible event) 记为 $\emptyset$

随机事件

随机事件由集合定义，那么随机事件也可进行集合的运算

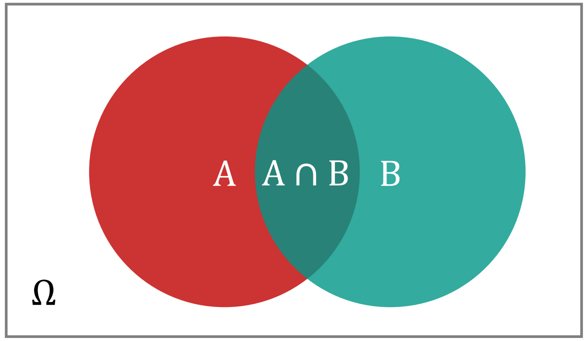
对于随机事件 A 与随机事件 B

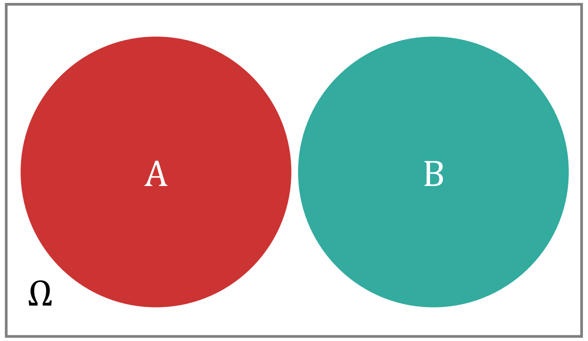
- $A \subset B$ 表示事件 A 发生必然导致事件 B 发生

  若 $A \subset B \wedge B \subset A$ 那么 $A = B$

- $A \cap B$ 表示 A, B 事件都发生，也可记为 AB
- $A \cup B$ 表示 A, B 事件至少发生一个，也可记为 $A + B$
- $A - B$ 表示 A 事件发生而 B 事件不发生，也可记为 $A\bar{B}$
- $A \mid B$ 表示 在 B 发生的前提下 A 发生
- 若 $A \cap B = \emptyset$ 称 A, B 为 **互斥事件** (mutually exclusive events)，也称 A, B **互不相容**
- 若 $A \cap B = \emptyset \ \wedge \ A \cup B = \Omega$ 称 A, B 为 **互逆/对立事件** (complementary events)

  $\bar{A}$ 表示事件 A 不发生，$\bar{A} \cap A$ 且 $A \cup \bar{A} = \Omega$ 故 $\bar{A}$ 与 A 为 **对立事件**





互斥事件

随机事件

对于 A, B, C 三个随机事件

如下给出事件的形式化描述

- 只有 A 发生: $A \cap \bar{B} \cap \bar{C}$ 或 $A - B - C$
- A, B, C 同时发生: $A \cap B \cap C$
- A, B, C 不同时发生 (A, B, C 至少有一个不发生) : $\overline{A \cap B \cap C}$ 或 $\bar{A} \cup \bar{B} \cup \bar{C}$
- A, B, C 同时不发生: $\bar{A} \cap \bar{B} \cap \bar{C}$
- A, B, C 恰有一个发生: $(A \cap \bar{B} \cap \bar{C}) \cup (\bar{A} \cap B \cap \bar{C}) \cup (\bar{A} \cap \bar{B} \cap C)$
- A, B, C 至少有一个发生: $A \cup B \cup C$ 或 $\overline{\bar{A} \cap \bar{B} \cap \bar{C}}$
- A, B, C 至多有一个发生: $(\bar{A} \cap \bar{B} \cap \bar{C}) \cup (A \cap \bar{B} \cap \bar{C}) \cup (\bar{A} \cap B \cap \bar{C}) \cup (\bar{A} \cap \bar{B} \cap C)$

事件运算

幂等律

$$A \cup A = A \qquad A \cap A = A$$

交换律

$$A \cup B = B \cup A \qquad A \cap B = B \cap A$$

结合律

$$(A \cup B) \cup C = A \cup (B \cup C)$$

$$(A \cap B) \cap C = A \cap (B \cap C)$$

分配律

$$A \cap (B \cup C) \cup C = (A \cap B) \cup (A \cap C)$$

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

对偶律/德摩根律(De Morgan′s laws)

$$\overline{A \cup B} = \overline{A} \cap \overline{B} \qquad \overline{A \cap B} = \overline{A} \cup \overline{B}$$

概率

古典定义

若试验满足以下两条，称其为古典试验

- 试验基本结果有限
- 试验的每个基本结果出现等可能

对于古典试验中的事件 A，它的概率定义为 $\mathrm{P(A)} = \frac{m}{n}$

其中 n 表示该试验中所有可能出现的基本结果的总数目

m 表示事件 A 包含的试验基本结果数

统计定义

在一定条件下进行了 n 次试验，事件 A 发生了 $\mathrm{N_A}$ 次

若随着 n 逐渐增大频率 $\frac{\mathrm{N_A}}{n}$ 逐渐稳定在某一数值 p 附近

概率

那么数值 p 称为事件 A 在该条件下发生的概率, 记做 $\mathrm{P(A)} = p$

概率的统计定义存在数学上的不严谨性，在实际中几乎不可能每一个事件做大量重复的试验来计算频率

苏联数学家柯尔莫哥洛夫 (Kolmogorov) 于 1933 年给出了概率的公理化定义

公理化定义

设 E 是随机试验 Ω 为其样本空间，对 Ω 的每一个事件 A 赋予一个 [0,1] 范围内的实数

记为 P(A) 称为事件 A 的 **概率** (probability)，P(A) 是一个从集合到实数的映射

对于一个事件 A 其发生概率 P(A) 满足以下公理

非负性

$$\mathrm{P(A)} \in [0,1]$$

规范性

样本空间的概率值为 1 即

概率公理

$$P(\Omega) = 1$$

可列可加性

若事件 $A_1, A_2, \cdots, A_n, \cdots$ 互不相容（即 $\forall i \neq j$ 都有 $A_i \cap A_j = \emptyset$）

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

还存在以下性质

- $P(\emptyset) = 0$

  证明

  $\Omega = \Omega \cup \emptyset \cup \emptyset \cdots \cup \emptyset \cup \cdots$，根据 **可列可加性** 有

  $$P(\Omega) = P(\Omega) + P(\emptyset) + P(\emptyset) \cdots + P(\emptyset) + \cdots$$

  $$\Rightarrow P(\emptyset) + P(\emptyset) \cdots + P(\emptyset) + \cdots = 0$$

概率性质

根据 **非负性** 必有

$$P(\emptyset) = 0$$

- **有限可加性**

对于有限个互不相容的事件 $A_1, A_2, \cdots, A_n$

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$$

证明

$A_1 \cup A_2 \cup \cdots \cup A_n = A_1 \cup A_2 \cup \cdots \cup A_n \cup \emptyset \cup \emptyset \cdots \cup \cdots$ ，根据 **可列可加性** 有

$$P(A_1 \cup A_2 \cup \cdots \cup A_n \cup \emptyset \cup \emptyset \cdots \cup \cdots) = P\left(\bigcup_{i=1}^{n} A_i\right) + P(\emptyset) + P(\emptyset) \cdots + P(\emptyset) + \cdots$$

$$= P\left(\bigcup_{i=1}^{n} A_i\right)$$

概率性质

- $P(\overline{A}) = 1 - P(A)$

  证明

  显然 $A, \overline{A}$ 互不相容 且 $A \cup \overline{A} = \Omega$

  根据 **规范性** 及 **有限可加性** 有

  $$P(\Omega) = P(A) + P(\overline{A}) = 1 \Rightarrow P(\overline{A}) = 1 - P(A)$$

- **单调性**

  若 $A \subset B$ 则有 $P(B - A) = P(B) - P(A)$ 且有 $P(A) \le P(B)$

  证明

  $B = A \cup (B - A)$ 显然 $A \cap (B - A) = \emptyset$ 即 $A, B - A$ 互不相容

  根据 **有限可加性** 有 $P(B) = P(A) + P(B - A)$, 又由于 **非负性**

  $$P(B - A) \ge 0 \Rightarrow P(B) - P(A) \ge 0 \Rightarrow P(B) \ge P(A)$$

- $P(A - B) = P(A) - P(AB)$

概率性质

证明

$A - B = A - AB$ 且 显然 $AB \subset A$

根据 **单调性** 有 $P(A\overline{B}) = P(A - AB) = P(A) - P(AB)$

- **加法公式**

$$P(A \cup B) = P(A) + P(B) - P(AB)$$

证明

$A \cup B = A \cup (B - AB)$ 显然 $A \cap (B - AB) = \emptyset$ 即 $A, B - AB$ 互不相容

根据 **有限可加性** 有 $P(A \cup B) = P(A) + P(B - AB) = P(A) + P(B) - P(AB)$

可推广至多个事件

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{r=1}^{\infty}\left((-1)^{r+1}\left(\sum_{i_1<\cdots<i_r} P\left(A_{i_1} \cdots A_{i_t}\right)\right)\right)$$

不难通过数学归纳法证明

条件概率

条件概率

P(B | A) 为事件 A 发生的前提下事件 B 发生的概率

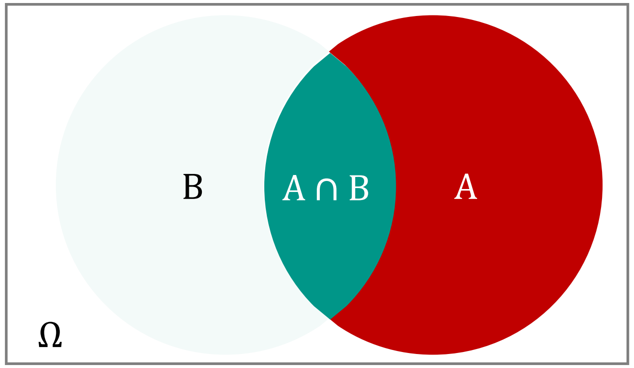
简称 **条件概率** (Conditional Probability)

$$P(B \mid A) = \frac{P(AB)}{P(A)}$$

当事件 A 发生后试验条件发生改变，新试验条件下 A 成为样本空间

A 的样本点具有等可能性且 A 发生后 AB 是 A 的子集

$$P(B|A) = \frac{n(AB)}{n(A)} = \frac{\frac{n(AB)}{n(\Omega)}}{\frac{n(A)}{n(\Omega)}} = \frac{P(AB)}{P(A)}$$



需保证 $P(A) > 0$

乘法公式

实验舱 青少年编程 走近科学 走进名校 LAB

乘法公式

若 $P(A) > 0$ 根据 **条件概率** 公式不难得出

$$P(AB) = P(A)\,P(B \mid A) = P(B)\,P(A \mid B)$$

若 $P(A_1A_2 \cdots A_{n-1}) > 0$ 有

$$P(A_1A_2 \cdots A_n) = P(A_1)\,P(A_2 \mid A_1)\,P(A_3 \mid A_1\,A_2)\ \cdots P(A_n \mid A_1\,A_2 \cdots A_{n-1})$$

上述公式被称为 **乘法公式** (muitipiicatme formula of probability)

证明

根据 **单调性** 有 $P(A_1) \geq P(A_1A_2) \geq \cdots \geq P(A_1A_2 \cdots A_{n-1}) > 0$

$$P(A_1)\,P(A_2 \mid A_1)\,P(A_3 \mid A_1\,A_2)\ \cdots P(A_n \mid A_1\,A_2 \cdots A_{n-1})$$

$$= P(A_1)\,\frac{P(A_1\,A_2)}{P(A_1)}\,\frac{P(A_1\,A_2\,A_3)}{P(A_1\,A_2)} \cdots \frac{P(A_1\,A_2 \cdots A_n)}{P(A_1\,A_2 \cdots A_{n-1})} = P(A_1A_2 \cdots A_n)$$

乘法公式

设罐中有 b 个黑球、r 个红球，每次随机取出一个球后放回，同时加入 c 个同色球和 d 个异色球

记事件 B_i 为第 i 次取出黑球，记事件 R_j 为第 j 次取出红球

若连续从罐中取出三个球其中有两个红球一个黑球，根据 **乘法公式** 有

$$P(B_1R_2R_3) = P(B_1)\,P(R_2 \mid B_1)\,P(R_3 \mid B_1\,R_2) = \frac{b}{b+r} \times \frac{r+d}{b+r+c+d} \times \frac{r+d+c}{b+r+2c+2d}$$

$$P(R_1B_2R_3) = P(R_1)\,P(B_2 \mid R_1)\,P(R_3 \mid R_1B_2) = \frac{r}{b+r} \times \frac{b+d}{b+r+c+d} \times \frac{r+d+c}{b+r+2c+2d}$$

$$P(R_1R_2B_3) = P(R_1)\,P(R_2 \mid R_1)\,P(B_3 \mid R_1R_2) = \frac{r}{b+r} \times \frac{r+c}{b+r+c+d} \times \frac{b+2d}{b+r+2c+2d}$$

不难发现概率与黑球再第几次取出有关

该问题也被称为 **波利亚罐子模型** (Polya′s urn scheme)

乘法公式

该问题存在一些特殊情况

当 $c=-1, d=0$ 时，即为 **不放回抽样**，此时前次抽取结果将影响后次抽取结果

但只要抽出红球和黑球个数确定，概率不依赖抽出球的顺序

$$P(B_1R_2R_3)=P(R_1B_2R_3)=P(R_1R_2B_3)=\frac{br(r-1)}{(b+r)(b+r-1)(b+r-2)}$$

当 $c=0, d=0$ 时称为 **放回抽样**，此时前次抽取结果不影响后次抽取结果，概率都相等

$$P(B_1R_2R_3)=P(R_1B_2R_3)=P(R_1R_2B_3)=\frac{br^2}{(b+r)^3}$$

当 $c>0, d=0$ 时称为 **传染病模型**，此时每次取出球都会增加下一次取出该颜色球的概率

$$P(B_1R_2R_3)=P(R_1B_2R_3)=P(R_1R_2B_3)=\frac{br(r+c)}{(b+r)(b+r+c)(b+r+2c)}$$

乘法公式

可以发现 $d=0$ 时，只要抽出红球和黑球个数确定，概率不依赖抽出球的顺序，概率都相同

当 $c=0, d>0$ 时，称为 **安全模型**，可理解为：

每当安全事故发生（红球被取出）安全工作就抓紧，下次出现事故的概率降低

而当安全事故未发生（黑球被取出）安全工作就放松，下次出现事故的概率增大

此时上述三种概率为

$$P(B_1R_2R_3)=\frac{b}{b+r}\times\frac{r+d}{b+r+d}\times\frac{r+d}{b+r+2d}$$

$$P(R_1B_2R_3)=\frac{r}{b+r}\times\frac{b+d}{b+r+d}\times\frac{r+d}{b+r+2d}$$

$$P(R_1R_2B_3)=\frac{r}{b+r}\times\frac{r}{b+r+d}\times\frac{b+2d}{b+r+2d}$$

#2848、生日悖论

题目描述

假设你在一个有 23 个人的聚会上

聚会中至少有两个人生日相同的概率是多少?

令人惊讶的是,结果超过了 0.5

你现在在其他星球上一年有 N 天

你必须找到必须邀请参加聚会的最少人数

以使聚会中至少有两个人生日相同的概率至少为 0.5

输入格式

第一行输入一个正整数 T 表示 T 组询问

接下来每行一个正整数 N ,表示这个星球一年有 N 天

输出格式

每组询问输出一行一个整数,表示最少需要邀请的人数

输入样例

```
2
365
669
```

输出样例

```
22
30
```

数据规模

对于全部的数据 $1 \le T \le 20000, 1 \le N \le 10^5$

记第 i 人与前 $i-1$ 人生日都不同为事件 A_i

显然 $P(A_1) = 1$

#2848、生日悖论

根据 **乘法公式**

$$P(A_1A_2\cdots A_i)=P(A_1)\,P(A_2\mid A_1)\,P(A_3\mid A_1\ A_2)\ \cdots P(A_n\mid A_1\ A_2\cdots A_{n-1})$$

第二个人有 $n-1$ 种选择，第三个人有 $n-2$ 种选择

即 $P\left(A_\mathrm{j}\mid A_1\ A_2\cdots A_{j-1}\right)=\frac{n-j+1}{n}$

$$P(A_1A_2\cdots A_i)=\frac{n}{n}\times\frac{n-1}{n}\times\cdots\times\frac{n-i}{n}$$

显然事件 $\overline{A_1A_2\cdots A_i}$ 表示前 i 人中至少两人生日相同

那么 $P(\overline{A_1A_2\cdots A_i})=1-P(A_1A_2\cdots A_i)$

直接递推计算即可

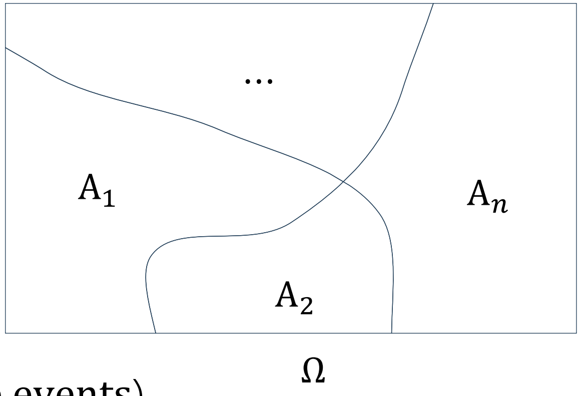
经过计算 $N=100000$ 仅需 372 人，当概率超过 0.5 时停止即可

全概率公式

全概率公式

若随机事件 $A_1, A_2, \cdots, A_n$ 满足

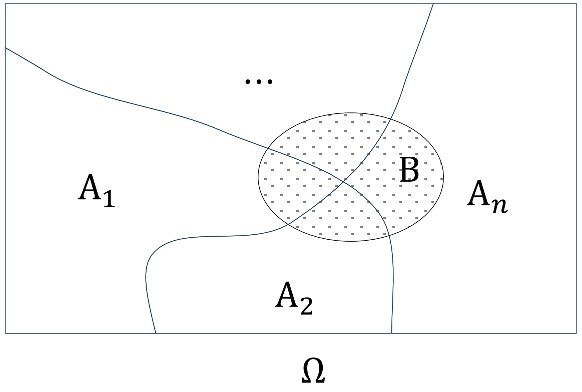
- $\forall i \neq j$ 都有 $A_i \cap A_j = \emptyset$
- $\Omega = \bigcup_{i=1}^{n} A_i$



那么称 $A_1, A_2, \cdots, A_n$ 为样本空间 Ω 的一个 **分割** ，或称为 **完备事件组** (collectively exhaustive events)

若事件 $A_1, A_2, \cdots, A_n$ 为 **完备事件组** 且 $P(A_i) > 0$, 对于任意事件 B 有

$$P(B) = \sum_{i=1}^{n} \left(P(A_i)\, P(B \mid A_i) \right)$$



证明

不难发现

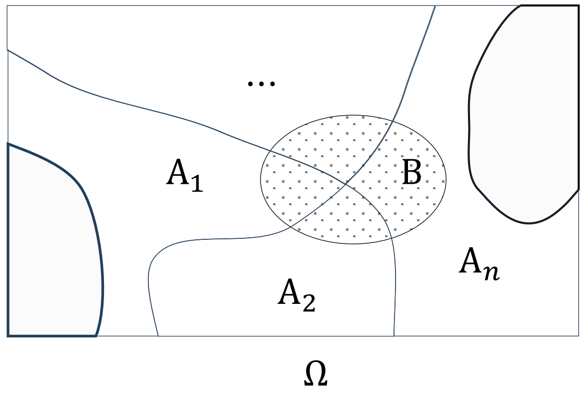
$$B = B \cap \Omega = B \cap \left(\bigcup_{i=1}^{n} A_i \right) = \left(\bigcup_{i=1}^{n} BA_i \right)$$

全概率公式

而 $B\cap A_1, B\cap A_2, \cdots, B\cap A_n$ 互不相容，根据 **有限可加性** 有

$$P(B) = \sum_{i=1}^{n} P(BA_i)$$

由于已保证 $P(A_i) > 0$，根据 **乘法公式** 有 $P(BA_i) = P(A_i)\,P(B\mid A_i)$



$$P(B) = \sum_{i=1}^{n} P(BA_i) = \sum_{i=1}^{n}\big(P(A_i)\,P(B\mid A_i)\big)$$

称上式为 **全概率公式**（total probability theorem）

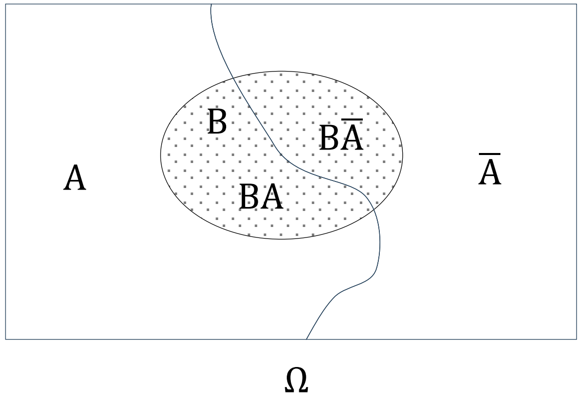
将 **完备事件组** 换为 $A_1, A_2, \cdots, A_n$ **互不相容** 且 $B \subset \bigcup_{i=1}^{n} A_i$ 时 **全概率公式** **依然成立**

可理解为： B 为某一过程的结果，将事件 $A_1, A_2, \cdots, A_n$ 视为产生该结果的若干原因

根据 **全概率公式** 可将复杂事件分解为若干 **互不相容** 的简单事件

全概率公式

在 n 张彩票中仅有一张可中奖，求第二个人中奖的概率



设第 i 人中奖的事件为 A_i，要求出 $P(A_2)$

A_1 是否发生影响 A_2 的发生，有

$$P(A_2 \mid A_1) = 0 \qquad P(A_2 \mid \overline{A_1}) = \frac{1}{n-1}$$

显然 $P(A_1) = \frac{1}{n} > 0 \quad P(\overline{A_1}) = \frac{n-1}{n} > 0$，根据 **全概率公式**

$$P(A_2) = P(A_1) \times P(A_2 \mid A_1) + P(\overline{A_1}) \times P(A_2 \mid \overline{A_1}) = \frac{1}{n} \times 0 + \frac{n-1}{n} \times \frac{1}{n-1} = \frac{1}{n}$$

这表明中奖概率与先后次序无关，类似的可以得出 $P(A_3) = \cdots = P(A_n) = \frac{1}{n}$

后者可能处于不利局面(前者中奖)，也可能处于有利局面(前者未中奖增大了中奖机会)

但经过 **全概率公式**(加权平均)综合后机会均等

若有 n 张彩票，其中 k 张可中奖 $P(A_1) = P(A_2) = \cdots = P(A_n) = \frac{k}{n}$

未成年人不得购买彩票及兑奖

Bayes 公式

Bayes 公式

若事件 $A_1, A_2, \cdots, A_n$ 为 **完备事件组**，同时有 $P(B) > 0$ 且 $P(A_i) > 0$

$$P(A_i \mid B) = \frac{P(A_i)\,P(B \mid A_i)}{\sum_{j=1}^{n}\left(P(A_j)\,P(B \mid A_j)\right)}$$

证明

根据 **条件概率** 定义

$$P(A_i \mid B) = \frac{P(A_i B)}{P(B)}$$

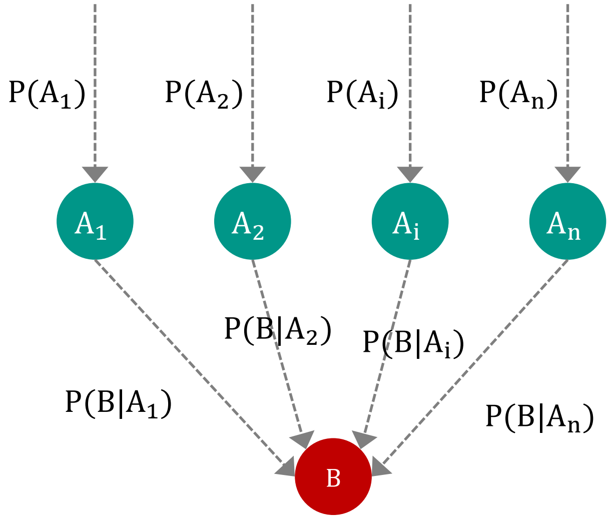
根据 **全概率公式** $P(B) = \sum_{j=1}^{n}\left(P(A_j)\,P(B \mid A_j)\right)$

$$P(A_i \mid B) = \frac{P(A_i B)}{P(B)} = \frac{P(A_i B)}{\sum_{j=1}^{n}\left(P(A_j)\,P(B \mid A_j)\right)}$$

Bayes 公式

再根据 **乘法公式** $P(A_iB) = P(A_i)\,P(B\mid A_i)$

$$P(A_i \mid B) = \frac{P(A_i)\,P(B\mid A_i)}{\sum_{j=1}^{n}\left(P(A_j)\,P(B\mid A_j)\right)}$$



上述公式被称为 **Bayes 公式** (Bayes′ theorem)

将 **Bayes 公式** 中的 $P(A_i)$ 称为 A_i 的 **先验概率**

将 **Bayes 公式** 中的 $P(A_i \mid B)$ 称为 A_i 的 **后验概率**

Bayes 公式 专门用于计算 **后验概率** ，即通过 B 的发生这个新信息来修正 A_i 的概率

《伊索寓言》中有一则“孩子与狼”的故事，使用 **Bayes 公式** 分析此寓言中村民对这个小孩的可信度是如何下降的

记事件 A 为小孩说谎，事件 B 为小孩可信

Bayes 公式

设过去村民对小孩的印象 $P(B) = 0.8 \quad P(\overline{B}) = 0.2$

考虑使用 **Bayes 公式** 求出 $P(B \mid A)$ 即小孩说了一次谎后，村民对其信任度的改变

设 $P(A \mid B) = 0.1$ 表示小孩可信且说谎的概率为 0.1， $P(A \mid \overline{B}) = 0.5$ 表示小孩不可信且说谎的概率为 0.5

第一次村民并没有发现狼即小孩说谎了，根据 **Bayes 公式** 有

$$P(B \mid A) = \frac{P(B)\,P(A \mid B)}{P(B)\,P(A \mid B) + P(\overline{B})\,P(A \mid \overline{B})} = \frac{0.8 \times 0.1}{0.8 \times 0.1 + 0.2 \times 0.5} = 0.444$$

此时村民对小孩的印象修正为 $P(B) = 0.444, P(\overline{B}) = 0.556$

在此基础上，第二次村民没有发现狼即小孩又说谎了，根据 **Bayes 公式** 有

$$P(B \mid A) = \frac{0.444 \times 0.1}{0.444 \times 0.1 + 0.556 \times 0.5} = 0.138$$

此时村民对小孩的信任度从 $0.8 \rightarrow 0.138$

所以第三次发生悲剧 !!!

#3975、 盒子取球

题目描述

有四个一样的不透明盒子

第一个盒子内有 n_1 个小球，其中 w_1 个白球，剩下 $n_1 - w_1$ 个都是黑球

第二个盒子内有 n_2 个小球，其中 w_2 个白球，剩下 $n_2 - w_2$ 个都是黑球

第三个盒子内有 n_3 个小球，其中 w_3 个白球，剩下 $n_3 - w_3$ 个都是黑球

第四个盒子内有 n_4 个小球，其中 w_4 个白球，剩下 $n_4 - w_4$ 个都是黑球

小球质地均匀，从盒子外部无法看到盒内的情况

现在 Mas 随机从四个箱子中取出一个小球，请你计算有多大的概率取出白球?

输入格式

第一行输入两个空格分隔的整数 n_1, w_2

第二行输入两个空格分隔的整数 n_2, w_2

第三行输入两个空格分隔的整数 n_3, w_3

第四行输入两个空格分隔的整数 n_4, w_4

输出格式

输出一个实数表示答案

答案并非

$$\frac{w_1 + w_2 + w_3 + w_4}{n_1 + n_2 + n_3 + n_4}$$

若第一个箱子中有 94 个球，其中 1 个白球

第二个箱子中有 2 个球，其中 1 个白球;第三个箱子中有 2 个球，其中 1 个白球;第四个箱子中有 2 个球，其中 1 个白球

取出白球的概率并非 4%

凭直觉 或 **蒙特卡罗方法** (Monte Carlo method) 都可察觉出 4% 过低

蒙特卡罗方法 也称统计模拟方法，是一种用随机数 (或伪随机数) 来解决计算问题的方法

#3975、 盒子取球

设事件 A_i 表示选择箱子 i , 事件 B 表示选出白球 , 那么事件 $B \mid A_1$ 表示从第 i 个箱子取出白球

显然 $P(A_1) = P(A_2) = P(A_3) = P(A_4) = \frac{1}{4}$

同时 $P(B \mid A_1) = \frac{w_1}{n_1}$ $P(B \mid A_2) = \frac{w_2}{n_2}$ $P(B \mid A_3) = \frac{w_3}{n_3}$ $P(B \mid A_4) = \frac{w_4}{n_4}$

根据 **全概率公式**

$$P(B) = \frac{1}{4} \times \frac{w_1}{n_1} + \frac{1}{4} \times \frac{w_2}{n_2} + \frac{1}{4} \times \frac{w_3}{n_3} + \frac{1}{4} \times \frac{w_4}{n_4}$$

时间复杂度 $O(1)$

已知取出白球 , 从第 i 个箱子取出的概率为 ?

根据 **Bayes 公式** 求解即可

$$P(A_i \mid B) = \frac{P(A_i)\, P(B \mid A_i)}{P(B)}$$

事件的独立性

一般情况下 $P(B \mid A) = \frac{P(AB)}{P(A)} \neq P(B)$ 即事件 A 发生对事件 B 发生产生影响

然而在有些情况下，事件 A 的发生对事件 B 的发生没有任何影响，即 $P(B \mid A) = P(B)$

根据 **条件概率公式**

$$P(B) = P(B \mid A) = \frac{P(AB)}{P(A)} \Rightarrow P(AB) = P(A)\,P(B)$$

此时称事件 A 与事件 B **相互独立** (mutually independent)

若事件 A 与事件 B 相互独立，那么 A 与 $\bar{B}$ 相互独立、 $\bar{A}$ 与 B 相互独立、$\bar{A}$ 与 $\bar{B}$ 相互独立

证明

根据概率性质有 $P(A\bar{B}) = P(A - AB) = P(A) - P(AB)$

又由于事件 A 与事件 B 相互独立 有 $P(AB) = P(A)\,P(B)$，那么

$$P(A\bar{B}) = P(A - B) = P(A) - P(AB) = P(A) - P(A)\,P(B) = P(A)\left(1 - P(B)\right) = P(A)\,P(\bar{B})$$

不难证明其它情况

独立性 & 互斥

从一副扑克 (不含大/小王) 中随机抽取一张牌，事件 A 表示抽到 10，事件 B 表示抽到黑色牌事件 A 与 B 是否相互独立?

事件 AB 表示抽到黑色 10 有 $P(AB) = \frac{2}{52}$，而 $P(A) = \frac{4}{52}, P(B) = \frac{1}{2}$ 显然 $P(AB) = P(A)\,P(B)$ 故事件 A 与 B **相互独立**

也可根据实际情况判断事件是否独立

事件 A 与事件 B 相互独立: 与概率相关，反映事件的概率属性

$$P(AB) = P(A)\,P(B)$$

事件 A 与事件 B 互不相容: 与概率无关，与事件的运算相关

$$A \cap B = \emptyset \Rightarrow P(A \cap B) = 0$$

即互斥与独立并非同一概念，如上文抽取扑克牌，事件 A 与 事件 B **独立** 但 **不互斥**

若 $P(A) \neq 0 \land P(B) \neq 0$ 不难证明

- 事件 A 与事件 B **独立**，那么 A 和 B 不 **互斥**
- 事件 A 与事件 B **互斥**，那么 A 和 B 不 **独立**

#1665、留学offer

实验舱 青少年编程 LAB 走近科学 走进名校

题目描述

Mas 想出国，他需要去申请学校了

要申请国外的任何大学要交纳一定的申请费用,这可是很惊人的

Mas 没有多少钱，总共只攒了 n 万美元

他将在 m 个学校中选择若干的(当然要在他的经济承受范围内)

每个学校都有不同的申请费用 a (万美元),并且 Mas 估计了他得到这个学校 offer 的可能性 b

不同学校之间是否得到 offer 不会互相影响

请你计算 Mas 可以收到至少一份 offer 的最大概率

如果 Mas 选择了多个学校，得到任意一个学校的 offer 都可以

输入格式

第一行有两个正整数 n, m

后面的 m 行,每行都有两个数据 a_i (整型), b_i (实型)

分别表示第 i 个学校的申请费用和可能拿到 offer 的概率

输出格式

输出一个实数表示 Mas 可能得到至少一份 offer 的最大概率

用百分数表示,精确到小数点后一位

数据规模

对于全部的数据 $0 \le n \le 10000, 0 \le m \le 10000$

设 $\mathrm{dp}[i][j]$ 为考虑大学 $1 \sim i$ 资金为 j 时的最小失败概率

显然 $dp[0][*] = 1$

不同大学间申请相互独立

$$\mathrm{dp}[i][j] = \min\left(\mathrm{dp}[i-1][j], \mathrm{dp}[i-1][j-a_i] \times (1-b_i)\right)$$

最终答案为 $1 - \mathrm{dp}[m][n]$

本质为 $0-1$ 背包，时间复杂度 $O(nm)$

#2847、抛硬币

题目描述

有 n 枚硬币排成一排

现在同时抛所有硬币

第 i 枚硬币向上的概率是 p_i

向下的概率是 $1-p_i$

求向上的硬币数量比向下的多的概率

输入格式

第一行输入一个正整数 n

第二行输入 n 个浮点数 p_i

输出格式

输出一个浮点数(保留 12 位小数),表示向上的硬币数量比向下的多的概率

数据规模

对于全部的数据 $1 \le n \le 3000, 0 \le p_i \le 1$,保证 n 为奇数

设 dp[i][j] 表示仅考虑前 i 个硬币有 j 枚硬币朝上的概率

显然抛硬币互不干扰，即事件间独立

若第 j 枚朝上有

$$\mathrm{dp}[i-1][j-1] \times p_i$$

若第 j 枚朝下有

$$\mathrm{dp}[i-1][j] \times (1-p_i)$$

根据 **全概率公式** ，状态转移方程

$$\mathrm{dp}[i][j] = \mathrm{dp}[i-1][j-1] \times p_i + \mathrm{dp}[i-1][j] \times (1-p_i)$$

最终答案为

$$\sum_{i=0}^{n} (\,[i > n-i] \times \mathrm{dp}[n][i]\,)$$

#2852、生物繁衍

题目描述

一开始有 k 只特殊生物，这种特殊生物只能存活 1 天

当时在其死亡时会产生 $[0, n-1]$ 范围内只生物

其中有 p_i 的概率产生 i 只这种生物(也只能活一天)

请你求出 m 天内所有生物都死亡的概率(包括 m 天前死亡的情况)

输入格式

第一行输入一个整数 T，表示 T 组询问

每组询问第一行输入三个整数 n, k, m

每组询问第二行输入 $n-1$ 个浮点数 p_i

输出格式

对于每组询问输出一行

每行输出一个浮点数(保留 7 位小数),表示 m 天内所有生物都死亡的概率

记事件 A_i 表示初始仅有一只 i 天内全部死亡

每只生物的死亡与否互不干扰

仅需考虑一只的情况 $P(A_i)$

k 只全部死亡的概率为 $P(A_i)^k$

数据规模

对于全部的数据 $1 \le T \le 100, 1 \le n, k \le 10, 1 \le m \le 1000$

#2852、生物繁衍

初始时 $\mathrm{P}(\mathrm{A}_1) = p_0$

记事件 B_j 表示初始仅有一只生物产生 j 只生物，有 $\mathrm{P}(\mathrm{B}_j) = p_j$

$\mathrm{A}_i \mid \mathrm{B}_j$ 表示初始仅有一只生物产生 j 只生物 i 天内全部死亡

由于生物之前的死亡对后续生物不产生干扰，仅需令这 j 只在 $i-1$ 天内死亡，即

$$\mathrm{P}(\mathrm{A}_i \mid \mathrm{B}_j) = \mathrm{P}(\mathrm{A}_{i-1})^j$$

根据 **全概率公式**

$$\mathrm{P}(\mathrm{A}_i) = \sum_{j=0}^{n-1} \left(\mathrm{P}(\mathrm{B}_j)\mathrm{P}(\mathrm{A}_i \mid \mathrm{B}_j)\right) = \sum_{j=0}^{n-1} \left(\mathrm{P}(\mathrm{B}_j)\, \mathrm{P}(\mathrm{A}_{i-1})^j \right)$$

令 $\mathrm{dp}_i = \mathrm{P}(\mathrm{A}_i)$ 递推求解即可，答案为 dp_m^k

时间复杂度 $O(nm)$

随机变量

随机变量 (random variable) 是取值由随机事件决定的变量，是从样本空间 $\Omega = \{\omega\}$ 到实数集 $\mathbb{R}$ 的映射

即 随机变量 $X : \Omega(X) \to \mathbb{R}$，而 X 实际是 $X(\omega)$ 的简记

随机变量 X 取值 α (简记 $X = \alpha$) 可理解为某事件发生时该随机变量取值为 α

随机变量按其值域是否可数分为

- **离散型随机变量** (discrete random variable)
- **连续型随机变量** (continuous random variable)

若随机变量 X 的取值 **有限** 或 **无穷可列**，那么 X 称为 离散型随机变量

设 X 表示 “骰子的点数” $\Omega = \{1,2,3,4,5,6\}$ 那么 $X(\omega) = i$

该随机变量取值有限

设 X 表示 “硬币抛出第一个正面前会抛出多少反面” $\Omega = \{\omega_i \mid i = 0,1,\ldots\}$ 那么 $X(\omega_k) = k$

该随机变量取值无穷可列

本课主要讨论 离散型随机变量

随机变量

$X = \alpha$ 对应着一个能实现该命题的 基本事件 集合，也有与之对应的概率 $P(X = \alpha)$

由于事件都为 基本事件, 那么事件必然 互不相容 根据 **有限可加性**

$$P(X = \alpha) = \sum_{X(\omega) = \alpha} P(\omega)$$

若 $X_1(\omega), X_2(\omega), \cdots, X_n(\omega)$ 是定义在同一样本空间 $\Omega = \{\omega\}$ 上的 n 个随机变量

则称

$$\mathbf{X} = \left(X_1(\omega), X_2(\omega), \cdots, X_n(\omega)\right)$$

为 **n 元随机变量 / n 维随机变量** (n − dimensional random vector)

对于分别在样本空间 $\Omega_1 = \{\omega_1\}$， $\Omega_2 = \{\omega_2\}$ 的随机变量，仅能在乘积空间

$$\Omega_1 \times \Omega_1 = \{ (\omega_1, \omega_2) : \omega_1 \in \Omega_1, \omega_2 \in \Omega_2\}$$

中讨论

随机变量

对于 **二元随机变量** X 和 Y 都在同一空间，即 ω 都为同一 样本点/事件 但两者映射的值可能不同

$(X=\alpha, Y=\beta)$ 也对应着一个能实现该命题的 基本事件 集合，也有与之对应的概率 $P(X=\alpha, Y=\beta)$

$$P(X=\alpha, Y=\beta)=\sum_{\substack{X(\omega)=\alpha \\ Y(\omega)=\beta}} P(\omega)$$

存在如下性质

$$\sum_{\beta} P(X=\alpha, Y=\beta)=\sum_{\beta} \sum_{\substack{X(\omega)=\alpha \\ Y(\omega)=\beta}} P(\omega)=\sum_{X(\omega)=\alpha} P(\omega)=P(X=\alpha)$$

考察 $Y(\omega)$ 所有取值 β 时，若 ω 不能使得 $X(\omega)=\alpha$ 则 样本点 对应概率不被计入

仅有 $X(\omega)=\alpha$ 的事件 ω 的概率被计入，既

$$P(X=\alpha)=\sum_{\beta} P(X=\alpha, Y=\beta)$$

数学期望

若 $\Omega = \{\omega\}$ 表示 X 所在的样本空间，**离散型随机变量** X 的 **数学期望** (expected value) 定义如下

$$E(X) = \sum_{\omega \in \Omega} \big(X(\omega) \times P(\omega)\big)$$

记 Range(X) 为随机变量 X 的值域

若在至于范围内考察 α 也可描述为

$$E(X) = \sum_{\alpha \in \mathrm{Range}(X)} \big(\alpha \times P(X = \alpha)\big)$$

数学期望就是随机变量 X 的取值与概率的乘积之和（X 的取值按概率的加权平均）

也将 数学期望 称为 **均值** (mean)

数学期望在实际中应用广泛，E(X) 常作为 X 的的分布的代表（一种统计指标）参与同类指标的比较

数学期望

有一种押注游戏，其规则如下: 庄家从 6 副 (每副 52 张) 扑克中随机发给玩家两张

如果玩家下注 a 元，当得到的两张牌是一对时庄家赔你十倍；否则输掉玩家的赌注

如果玩家下注 100 元，玩家和庄家在每局中各期望赢多少元?

设随机变量 X, Y 分别表示玩家和庄家在一局中的收益，$a = 100$

$$P(X = 10a) = \frac{13C_{4\times 6}^2}{C_{52\times 6}^2} = 0.074 \qquad P(X = -a) = 1 - 0.074$$

那么

$$E(X) = 0.074 \times 10a + (1 - 0.074) \times -a = -18.6$$

$$E(Y) = 0.074 \times -10a + (1 - 0.074) \times a = 18.6 = -E(X)$$

当只使用一副扑克，可以计算出玩家每局期收益 -35.32 元

严禁赌博

期望性质

实验舱 青少年编程 走近科学 走进名校

- 若 a,b 为常数 X, Y 为随机变量

$$\mathrm{E}(a\mathrm{X}+b\mathrm{Y})=a\mathrm{E}(\mathrm{X})+b\mathrm{E}(\mathrm{Y})$$

证明

$$\mathrm{E}(a\mathrm{X}\ +\ b\mathrm{Y})=\sum_{x\,\in\,\mathrm{Range(X)}}\ \sum_{y\,\in\,\mathrm{Range(Y)}}\big((ax+by)\times \mathrm{P}(\mathrm{X}=x,\mathrm{Y}=y)\,\big)$$

$$=a\times\left(\sum_{x\,\in\,\mathrm{Range(X)}}\left(x\times\sum_{y\,\in\,\mathrm{Range(Y)}}\mathrm{P}(\mathrm{X}=x,\mathrm{Y}=y)\right)\right)+b\times\left(\sum_{y\,\in\,\mathrm{Range(Y)}}\left(y\times\sum_{x\,\in\,\mathrm{Range(X)}}\mathrm{P}(\mathrm{X}=x,\mathrm{Y}=y)\right)\right)$$

$$=a\times\left(\sum_{x\,\in\,\mathrm{Range(X)}}\big(x\times\ \mathrm{P}(\mathrm{X}=x)\big)\right)+b\times\left(\sum_{y\,\in\,\mathrm{Range(Y)}}\big(y\times \mathrm{P}(\mathrm{Y}=y)\big)\right)=a\mathrm{E}(\mathrm{X})+b\mathrm{E}(\mathrm{Y})$$

利用该性质可将一个随机变量拆分成若干个随机变量

分别求这些随机变量的期望值，最后相加得到期望

期望性质

特殊的，若 c 为常数有

$$\mathrm{E}(c)=c \qquad \mathrm{E}(\mathrm{X}+c)=\mathrm{E}(\mathrm{X})+c$$

- 若 X, Y 为相互独立的随机变量（即对于任意使得 $\mathrm{X}=\alpha$ 和 $\mathrm{Y}=\beta$ 成立的事件都相互独立）

$$\mathrm{E}(\mathrm{XY})=\mathrm{E}(\mathrm{X})\,\mathrm{E}(\mathrm{Y})$$

证明

$$\mathrm{E}(\mathrm{XY})=\sum_{\substack{\alpha\in\mathrm{Range}(\mathrm{X})\\ \beta\in\mathrm{Range}(\mathrm{Y})}}\big(\alpha\times\beta\times\mathrm{P}(\mathrm{X}=\alpha,\mathrm{Y}=\beta)\big)=\sum_{\substack{\alpha\in\mathrm{Range}(\mathrm{X})\\ \beta\in\mathrm{Range}(\mathrm{Y})}}\big(\alpha\times\beta\times\mathrm{P}(\mathrm{X}=\alpha)\,\mathrm{P}\,(\mathrm{Y}=\beta)\big)$$

$$=\left(\sum_{\alpha\in\mathrm{Range}(\mathrm{X})}\big(\alpha\times\mathrm{P}(\mathrm{X}=\alpha)\big)\right)\times\left(\sum_{\beta\in\mathrm{Range}(\mathrm{Y})}\big(\beta\times\mathrm{P}(\mathrm{Y}=\beta)\big)\right)$$

$$=\mathrm{E}(\mathrm{X})\,\mathrm{E}(\mathrm{Y})$$

期望性质

在掷两枚骰子的点数实验中，样本空间是由 36 个样本输出组成的集合，每个样本点可以写作 (a,b)，其中 $1 \le a,b \le 6$

定义“掷出的点数之和”为 X ,那么随机变量 X 的取值为 $2 \sim 12$

随机事件可描述为“掷出 X 点”，即由 $a+b=\mathrm{X}$ 的样本点 (a,b) 构成的子集

掷出 8 点的概率 $\mathrm{P(X=8)}=\frac{5}{36}$，则掷出的点数的数学期望为

$$\frac{1}{36}\times 2+\frac{2}{36}\times 3+\frac{3}{36}\times 4+\frac{4}{36}\times 5+\frac{5}{36}\times 6+\frac{6}{36}\times 7+\frac{5}{36}\times 8+\frac{4}{36}\times 9+\frac{3}{36}\times 10+\frac{2}{36}\times 11+\frac{1}{36}\times 12=7$$

设随机变量 X 表示掷一枚骰子的点数，其期望值为

$$\mathrm{E(X)}=\frac{(1+2+3+4+5+6)}{6}=3.5$$

掷两枚骰子的点数可表示为随机变量 2X ,则有

$$\mathrm{E(2X)}=\mathrm{E(X+X)}=3.5+3.5=7$$

#2851、格点染色

题目描述

有 n 个格点，初始是全为白色

进行 m 次操作，每次等概率随机选一个格点染黑

> 已经染黑的可以再次被选择染黑

求操作完后期望有多少点被染黑了

输入格式

第一行输入一个正整数 T，表示 T 组询问

接下来 T 行,每行输入两个正整数 n, m

输出格式

对于每组询问输出一行,输出一个浮点数(保留 5 位小数)

表示期望的黑色格点数量

数据规模

对于全部数据 $1 \le T \le 1000, 1 \le n, m \le 10^5$

思路1

设随机变量 $X_i \in \{0,1\}$ 表示第 i 个点是否被染黑

根据 期望可加性

$$E(X_1 + X_2 + \cdots + X_n) = E(X_1) + E(X_2) + \cdots + E(X_n)$$

其中 $E(X_i) = 1 \times P(X_i = 1) + 0 \times P(X_i = 0)$

第 i 个点从未被染黑的概率 $P(X_i = 0) = \left(\frac{n-1}{n}\right)^m$

那么 $P(X_i = 1) = 1 - \left(\frac{n-1}{n}\right)^m$

即 $E(X_i) = 1 - \left(\frac{n-1}{n}\right)^m$，答案为 $n - n \times \left(\frac{n-1}{n}\right)^m$

若使用快速幂求解 $\left(\frac{n-1}{n}\right)^m$，一次询问时间复杂度 $O(\log m)$

#2851、格点染色

思路2

设事件 A_i 表示第 i 个点涂成黑色的概率

令 $dp_i = P(A_i)$ 初始时有 $dp_1 = 1$, 根据 全概率公式

$$dp_i = dp_{i-1} \times P(A_i \mid A_{i-1}) + (1 - dp_{i-1}) \times P(A_i \mid \overline{A_{i-1}})$$

由于各次涂色独立且格子无区别 , **第 $i-1$ 次涂黑且第 i 次未涂黑** 与 **第 $i-1$ 次未涂黑且第 i 次涂黑** 等价

可理解将两次操作对调 , 即 $P(A_i \mid \overline{A_{i-1}}) = dp_{i-1}$

若 A_{i-1} 发生那么 A_i 发生的概率将减少 $\frac{1}{n}$, 即 $P(A_i \mid A_{i-1}) = dp_{i-1} - \frac{1}{n}$

设随机变量 $X_i \in \{0,1\}$ 表示第 i 次涂色是否成功 , 根据 期望可加性

$$E(X_1 + X_2 + \cdots + X_m) = E(X_1) + E(X_2) + \cdots + E(X_m)$$

其中 $E(X_i) = 1 \times P(X_i = 1) + 0 \times P(X_i = 0) = dp_i$

答案为 $\sum_{i=1}^{m} dp_i$, 一次询问时间复杂度 $O(m)$

期望性质

条件期望

当 $X = \alpha$ 时随机变量 Y 的 **条件期望** (conditional expectation) 以 $E(Y \mid X = \alpha)$ 表示

$$E(Y \mid X = \alpha) = \sum_{\beta} \left(\beta \times P(Y = \beta \mid X = \alpha)\right)$$

根据条件概率公式

$$E(Y \mid X = \alpha) = \sum_{\beta} \left(\beta \times \frac{P(Y = y, X = \alpha)}{P(X = \alpha)}\right)$$

$E(Y \mid X = \alpha)$ 为一数值，**而 $E(Y \mid X)$ 为一随机变量**

特殊的有

$$E(X) = E(X \mid A)\, P(A) + E(X \mid \bar{A})\, P(\bar{A})$$

期望计算

硬币抛出正面之前期望要抛出多少次反面?

令随机变量 X 表示 “抛出正面前的反面次数” ，令事件 A 表示 “第一次抛出了正面” 那么

$$E(X) = P(A)\,E(X \mid Y = A) + P(\bar{A})\,E(X \mid Y = \bar{A})$$

不难想到 $E(X \mid Y = A) = 0$

而 $E(X \mid Y = \bar{A}) = E(1 + X) = 1 + E(X)$

第一次失败后第二次开始 与 重新开始 无异 , 根据期望的可加性 $E(1 + X) = 1 + E(X)$

$$E(X) = \frac{1}{2} \times 0 + \frac{1}{2} \times \left(1 + E(X)\right)$$

解得 $E(X) = 1$, 即期望要抛 2 次才能抛出正面

推广到不停地做一件事情 , 每次有 p 的概率成功 , 期望在第 $\frac{1}{p}$ 次时成功

期望性质

全期望定理

$$E(Y) = E\big(E(Y \mid X)\big)$$

证明

记 I(X) 为随机变量 X 的值域

$$E\big(E(Y \mid X)\big) = \sum_{\alpha \in I(X)} \big(P(X=\alpha) \times E(Y \mid X=\alpha)\big) = \sum_{\alpha \in I(X)} \left(P(X=\alpha) \times \sum_{\beta \in I(Y)} \left(\beta \times \frac{P(Y=\beta, X=\alpha)}{P(X=\alpha)}\right)\right)$$

$$\sum_{\alpha \in I(X)} \sum_{\beta \in I(Y)} \big(\beta \times P(Y=\beta, X=\alpha)\big) = \sum_{\beta \in I(Y)} \left(\beta \times \sum_{\alpha \in I(X)} P(Y=y, X=\alpha)\right)$$

$$= \sum_{\beta \in I(Y)} \big(\beta \times P(Y=\beta)\big) = E(Y)$$

上述性质被称为 **全期望定理** (Law of total expectation)

这意味着可根据已求出的期望推出其它状态的期望

#2850、骰子的期望

题目描述

有一个 n 面指定质地均匀的骰子

请你求出抛出 n 个面所需要的期望次数

输入格式

第一行输入一个正整数 T ，表示 T 组询问

接下来输入 T 行,每行一个正整数 n

输出格式

每组询问输出一行一个浮点数(保留两位小数),表示期望的次数

输入样例

```
2
1
12
```

输出样例

```
1.00
37.24
```

设 dp_i 为已抛出 i 个面为了抛出所有面的期望次数

显然 $dp_n = 0$

各次抛骰子相互独立

数据规模

对于全部的数据 $1 \leq T \leq 200, 1 \leq n \leq 1000$

#2850、骰子的期望

对于第 i 次抛骰子

- 抛出未出现的面概率为 $\frac{n-i}{n}$, 由于抛出了新的面仅需要抛 $\frac{n-i}{n} \times (1 + \mathrm{dp}_{i+1})$ 次
- 抛出已出现的面概率为 $\frac{i}{n}$, 由于抛出了已出现的面 , 还需要抛 $\frac{i}{n} \times (1 + \mathrm{dp}_i)$ 次

加上抛出当前这次

$$\mathrm{dp}_i = (1 + \mathrm{dp}_i) \times \frac{i}{n} + (1 + \mathrm{dp}_{i+1}) \times \frac{n-i}{n}$$

整理得

$$\mathrm{dp}_i = \mathrm{dp}_{i+1} + \frac{n}{n-i}$$

答案为 $\mathrm{dp}[0]$, 需 $n \to 0$ 递推

时间复杂度 $O(n)$

#2849、绿豆蛙的归宿

题目描述

给出张 n 个点 m 条边的 DAG

起点为 1，终点为 n，每条边都有一个长度，并且从起点出发能够到达所有的点，所有的点也都能够到达终点

绿豆蛙从起点出发，走向终点

到达每一个顶点时，如果该节点有 k 条出边，绿豆蛙可以选择任意一条边离开该点，并且走向每条边的概率为 $\frac{1}{k}$

现在绿豆蛙想知道，从起点走到终点的所经过的路径总长度期望是多少？

输入格式

输入的第一行是两个整数，分别代表图的点数 n 和边数 m

第 $2 \sim m+1$ 行,每行有三个整数 u, v, w，代表存在一条 $u \to v$ 长度为 w 的有向边

输出格式

输出一行一个实数代表答案,保留两位小数

数据规模

对于 20% 的数据，保证 $n \le 10^2$

对于 40% 的数据，保证 $n \le 10^3$

对于 60% 的数据，保证 $n \le 10^4$

对于 100% 的数据，保证 $1 \le n \le 10^5$，$1 \le m \le 2 \times n$，$1 \le u, v \le n, 1 \le w \le 10^9$

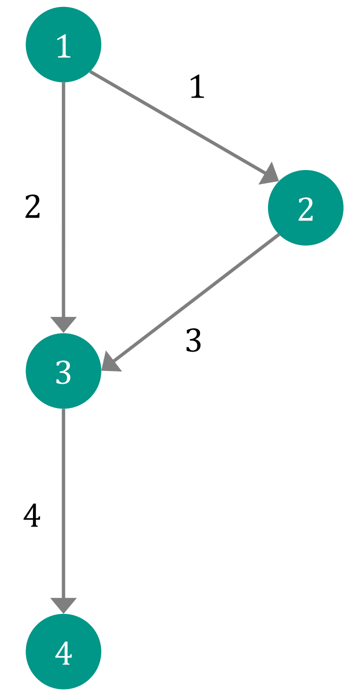
给出的图无重边和自环

输入样例

```
4 4
1 2 1
1 3 2
2 3 3
3 4 4
```

输出样例

```
7.00
```



从起点 1 到达各个点的概率为?

#2849、绿豆蛙的归宿

若将边反向建立反图 G，n 变为起点

设 $\mathrm{dp}[i]$ 为 i 到 n 的期望路径长度，初始时 $\mathrm{dp}_n = 0$

对于一个点 v 存在 k 条有向边，起点分别为 $u_1, u_2, \dots, u_k$，边长分别为 $w_1, w_2, \dots, w_k$ 那么

$$\mathrm{E}(v) = \frac{\mathrm{E}(u_1) + w_1}{k} + \frac{\mathrm{E}(u_2) + w_2}{k} + \cdots + \frac{\mathrm{E}(u_k) + w_k}{k}$$

状态转移方程

$$\mathrm{dp}[v] = \frac{1}{k} \times \sum_{i=1}^{k} (\mathrm{dp}[u_i] + w_i)$$

在反图 G 上做拓扑排序转移即可，答案为 $\mathrm{dp}[1]$

时间复杂度 $O(n+m)$

能否从 $1 \to n$ 计算?

#3333、游走

题目描述

给定一个 n 个点 m 条边的无向连通图,点编号 $1 \sim n$,边编号 $1 \sim m$

小 Z 在该图上进行随机游走,初始时小 Z 在 1 号顶点

每一步小 Z 以相等的概率随机选择当前顶点的某条边,沿着这条边走到下一个顶点,获得等于这条边的编号的分数

当小 Z 到达 n 号顶点时游走结束,总分为所有获得的分数之和

现在请你对这 m 条边进行编号,使得小 Z 获得的总分的期望值最小

输入格式

第一行输入两个整数,分别表示该图的顶点数 n, m

接下来 m 行每行两个整数 u, v ,表示 u, v 间存在一条边

输出格式

输出一行一个实数表示答案,保留三位小数

数据规模

对于 30% 的数据,保证 $n \leq 10$

对于 100% 的数据,保证 $2 \leq n \leq 500, 1 \leq m \leq 125000, 1 \leq u, v \leq n$

给出的图无重边和自环,且从 1 出发可以到达所有的节点

记 d_u 为点 u 的度，f_u 为点 u 期望经过的次数

g_i 为第 i 条边期望经过的次数

不难想到若 g_i 越大应尽可能分配较小的编号

对于第 i 条边 $u \leftrightarrow v$ 有

$$g_i = \frac{f_u}{d_u} + \frac{f_v}{d_v}$$

#3333、游走

考来求出 f_u，对于点 u 考虑与之相邻的边有

$$f_u = \sum_{\langle u,v\rangle \in \mathrm{E}} \frac{f_v}{d_v}$$

特殊的由于起点为 1

$$f_u = 1 + \sum_{\langle u,v\rangle \in \mathrm{E}} \frac{f_v}{d_v}$$

又由于终点为 n，所以其它点不能从 n 转移得到（不能考虑 $\frac{f_n}{d_n}$ 的贡献）

发现上述计算式存在环状依赖，无法递推求解

可将 f_u 计算式作为线性方程，那么可得到 $n-1$ 个线性方程组

高斯消元求解即可（本题对精度要求较高）

时间复杂度 $O(n^3 + m\log m)$

实验舱
青少年编程
LAB
走近科学 走进名校

谢谢观看