数学活动课程讲座

概率方法在数学竞赛中的运用

武炳杰

（复旦大学数学科学学院 2013 级直博，200433）

中图分类号：O211 文献标识码：A 文章编号：1005－6416(2015)04－0002－07

（本讲适合高中）

在中学课本中，同学们接触了一些简单的有关古典概率的概念. 本文通过概率模型及一些竞赛题阐述概率方法在代数、数论及组合问题中的运用与解题技巧.[1]

设 $P(A)$ 表示事件 A 发生的概率.

对一个取值范围是有限集 S 的随机变量 X，定义 X 的数学期望为

$$E(X)=\sum_{i\in S} iP(X=i).$$

有以下主要结论：

(1) 若基本事件全体所构成的样本空间 Ω 可划分为 $\{A_1,A_2,\cdots,A_n\}$，则 $A_1,A_2,\cdots,A_n$ 互不相交且 $P(\bigcup_{i=1}^{n}A_i)=1$.

(2) 在概率空间中的随机变量 X，总存在一些事件，使得 $X\geqslant E(X)$，也存在一些事件，使得 $X\leqslant E(X)$.

利用数学期望解题，应注意其可加性，这也给了同学们分解为示性函数、求和提供了可行性.

可加性 对于任意两个随机变量 X、Y（注意未必独立），总有

$$E(X+Y)=E(X)+E(Y).$$

证明

$$\begin{aligned}
E(X+Y)&=\sum_{x,y}(x+y)P(X=x,Y=y)\\
&=\sum_{x,y}xP(X=x,Y=y)+\sum_{x,y}yP(X=x,Y=y)\\
&=\sum_{x}\Big(x\sum_{y}P(X=x,Y=y)\Big)+\sum_{y}\Big(y\sum_{x}P(X=x,Y=y)\Big)\\
&=\sum_{x}xP(X=x)+\sum_{y}yP(Y=y)\\
&=E(x)+E(y).
\end{aligned}$$

1 运用概率模型

在遍历所有项的和式中，有时能套用概率模型.

例 1 设 S 为平面上的有限点集，任三点不共线，对于顶点属于集合 S 的每一个凸多边形 P，设 P 的顶点数为 $a(P)$，属于集合 S 且在多边形 P 外部的点的个数为 $b(P)$. 证明：对于任意实数 $x(0<x<1)$，均有

$$\sum_{P}x^{a(P)}(1-x)^{b(P)}=1,$$

其中，P 遍历了集合 S 中的所有凸多边形（包括三角形，且一条线段、一个点和空集也分别认为是凸二边形，凸一边形和凸零边形）.[2]

（第 47 届 IMO 预选题）

证明 将集合 S 中的点随机染色，每个点染黑的概率为 x，染白的概率为 $1-x$.

对于任意凸多边形 P，令事件 E_P 表示随机染色后多边形 P 边界上的点均染黑、外部的点均染白的事件. 显然，对任何凸多边形 P、P'，事件 E_P、$E_{P'}$ 互不包含，故 $\sum_{P}x^{a(P)}(1-x)^{b(P)}$ 表示 $P(\bigcup_{P}E_P)$. 而对于任何一种染法，可取凸多边形 P 为所有黑点的凸包，故无论怎样染色，总有一

收稿日期：2014－10－31 修回日期：2014－12－26

个事件 E_P 会发生. 于是,$P(\bigcup_P E_P)=1$.

从而,原命题得证.

例 2 设 $n\in\mathbf{Z}_+$,实数 $a_1,a_2,\cdots,a_n\in(0,1)$,对于任意子集 $S\subseteq\{1,2,\cdots,n\}$,定义集函数

$$f(S)=\left(\prod_{i\in S}a_i\right)\left[\prod_{j\notin S}(1-a_j)\right].$$

若 $\sum\limits_{|S|\text{为奇数}}f(S)=\frac{1}{2}$,证明:存在某些 k,使得 $a_k=\frac{1}{2}$($\sum\limits_{|S|\text{为奇数}}f(S)$ 表示子集 S 遍历了 $\{1,2,\cdots,n\}$ 的所有有奇数个元素的子集).

(2014,美国国家集训队训练题)

证明 令 $X=\sum\limits_{|S|\text{为奇数}}f(S)$.

考虑 n 枚硬币,每枚正面朝上的概率分别为 $a_1,a_2,\cdots,a_n$. 则 X 表示有奇数枚硬币正面朝上的概率.

显然,若存在一个 $a_k=\frac{1}{2}$,则 $X=\frac{1}{2}$. 从而,$X-\frac{1}{2}$ 可被 $\prod\limits_{i=1}^{n}\left(a_i-\frac{1}{2}\right)$ 整除.

比较两者的次数,均为 n,故两者只相差一个首项系数.

于是,结论成立.

2 运用数学期望的分解

运用数学期望解组合问题的步骤是先将问题看作随机事件,把所求量视为一个随机变量,并将其分解为若干个示性随机变量,再分别进行概率计算,最终累和即可.

例 3 在 $n\times n$ 的方格表中,$1,2,\cdots,n$ 各出现 n 次. 证明:一定存在一行或一列,至少有 $[\sqrt{n}]$ 个不同的数,其中,$[x]$ 表示不超过实数 x 的最大整数.

(2004,美国数学夏令营)

证明 随机地选出一行或一列,这样一共有 $2n$ 种选法. 令随机变量 X 表示该行或列中含有不同数的个数.

记 I_i 为示性随机变量,即若 i 出现在该行,则 $I_i=1$,否则,$I_i=0$.

故 $X=\sum\limits_{i=1}^{n}I_i$.

从而,$E(X)=\sum\limits_{i=1}^{n}E(I_i)$.

注意到,

$$E(I_i)=0\times P(I_i=0)+1\times P(I_i=1)$$
$$=P(I_i=1).$$

为了计算其下界,考虑最坏状态,即 n 个 i 恰分布在 $[\sqrt{n}]\times[\sqrt{n}]$ 的方格表中. 则 n 个 i 至少要占据 $2[\sqrt{n}]$ 个不同的行或列.

而任意一行或一列的选法有 $2n$ 种,故

$$P(I_i=1)\geqslant\frac{2[\sqrt{n}]}{2n}.$$

进一步,$E(X)=nP(I_i)\geqslant[\sqrt{n}]$.

因此,总有一行或一列,其包含至少 $[\sqrt{n}]$ 个不同的数.

例 4 设 $S=\{0,1,\cdots,N^2-1\}$,A 为集合 S 的一个 N 元子集. 证明:存在集合 S 的一个 N 元子集 B,使得集合

$$A+B=\{a+b\mid a\in A,b\in B\}$$

中的元素模 N^2 的余数的数目不少于集合 S 元素个数的一半.

(第 40 届 IMO 预选题)

证明 将等概率地、独立地从 N^2 个剩余系的元素中取 N 个数的集合记为 B(由于是独立地选,则每次每个数被取出的概率为 $\frac{1}{N^2}$,虽可能有元素会重复地选出,但若能证明存在一个不超过 N 个元素的子集符合题意,则题目中要求的 N 元子集 B 亦自然存在). 用随机变量 X 表示集合 $A+B$ 中的不同元素个数.

对任意一元素 i,恰有 N 个不同的元素 b,使得 $i\in A+b$,故 i 不出现在集合 $A+B$ 中的概率为 $\left(1-\frac{N}{N^2}\right)^N$. 从而,$i$ 出现在集合 $A+B$

中的概率为 $1-\left(1-\frac{N}{N^2}\right)^N$.

类似于例3,用 I_i 表示示性随机变量,即若 i 出现在集合 $A+B$ 中,则 $I_i=1$,否则,$I_i=0$.

故 $E(X)=\sum\limits_{i=1}^{N^2}E(I_i)=N^2\left[1-\left(1-\frac{1}{N}\right)^N\right]$.

由 $\left(1-\frac{1}{N}\right)^{-N}=\left(\frac{N}{N-1}\right)^N=\left(1+\frac{1}{N-1}\right)^N>\mathrm{e}$,则 $\left(1-\frac{1}{N}\right)^N<\frac{1}{\mathrm{e}}<\frac{1}{2.7}$.

代入后得

$$E(X)>N^2\left(1-\frac{1}{2.7}\right)>\frac{N^2}{2}.$$

因此,存在一个符合题意的集合 B,使得 X 的值大于 $\frac{N^2}{2}$.

3 结合 C_x^k 的凸性

下面的例子需要用到 C_x^k 的凸性.

如当 $k=2$ 时,若 $\sum\limits_{i=1}^{n}x_i=S(x_i\in\mathbf{Z}_+)$,则

$$\sum_{i=1}^{n}\mathrm{C}_{x_i}^2=\frac{\sum\limits_{i=1}^{n}x_i^2-\sum\limits_{i=1}^{n}x_i}{2}$$

$$\geqslant\frac{\frac{\left(\sum\limits_{i=1}^{n}x_i\right)^2}{n}-\sum\limits_{i=1}^{n}x_i}{2}=\frac{S^2-Sn}{2n}=n\mathrm{C}_{\frac{S}{n}}^2.$$

例5 杜马中有1 600位议员,他们组成了16 000个委员会,每个委员会有80名成员.证明:可以找出两个委员会,两者中至少有4位公共成员.

(1995—1996,全俄数学奥林匹克)

证明 随机地选出一对委员会,这样一共有 $\mathrm{C}_{16\,000}^2$ 种可能,用随机变量 X 表示这对委员会中相同委员的人数.记 I_i 为示性随机变量,即若 i 为这对委员会的公共委员时,$I_i=1$,否则,$I_i=0$.

故 $X=\sum\limits_{i=1}^{1\,600}I_i$.

从而,由随机变量的可加性知

$$E(X)=\sum_{i=1}^{1\,600}E(I_i).$$

经计算得

$$E(I_i)=P(I_i=1)=\frac{\mathrm{C}_{n_i}^2}{\mathrm{C}_{16\,000}^2},$$

其中,n_i 表示委员 i 参加的委员会个数.

注意到,$\sum\limits_{i=1}^{1\,600}n_i=16\,000\times80$.

由 $\mathrm{C}_{n_i}^2$ 的凸性知

$$E(X)=\sum_{i=1}^{1\,600}\mathrm{C}_{n_i}^2\geqslant1\,600\cdot\frac{\mathrm{C}_{\frac{16\,000\times80}{1\,600}}^2}{\mathrm{C}_{16\,000}^2}=3.995.$$

由于 X 为整数,故一定有一对委员会,其相同人数大于或等于4.

4 运用到一些代数不等式与组合结合的问题

例6是一道关于平衡向量的题.

例6 n 个 n 维向量 $\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_n\in\mathbf{R}^n$,每一个的长度为1,即

$$|\boldsymbol{v}_i|=\sqrt{v_{1i}^2+v_{2i}^2+\cdots+v_{ni}^2}=1.$$

证明:存在 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n\in\{1,-1\}$,使得

$$|\varepsilon_1\boldsymbol{v}_1+\varepsilon_2\boldsymbol{v}_2+\cdots+\varepsilon_n\boldsymbol{v}_n|\leqslant\sqrt{n}.$$

证明 独立地、等概率地从 $\{1,-1\}$ 中取出 n 个数 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$,记随机变量

$$X=|\varepsilon_1\boldsymbol{v}_1+\varepsilon_2\boldsymbol{v}_2+\cdots+\varepsilon_n\boldsymbol{v}_n|^2.$$

展开后得

$$\begin{aligned}X&=(\varepsilon_1\boldsymbol{v}_1+\varepsilon_2\boldsymbol{v}_2+\cdots+\varepsilon_n\boldsymbol{v}_n)\cdot\\&\quad(\varepsilon_1\boldsymbol{v}_1+\varepsilon_2\boldsymbol{v}_2+\cdots+\varepsilon_n\boldsymbol{v}_n)^{\mathrm{T}}\\&=(\varepsilon_1\boldsymbol{v}_1+\varepsilon_2\boldsymbol{v}_2+\cdots+\varepsilon_n\boldsymbol{v}_n)\cdot\\&\quad(\varepsilon_1(\boldsymbol{v}_1)^{\mathrm{T}}+\varepsilon_2(\boldsymbol{v}_2)^{\mathrm{T}}+\cdots+\varepsilon_n(\boldsymbol{v}_n)^{\mathrm{T}})\\&=\sum_{i=1}^{n}\sum_{j=1}^{n}\varepsilon_i\varepsilon_j\boldsymbol{v}_i\cdot(\boldsymbol{v}_j)^{\mathrm{T}},\end{aligned}$$

其中,$(\boldsymbol{v})^{\mathrm{T}}$ 代表 $\boldsymbol{v}$ 的转置.

两边取数学期望,由线性可加性知

$$E(X)=\sum_{i=1}^{n}\sum_{j=1}^{n}E(\varepsilon_i\varepsilon_j)\boldsymbol{v}_i\cdot(\boldsymbol{v}_j)^{\mathrm{T}}.$$

当 $i\neq j$ 时,$E(\varepsilon_i\varepsilon_j)=E(\varepsilon_i)E(\varepsilon_j)=0$;

当 $i=j$ 时,$E(\varepsilon_i^2)=1$.

于是，$E(X)=\sum\limits_{i=1}^{n}\boldsymbol{v}_i\cdot(\boldsymbol{v}_i)^{\mathrm{T}}=n$.

故一定存在至少一组 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n\in\{1,-1\}$，使得

$$|\varepsilon_1\boldsymbol{v}_1+\varepsilon_2\boldsymbol{v}_2+\cdots+\varepsilon_n\boldsymbol{v}_n|\leqslant\sqrt{n}\ .$$

例 7 对于整数 $n(n\geqslant2)$，设实数 $x_1,x_2,\cdots,x_n$ 满足

$$x_1+x_2+\cdots+x_n=0,$$
$$x_1^2+x_2^2+\cdots+x_n^2=1.$$

对于任意的集合 $A\subseteq\{1,2,\cdots,n\}$，定义

$S_A=\sum\limits_{i\in A}x_i$（若 A 为空集，则 $S_A=0$）.

证明：对于任意的正实数 λ，满足 $S_A\geqslant\lambda$ 的集合 A 的个数最多为 $\dfrac{2^{n-3}}{\lambda^2}$，并确定使等号成立的一切 $x_1,x_2,\cdots,x_n$.[3]

（2012，美国数学奥林匹克）

证明 设集合 $\mathscr{F}$ 为 $\{1,2,\cdots,n\}$ 的所有子集构成的子集族.

显然，$|\mathscr{F}|=2^n$.

取集合 A 的示性函数 I_A，即若 $i\in A$，则 $I_A(i)=1$，否则，$I_A(i)=0$.

由题意知 $S_A=\sum\limits_{i=1}^{n}I_A(i)x_i$.

计算知

$$\sum_{A\in\mathscr{F}}(S_A)^2=\sum_{A\in\mathscr{F}}\left(\sum_{i=1}^{n}I_A(i)x_i\right)\left(\sum_{j=1}^{n}I_A(j)x_j\right)$$
$$=\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{A\in\mathscr{F}}(I_A(i)I_A(j))x_ix_j.$$

当 $i\neq j$ 时，有 2^{n-2} 个集合 A 同时包含 i、j；

当 $i=j$ 时，有 2^{n-1} 个集合 A 同时包含 i、j.

故 $\sum\limits_{A\in\mathscr{F}}(S_A)^2$

$$=2^{n-2}\left(2\sum_{i=1}^{n}x_i^2+\sum_{\substack{i,j\in\{1,2,\cdots,n\}\\ i\neq j}}x_ix_j\right)$$
$$=2^{n-2}\left(\sum_{i=1}^{n}x_i^2+\sum_{i=1}^{n}\sum_{j=1}^{n}x_ix_j\right)$$
$$=2^{n-2}\left[\sum_{i=1}^{n}x_i^2+\left(\sum_{i=1}^{n}x_i\right)^2\right]$$
$$=2^{n-2}(1-0)=2^{n-2}.$$

事实上，接下来的证明完全就是概率论中的切比雪夫不等式.

记 $B_\lambda=\{A\in\mathscr{F}\mid S_A\geqslant\lambda\}$，

$B'_\lambda=\{A\in\mathscr{F}\mid S_A\leqslant-\lambda\}$.

因为 $x_1+x_2+\cdots+x_n=0$，所以，集合 B_λ、B'_λ 可通过取补集的运算构成双射，即

$$|B_\lambda|=|B'_\lambda|.$$

构造集合函数

$$g(A)=\begin{cases}\lambda^2, & A\in B_\lambda\cup B'_\lambda;\\ 0, & A\notin B_\lambda\cup B'_\lambda.\end{cases}$$

于是，$S_A^2\geqslant g(A)$.

故 $\lambda^2|B_\lambda\cup B'_\lambda|=\sum\limits_{A\in\mathscr{F}}g(A)$

$$\leqslant\sum_{A\in\mathscr{F}}S_A^2=2^{n-2}.$$

从而，$|B_\lambda|\leqslant\dfrac{2^{n-3}}{\lambda^2}$.

这便完成了结论中不等式的证明.

对于等号成立，由最后一步不等式，知应该对任意的集合 A 均有 $S_A^2=g(A)$，这只可能在 $(x_1,x_2,\cdots,x_n)=(\lambda,-\lambda,0,\cdots,0)$ 或其一个任意排列时才会发生. 由平方和为 1 的式子，知正实数 $\lambda=\dfrac{\sqrt{2}}{2}$.

5 在图论问题中的运用

将图论中的一些概念和性质转化为概率表述，结合数学期望的线性可加性，便可得到计算式.

例 8 某社团中，每两人不是友好的就是敌对的. 设该社团有 n 个人及 q 个友好对，且任三人中至少有一对人是敌对的. 证明：该社团中至少有一名成员，其敌人所组成的集合中，友好对的数且不多于 $q\left(1-\dfrac{4q}{n^2}\right)$.

（1995，美国数学奥林匹克）

证明 随机取一人 v，并用随机变量 X 表示与此人敌对的人中的友好对数目.

利用算两次得

$$E(X)=\sum_{\text{友好对}e}E(I_e)$$

$$=\sum_{\text{友好对}e}\frac{1}{n}|\{v|\text{边}\ e\ \text{的两端均与}\ v\ \text{敌对}\}|.$$

因为没有三人两两友好,所以,由上式可继续推得

$$E(X)=\sum_{\text{友好对}e}\frac{1}{n}[(n-2)-|\{v|\text{边}\ e\ \text{恰有一端与}\ v\ \text{友好}\}|]$$

$$=q\cdot\frac{n-2}{n}-\frac{1}{n}\sum_{\text{友好对}e}|\{v|\text{边}\ e\ \text{恰有一端与}\ v\ \text{友好}\}|.$$

用 d_v 表示与 v 友好的人数,K 表示社团中恰有两个友好对的三人组总数.

由前述的 C_x^2 的凸性知

$$K=\sum_v C_{d_v}^2\geqslant nC_{\frac{2q}{n}}^2.$$

由于每个这样的三人组有两条边,从而,被算了两次.

故 $E(X)=q\cdot\frac{n-2}{n}-\frac{1}{n}\cdot2K$

$$\leqslant q\cdot\frac{n-2}{n}-\frac{1}{n}\cdot2nC_{\frac{2q}{n}}^2=q\left(1-\frac{4q}{n^2}\right).$$

例 9 在一个有 n 个顶点的图 G 中,若不包含 p 阶完全图 K_p,则图 G 的边数

$$|E|\leqslant\left(1-\frac{1}{p-1}\right)\frac{n^2}{2},$$

其中,p 阶完全图 K_p 指的是有 p 个顶点且这 p 个顶点两两间均有边的图. [4]

【注】本题即为托兰(Turan)定理.

证明 记图 G 的顶点为 $v_1,v_2,\cdots,v_n$,对应的度数分别为 $d_1,d_2,\cdots,d_n$. 用 $\omega(G)$ 表示图 G 中最大完全图的顶点数,用排列τ表示图 G 中 n 个顶点的一个随机排列,显然有$\frac{1}{n!}$种.

对于排列τ,用集合 C_τ表示图 G 中的某些顶点,这些顶点在排列τ下,均与其前面的顶点相邻. 设随机变量 X 表示 C_τ的顶点个数. 事实上,C_τ的顶点构成了图 G 的一个完全图.

再用 X_i 为示性随机变量,即若 $v_i\in C_\tau$时,$X_i=1$,否则,$X_i=0$. 于是,$X=\sum\limits_{i=1}^{n}X_i$.

注意到,$v_i\in C_\tau$当且仅当所有 $n-1-d_i$ 个与 v_i 不相邻的顶点在它之后,而该事件的概率恰等于 v_i 在与这 $n-1-d_i$ 个顶点的排列中总排在第 1 个,即$\frac{1}{n-d_i}$,所以,

$$E(X_i)=\frac{1}{n-d_i}.$$

故 $E(|C_\tau|)=E(X)$

$$=\sum_{i=1}^{n}E(X_i)=\sum_{i=1}^{n}\frac{1}{n-d_i}.$$

又最大完全图一定比取出是完全图的随机变量的数学期望要大,结合柯西不等式知

$$|\omega(G)|\geqslant E(|C_\tau|)=\sum_{i=1}^{n}\frac{1}{n-d_i}$$

$$\geqslant\frac{n^2}{\sum\limits_{i=1}^{n}(n-d_i)}=\frac{n^2}{n^2-2|E|}.$$

由题意得$|\omega(G)|\leqslant p-1$.

代入后化简得

$$|E|\leqslant\left(1-\frac{1}{p-1}\right)\frac{n^2}{2}.$$

6 对第 55 届 IMO 第 6 题的探索

以下的概率证法是基于美国队领队 Po-Shen Loh 在 2014 年 IMO 上提出的.

例 10 若平面上的一簇直线中没有两条直线平行,没有三条直线共点,则称其为“处于一般位置”. 一簇处于一般位置的直线将平面分割成若干区域,其中面积有限的区域称为这簇直线的“有限区域”. 证明:对于充分大的 n 及任意处于一般位置的 n 条直线,均可将其中至少$\frac{2\sqrt{n}}{3}$条直线染成蓝色,使得每一个有限区域的边界均不全为蓝色的.

(改编自第 55 届 IMO 第 6 题)

证明 首先,计数 n 条一般位置的直线至多能分割出多少三角形和多边形(三条边以上的图形). 由于一共有 C_n^2 个交点,每个交点至多为两个三角形的顶点,而每个三角形共有三个顶点,于是,三角形的个数至多为 $\frac{2}{3}C_n^2$,即不超过 $\frac{1}{3}n^2$(这样的放缩是因为题目要求估计充分大的 n,这表明,阶数较小的量是可以忽略的,下同).

对于多边形,由欧拉定理

$$V-E+F=2$$

(V、E、F 分别代表顶点个数、边数、面数,平面上为划分出来的小区域加一块涉及无穷远点的面,至于如何从立体图形转化到平面图形,可由欧拉球极投影来想像),而

$$E=(n-2)n, V=C_n^2,$$

则 $F=\frac{n^2-3n}{2}+2$,即划分出的有限图形不超过 $\frac{n^2}{2}$ 块(其中还包括了三角形).

假设有概率为 p 的可能将每条直线染为蓝色. 则

$$E(\text{蓝色直线})=np,$$

$$E(\text{蓝色三角形})<\frac{1}{3}n^2p^3,$$

$$E(\text{蓝色有限区域})<\frac{1}{2}n^2p^4.$$

更进一步地放缩知

$$E(\text{蓝色有限区域(三边以上图形)})$$

$$<\frac{1}{2}n^2p^4.$$

接下来考虑问题:对于每一块蓝色有限区域(无论是三角形还是多边形),只需对其任一条蓝边染回原色即可使其不是全蓝,从而,使得没有全蓝区域的蓝色直线数的数学期望至少为

$$np-\left(\frac{1}{3}n^2p^3\right)-\left(\frac{1}{2}n^2p^4\right)=np\left(1-\frac{np^2}{3}-\frac{np^3}{2}\right).$$

因为题目要求估计的对象是充分大的 n,所以,可忽略上式中较小项 $\frac{np^3}{2}$.

设 $p=\frac{k}{\sqrt{n}}$.

代入后得 $\sqrt{n}$ 的系数为 $k\left(1-\frac{k^2}{3}\right)$.

由均值不等式知

$$\left[k\left(1-\frac{k^2}{3}\right)\right]^2$$

$$\leqslant\frac{3}{2}\left[\frac{\frac{2}{3}k^2+\left(1-\frac{k^2}{3}\right)+\left(1-\frac{k^2}{3}\right)}{3}\right]^3=\frac{4}{9}.$$

当 $\sqrt{n}$ 的系数取最大值 $\frac{2}{3}$ 时,$k=1$.

代入 $np\left(1-\frac{np^2}{3}-\frac{np^3}{2}\right)$ 后知至少能保留住 $\frac{2\sqrt{n}}{3}-\frac{1}{2}$ 条蓝色直线.

因此,当 n 充分大时,有方案染 $\frac{2\sqrt{n}}{3}$ 条直线为蓝色,使得每一个有限区域的边界均不全为蓝色的.

文[5]第六题的评注中,提到美国队领队Po-Shen Loh 基于非初等的概率方法对充分大的 n 证明了 $\sqrt{n\lg n}$ 的情形,并鼓励参赛者继续研究探索. 事实上,当时他运用的方法框架便是以上解法,其中,将计数部分放缩到了更佳的情形,用到了以下定理:

对于一个超图 G,含有 N 个顶点,平均度数为 d,所有的边的大小为3. 假设任何两个顶点间至多有一条边联结,则可找到至少 $c\frac{N}{\sqrt{d}}\sqrt{\lg d}$ 个顶点,两两之间无连线,其中,c 为常数.[6]

需要解释的是,超图不同于以往的图,其每条边不是联结两个顶点,而是若干个顶点. 在以上定理中,每条边便联结了三个顶点,并

且要求任意两个顶点间至多只有一条边.

回到 IMO 题目中,可以将 n 条直线视为超图的顶点,每块有限区域视为超图的边. 类似地,假设有概率为 p 的可能将每条边染为蓝色,将所有超过四条边的有限区域视为三角形或四边形. 从所有的蓝边中,去除联结四个顶点的边及同时联结两个顶点的重复 3—边(这样的重复边至多两条,因为原题中两条直线至多同时属于两个不同的三角形),然后,剩余的图便符合定理的要求且能估计出度数的平均值.

套用定理,将定理中独立的顶点视为原题中需要染为蓝色的边,便得到 $\sqrt{n\lg n}$ 的情形了.

练 习 题

1. 有 n 位同学参加了2014 年中国数学奥林匹克冬令营,每位同学有一张准考证. 现收取所有的准考证并打乱顺序后再随机地分发给每一位学生. 假设 S 表示依旧拿到自己的准考证的学生数目. 证明:S 的数学期望为 1.

提示:将 S 分解为每位学生的示性随机变量即可.

2. $A_1,A_2,\cdots,A_s$ 为 $\{1,2,\cdots,M\}$ 的一些子集. 假设 $A_1,A_2,\cdots,A_s$ 不相互包含(即没有一个是另一个的子集),令 $a_i=|A_i|$. 证明:

$$\sum_{i=1}^{s}\frac{1}{C_M^{a_i}}\leqslant 1.$$

(1981,圣彼得堡数学竞赛(十年级))

提示:令 $\sigma=(x_1,x_2,\cdots,x_M)$ 为 $\{1,2,\cdots,M\}$ 的一个随机排列,这样的全体记为空间 Ω. E_i 表示 A_i 中的元素出现在 σ 的前 a_i 个的事件,即 $\{x_1,x_2,\cdots,x_{a_i}\}=A_i$ 的概率.

由于 A_i 之间互相不包含,故事件 E_i 之间互相不包含.

又 $P(E_i)=\dfrac{a_i!\cdot(M-a_i)!}{M!}=\dfrac{1}{C_M^{a_i}}$,故

$$\sum_{i=1}^{s}\frac{1}{C_M^{a_i}}=P\left(\bigcup_{i=1}^{s}E_i\right)\leqslant P(\Omega)=1.$$

3. 证明:在一次有 799 支队参加的比赛中,存在 14 支队,可以被分成两组,使得第一组中的所有队伍胜了第二组中的所有队伍.

(2008,伊朗国家队选拔考试)

提示:随机地选取七支队,结合例 5 及 $C_{x_i}^7$ 的凸性即可得到结论.

4. n 个 n 维向量 $\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_n\in\mathbf{R}^n$,每一个的长度小于 1,即

$$|\boldsymbol{v}_i|=\sqrt{v_{1i}^2+v_{2i}^2+\cdots+v_{ni}^2}\leqslant 1.$$

任取 n 个实数 $p_1,p_2,\cdots,p_n\in[0,1]$,记向量 $\boldsymbol{w}=p_1\boldsymbol{v}_1+p_2\boldsymbol{v}_2+\cdots+p_n\boldsymbol{v}_n$.

证明:存在 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n\in\{1,0\}$,记向量 $\boldsymbol{v}=\varepsilon_1\boldsymbol{v}_1+\varepsilon_2\boldsymbol{v}_2+\cdots+\varepsilon_n\boldsymbol{v}_n$.

则 $|\boldsymbol{w}-\boldsymbol{v}|\leqslant\dfrac{\sqrt{n}}{2}$.

提示:独立地从 $\{1,0\}$ 中取出 n 个数 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$,其中,

$P(\varepsilon_i=1)=p_i,P(\varepsilon_i=0)=1-p_i$.

接下来仿例 6 即可.

5. 对 $3\leqslant k\leqslant n$,令 $X=\{1,2,\cdots,n\}$,F_k 为 X 中一些 k 元子集的集合族,使得其中任两个 k 元子集至多有 $k-2$ 个共同的元素.

(1997,中国台湾数学奥林匹克)

提示:类似例 8 的计算,放缩时需要运用不等式

$C_t^k\leqslant 3\times 2^{t-3}(3\leqslant k\leqslant t)$.

这可用数学归纳法证明.

参考文献:

[1] 宋 波,安永宏,杨志龙. 高三概率复习策略实验研究[J]. 数学教育学报,2013(4).

[2] 第 47 届 IMO 预选题(下)[J]. 中等数学,2007(11).

[3] 李 晕 译. 2012 美国数学奥林匹克[J]. 中等数学,2013(增刊二).

[4] Martin Aigner, Gunter M. Ziegler. Proofs from the Book [M] 4th ed. Springer—Verlag Berlin and Herdelberg GmbH & Co. K. 2010.

[5] 第 55 届 IMO 试题解答[J]. 中等数学,2014(9).

[6] R. Duke, H. Lefmann, V. Rodl. On Uncrowded Hypergraphs [J]. Random Structures and Algorithms, 1995(6).