信奥算法

概率问题

probability

数学期望也称为均值
英文常用mean或expectation

概率加权求和

每次投掷硬币
正反面概率均为1/2
正面赢10元
反面输5元

每次期望赢2.5元

随机变量x
x=1,概率均为10%
x=2,概率均为40%
x=3,概率均为50%

x的期望值
E[X]=1*0.1+2*0.4+3*0.5
=2.4

盲盒1

请同学简述题意
突出核心要点

简化问题

m=2,只有两种可能性AB,目标是抽中B

B	1次抽中	概率1/2
AB	2次抽中	概率1/4
AAB	3次抽中	概率1/8
AAAB	4次抽中	概率1/16
AAAAB	5次抽中	概率1/32
AAAAAB	6次抽中	概率1/64
AAAAAAB	7次抽中	概率1/128

简化问题

m=2,只有两种可能性AB,目标是抽中B

B	1次抽中	概率1/2
AB	2次抽中	概率1/4
AAB	3次抽中	概率1/8

抽中次数x的平均值
也叫作期望(Expectation)，简写E

$$E[X] = \sum_{i=1,2,\ldots} i \times \Pr(X=i) = \sum_{i=1,2,\ldots} i \times \frac{1}{2^i}$$

简化问题

$1\times \Pr(x=1)$	$1\times \Pr(x=2)$	$1\times \Pr(x=3)$	$1\times \Pr(x=4)$	$1\times \Pr(x=5)$	…… ……
	$1\times \Pr(x=2)$	$1\times \Pr(x=3)$	$1\times \Pr(x=4)$	$1\times \Pr(x=5)$	…… ……
		$1\times \Pr(x=3)$	$1\times \Pr(x=4)$	$1\times \Pr(x=5)$	…… ……
			$1\times \Pr(x=4)$	$1\times \Pr(x=5)$	…… ……
				$1\times \Pr(x=5)$	…… ……
					…… ……
					…… ……

算两次	横竖各一次

$$E[X]=\sum_{i=1,2,\ldots}\Pr(X\geq i)=\sum_{i=1,2,\ldots}\frac{1}{2^{i-1}}=2$$

$$E[X]=\sum_{i=1,2,\ldots} i\times\Pr(X=i)=\sum_{i=1,2,\ldots} i\times\frac{1}{2^{i}}$$

简化问题

从简化问题(概率1/2)
拓展到原问题(概率1/m)
是否把2均改成m即可?

$$E[X] = \sum_{i=1,2,\ldots} \Pr(X \geq i) = \sum_{i=1,2,\ldots} \frac{1}{2^{i-1}} = 2$$

$$E[X] = \sum_{i=1,2,\ldots} i \times \Pr(X = i) = \sum_{i=1,2,\ldots} i \times \frac{1}{2^{i}}$$

原问题

$1\times\Pr(x=1)$	$1\times\Pr(x=2)$	$1\times\Pr(x=3)$	$1\times\Pr(x=4)$	$1\times\Pr(x=5)$	$\cdots\ \cdots$
	$1\times\Pr(x=2)$	$1\times\Pr(x=3)$	$1\times\Pr(x=4)$	$1\times\Pr(x=5)$	$\cdots\ \cdots$
		$1\times\Pr(x=3)$	$1\times\Pr(x=4)$	$1\times\Pr(x=5)$	$\cdots\ \cdots$
			$1\times\Pr(x=4)$	$1\times\Pr(x=5)$	$\cdots\ \cdots$
				$1\times\Pr(x=5)$	$\cdots\ \cdots$
					$\cdots\ \cdots$
					$\cdots\ \cdots$

算两次 横竖各一次

$$E[X]=\sum_{i=1,2,\ldots}\Pr(X\geq i)=\sum_{i=1,2,\ldots}\left(\frac{m-1}{m}\right)^{i-1}=m$$

$$E[X]=\sum_{i=1,2,\ldots}i\times\Pr(X=i)=\sum_{i=1,2,\ldots}i\times\left(\frac{m-1}{m}\right)^{i-1}\times\frac{1}{m}$$

小结：几何分布

抽中次数x是随机变量
服从几何分布
Geometric distribution

Pr(X=k)定义为在n次01试验中
前k-1次皆失败0，第k次成功1的概率

已知单次试验成功1的概率为p
则 x的期望为1/p

小结：过关概率

$$E[X] = \sum_{i=1,2,\ldots} i \times \Pr(X = i)$$

期望定义公式

$$E[X] = \sum_{i=1,2,\ldots} \Pr(X \geq i)$$

尾部概率求和公式

tail integration

易于计算

盲盒2

请同学简述题意
突出核心要点

这是经典的"卡牌收集"问题
coupon collector's problem

思维框架

m=2,两种可能AB	简化问题
先抽一次 必然收集到一种	走一步看看 one-step analysis
目标改为收集另一种	转化经典问题
m=3,4,5,...	推广结论

路径分步走
一步一期望

收集齐m种卡牌，分为m步

收到第1种新卡	每次成功概率1
收到第2种新卡	每次成功概率(m-1)/m
收到第3种新卡	每次成功概率(m-2)/m
收到第4种新卡	每次成功概率(m-3)/m
······	······
收到第m种新卡	每次成功概率1/m

$$E[X] = E\left[\sum_{i=1,2,\ldots,m} X_i\right] = \sum_{i=1,2,\ldots,m} E[X_i]$$

X_i代表从拿到(i-1)种卡到拿到i种卡的次数

收到第1种新卡　　每次成功概率1

收到第2种新卡　　每次成功概率(m-1)/m

收到第3种新卡　　每次成功概率(m-2)/m

收到第4种新卡　　每次成功概率(m-3)/m

……　　……

收到第m种新卡　　每次成功概率1/m

$$E[X] = E\left[\sum_{i=1,2,\ldots,m} X_i\right] = \sum_{i=1,2,\ldots,m} E[X_i]$$

X_i代表从拿到(i-1)种卡到拿到i种卡的次数

$$= \frac{m}{m} + \frac{m}{m-1} + \frac{m}{m-2} + \cdots + \frac{m}{1}$$
$$= m \times \left(\frac{1}{m} + \frac{1}{m-1} + \frac{1}{m-2} + \cdots + \frac{1}{1}\right)$$

当m很大时

$E[X]$ 趋向于 mlog(m)

小结：分步走

走一步看看

one-step analysis

路径分步走

一步一期望

$$E[A+B]=E[A]+E[B]$$

A和B是任意两个随机变量
甚至A和B 可以相关

加法分拆

但$E[AB]=E[A]E[B]$的成立需要AB不相关

50%概率： A=1且B=10
50%概率： A=2且B=20

50%概率： A=1且B=20
50%概率： A=2且B=10

$$E[A+B]=E[A]+E[B]$$

A和B是任意两个随机变量
甚至A和B 可以相关

刷题进度

共有n道题要刷，每天你有80%可能性一题都做不出来，还有20%可能性正好做出一个新题。求完成这n题的期望天数。

利用期望的加法分拆

融会贯通

期望的加法分拆

$$E[A+B]=E[A]+E[B]$$

尾部概率求和公式

tail integration formula

$$X=I_{(X\geq 1)}+I_{(X\geq 2)}+I_{(X\geq 3)}+\cdots$$

$$E[X]=E\left[\sum_{i=1,2,\ldots} I_{(X\geq i)}\right]=\sum_{i=1,2,\ldots}\Pr(X\geq i)$$

思考题：连中问题

卡牌收集问题里，共m种卡牌，目标：
连续k次抽卡结果都一样
求抽卡次数期望

随机搜索二叉树

有一颗二叉树共n个节点，编号1到n，每条边长度为1。根节点是1号，目标节点是n号。现在从1号节点出发开始DFS, 到n号节点停止, 求DFS算法经过路径长度的期望。
注意：回溯步骤也算路径长度。

输入第一行为正整数n，n<=100000。接下去n-1行，每行两个正整数代表树上某条边的两个端点。

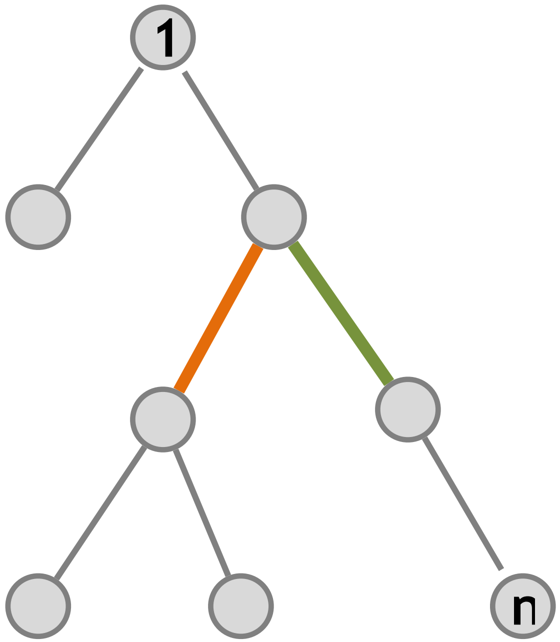
输出一个浮点数，保留2位小数。

输入样例:
3
1 2
1 3

输出样例:
2.00

请用纸和笔
写出算法

随机搜索二叉树



绿色边在每次DFS中恰访问一次

橙色边在所有DFS中访问概率为1/2

若访问了橙色边一定来回访问2次

逆序对

{1,2,3,...,n}这n个数做随机排列，在一个排列中，逆序对个数的期望是多少?

请用纸和笔
写出算法

每对数字，有1/2概率形成逆序对

小结

两种统计角度

每一个排列
内部统计

每一个对象
跨排列整体统计

均匀撒点

{1,2,3,...,n}这n个数字里等概率随机取一个A,
{1,2,3,...,n}这n个数字里等概率随机取一个B,
求E[|A-B|]

E[|A-B|] = E[max(A,B) - min(A,B)]
= E[max(A,B)] - E[min(A,B)]

请用纸和笔写出算法

让X代表min(A,B)

两种方法

$$E[X] = \sum_{i=1,2,\ldots} i \times \Pr(X = i) \quad = \sum_{i=1,2,\ldots,n} i \times \Pr(\text{最小值} = i)$$

$$E[X] = \sum_{i=1,2,\ldots} \Pr(X \geq i) \quad = \sum_{i=1,2,\ldots,n} \Pr(\text{最小值} \geq i)$$

均匀撒点

$\Pr(\text{最小值} = i)$

$$\frac{1}{n} \times \frac{n-i+1}{n} \times 2$$

✗

$$\frac{1}{n} \times \frac{1}{n} + \frac{1}{n} \times \frac{n-i}{n} \times 2$$

✓

$\Pr(\text{最小值} \geq i)$

$$\left(\frac{n-i+1}{n}\right)^2$$

$$E[X] = \sum_{i=1,2,\ldots,n} \Pr(X \geq i) = \sum_{i=1,2,\ldots,n} \left(\frac{n-i+1}{n}\right)^2$$

$$= \frac{1}{n^2} \sum_{i=1,2,\ldots,n} (n-i+1)^2 = \frac{1}{n^2} \sum_{i=1,2,\ldots,n} i^2 = \frac{(n+1)(2n+1)}{6n} \approx \frac{n}{3}$$

均匀撒点

[0,1]区间内取n个i.i.d.均匀分布的随机变量，
求最大值的期望，最小值的期望。

$\Pr(\text{最小值} \geq x)$		$(1-x)^n$
$\Pr(\text{最大值} \geq x)$	$1-\Pr(\text{最大值} < x)$	$1-x^n$

$$E[\text{最大值}] = \int_{x=0}^{1} \Pr(\text{最大值} \geq x)\, dx = 1 - \int_{x=0}^{1} x^n dx$$

$$= 1 - \frac{1}{n+1} x^{n+1} \Big|_0^1 = \frac{n}{n+1}$$

$$E[\text{最小值}] = \frac{1}{n+1}$$

随机过程

stochastic processes

皇家赌场1

请同学简述题意
突出核心要点

动态规划1 期望

定义状态

自然定义法,也叫抄原题大法

题目问什么,状态含义就是什么

问题答案

f[t][0]代表剩t次要抛,当前0个正面,
最优策略,最终正面的期望个数

f[i][j]代表剩i次要抛,当前j个正面,
最优策略,最终正面的期望个数

动态规划1 期望

f[i][j]代表剩i次要抛,当前j个正面,
最优策略,最终正面的期望个数

手算表格

输入	2 3
2个硬币抛3次	

输出	1.25

f[i][j]	j=0	j=1	j=2
i=0			
i=1			
i=2			
i=3			

请总结转移方程

动态规划1 期望

f[i][j]代表剩i次要抛,当前j个正面,
最优策略,最终正面的期望个数

i=0

$f[0][j]=j$

j=n

$f[i][n]=f[i-1][n]/2+f[i-1][n-1]/2$

i=1,2,..,t
j=0,1,..,n-1

$f[i][j]=f[i-1][j]/2+f[i-1][j+1]/2$

代码1

```
cin>>n>>t;
for(int j=0;j<=n;j++)f[0][j]=j;
for(int i=1;i<=t;i++){
    for(int j=0;j<n;j++)
        f[i][j]=(f[i-1][j]+f[i-1][j+1])/2;
    f[i][n]=(f[i-1][n]+f[i-1][n-1])/2;
}
cout<<fixed<<setprecision(2)<<f[t][0]<<endl;
```

动态规划2 概率

定义状态

具体化状态,比原问题更细致

原问题需要哪些要素,就补充计算

p[i][j]代表已经最优策略抛i次,
当前j个正面的概率

问题答案

$$\Pr(\text{最小值} \geq i)$$

动态规划2 概率

p[i][j]代表已经最优策略抛i次，
当前j个正面的概率

手算
表格

输入	2 3
2个硬币抛3次	

输出	1.25

p[i][j]	j=0	j=1	j=2
i=0			
i=1			
i=2			
i=3			

请总结转移方程

动态规划2 概率

p[i][j]代表已经最优策略抛i次，
当前j个正面的概率

i=0,j=0	p[0][0]=1
i=0,j>=1	p[0][j]=0
i>=1,j=0	p[i][0]=p[i-1][0]/2
1<=i<=t 1<=j<=n	p[i][j]=p[i-1][j]/2+p[i-1][j-1]/2
i>=1,j=n-1	p[i][n-1]**+=**p[i-1][n]/2

代码2

```
cin>>n>>t;
p[0][0]=1;
for(int i=1;i<=t;i++){
    p[i][0]=p[i-1][0]/2;
    for(int j=1;j<=n;j++)
        p[i][j]=p[i-1][j]/2+p[i-1][j-1]/2;
    p[i][n-1]+=p[i-1][n]/2;
}
long double ans=0;
for(int j=1;j<=n;j++)ans+=j*p[t][j];
cout<<fixed<<setprecision(2)<<ans<<endl;
```

DP小结

状态定义	基于期望
	基于概率

转移方程	走一步看看
	one-step analysis
	当前状态依赖走一步之后的状态

期望DP解盲盒2

请复述题意

期望DP

如何用DP求解

定义状态

自然定义法，也叫抄原题大法

题目问什么，状态含义就是什么

问题答案

f[m]代表剩m种卡牌要收集
收集齐m种所需次数的期望

f[i]代表剩i种卡牌要收集
收集齐m种所需次数的期望

期望DP

f[i]代表剩i种卡牌要收集
收集齐m种所需次数的期望

i=0	f[0]=0
i=1,2,..n	f[i]=i/m*f[i-1]+(m-i)/m*f[i] **+1**
解方程	m*f[i]=i*f[i-1]+(m-i)*f[i]+m
	f[i]=f[i-1]+m/i

拓展任务

{1,2,3,...,n}这n个数做随机排列,
在一个排列中，LIS的长度期望是多少?

请同学课后思考，提出自己的猜想

快快编程作业

1514
1515
1529
1513

思考题	LIS期望