第三节 伯努利试验与直线上的随机游动

一、伯努利概型

在许多问题中，我们对试验感兴趣的是某一类结果(事件A)是否出现．例如在产品抽样检查中注意的是抽到废品，还是抽到正品；在抛掷硬币时注意的是出现正面还是出现反面；在电视节目收视率的调查中注意的调查对象是否观看了节目.

在这类问题中我们可把一次试验的事件域取为

$$F=\left\{\phi, A, \overline{A}, \Omega\right\}$$

并称出现A为“成功”，$\overline{A}$为“失败”．这种只有两个结果的试验称为伯努利试验.

在伯努利试验中，首先要给出下面的概率：

$$P(A)=p,\quad P(\overline{A})=q=1-p.$$

其中： $p\geq 0,\ q\geq 0,\ p+q=1.$

现在考虑 n 重伯努利试验. 即满足

(1)每次试验至少出现两个可能结果之一：A与$\overline{A}$.

(2)事件A在每次试验中出现的概率p保持不变.

(3)各次试验相互独立,

(4)共进行n次试验.

n重伯努利试验的基本结果可以记作

$$\omega=(\omega_1,\omega_2,\cdots,\omega_n)$$

其中ω_i或者为A,或者为$\overline{A}$,这样的ω_i共有2^n个.

这2^n个样本点组成了样本空间Ω.

设样本点$\omega=(\omega_1,\omega_2,\cdots,\omega_n)$中有$k$个$A$，$n-k$个$\bar{A}$.
所以由独立性知

$$P(\omega)=p^k q^{n-k}.$$

每个样本点的概率可由上式得到，因而任何事件的概率都可计算出来.

例如：三重伯努利试验共有8个样本点:

$$(\bar{A},\bar{A},\bar{A}),\ (A,\bar{A},\bar{A}),\ (\bar{A},A,\bar{A}),\ (\bar{A},\bar{A},A),$$
$$(A,A,\bar{A}),\ (A,\bar{A},A),\ (\bar{A},A,A),\ (A,A,A).$$

概率分别为

$$p^0q^3 \qquad p^1q^2 \qquad p^1q^2 \qquad p^1q^2$$
$$p^2q^1 \qquad p^2q^1 \qquad p^2q^1 \qquad p^3q^0$$

历史上，伯努利概型是概率论中最早研究的模型之一，也是得到最多研究的模型之一, 在理论上和应用上都有重要的意义.

二、伯努利概型中的一些分布

1、伯努利分布

如果只进行一次伯努利试验，则或是事件A出现或是事件$\overline{A}$出现. 其概率为

$$P(A)=p, \qquad P(\overline{A})=q=1-p.$$

其中：$$p\geq 0,\ \ q\geq 0,\ \ p+q=1.$$

2、 二项分布

下面求在 n 重伯努利试验中事件 A 出现 k 次的概率. 记该事件为B_k，其概率记为 $b(k;n,p)$.

显然事件B_k共包含C_n^k个样本点. 其中任何一个样本点的概率都是$p^k q^{n-k}$. 因而事件B_k的概率为

$$b(k;n,p)=C_n^k p^k q^{n-k}, \qquad k=0,1,2,\cdots,n$$

注意到 $b(k;n,p)$ 是二项式 $(q+p)^n$ 展开式的一般项，因此称为二项分布. 显然有

$$\sum_{k=0}^{n} b(k;n,p)=\sum_{k=0}^{n} C_n^k p^k q^{n-k}=(q+p)^n=1$$

【例1】若在N件产品中有M件次品，现进行n次有放回的抽样调查，问共抽得k件次品的概率是多少？

解：由于抽样是有放回的, 因此这是n重伯努利试验，记A表示抽得次品这一事件，则

$$p = P(A) = \frac{M}{N}$$

因此所求的概率为

$$b(k;n,p) = C_n^k \left(\frac{M}{N}\right)^k \left(1 - \frac{M}{N}\right)^{n-k}.$$

3、几何分布

现在求在伯努利试验中首次成功(事件A第1次发生)出现在第k次试验的概率.

要使首次成功出现在第k次试验，必须而且只需在前$k-1$次试验中都出现$\overline{A}$，而在第k次试验出现A，因此该事件(记为W_k)可表示为

$$W_k = \overline{A}_1 \overline{A}_2 \cdots \overline{A}_{k-1} A_k$$

其中A_i表示事件A在第 i 次试验出现，$\overline{A}_i$表示事件A在第i次试验不出现，根据试验的独立性得

$$P(W_k) = P(\overline{A}_1 \overline{A}_2 \cdots \overline{A}_{k-1} A_k) = q^{k-1} p.$$

记 $$g(k;p)=q^{k-1}p,\qquad k=1,2,\cdots$$

$g(k;p)$是几何级数的一般项，因此上式称为几何分布. 显然有：

$$\sum_{k=1}^{\infty}g(k;p)=\sum_{k=1}^{\infty}q^{k-1}p=p\cdot\frac{1}{1-q}=1.$$

【例2】一个人要开门，他有n把钥匙，其中仅有一把可以开这扇门. 他随即地选取一把钥匙开门，即在每次试开时每把钥匙被选取的概率都是$1/n$. 求这人在第s次试开时才首次成功的概率是多少？

解 这是一个伯努利概型，所求概率为

$$g(s;\frac{1}{n})=(\frac{n-1}{n})^{s-1}\cdot\frac{1}{n}.$$

下面讨论的是更复杂一点的情况，即帕斯卡分布，它是几何分布的一种推广.

4、帕斯卡分布（负二项分布）

考虑在贝努利试验中，需要多少次试验才会出现第 r 次成功？

若以C_k表示第r次成功出现在第 k 次试验这一事件，并以$f(k;r,p)$记其概率.

C_k发生 $\Leftrightarrow$ 前面的$k-1$ 次试验中有$r-1$次出现成功，$k-r$次失败，而在第k次试验的结果是成功.

这两个事件的概率分别为:

$$C_{k-1}^{r-1}p^{r-1}q^{k-r} \quad 和 \quad p.$$

利用试验的独立性得

$$f(k;r,p)=C_{k-1}^{r-1}p^{r-1}q^{k-r}\cdot p=C_{k-1}^{r-1}p^{r}q^{k-r}.$$

注意到：

$$\sum_{k=r}^{\infty}f(k;r,p)=\sum_{k=r}^{\infty}C_{k-1}^{r-1}p^{r}q^{k-r}=\sum_{l=0}^{\infty}C_{r+l-1}^{l}p^{r}q^{l}$$

$$=\sum_{l=0}^{\infty}C_{-r}^{l}(-1)^{r}p^{r}q^{l}=p^{r}(1-q)^{-r}=1.$$

$f(k;r,p)$称为帕斯卡分布．特别当$r=1$时，即为几何分布.

【例3】（分赌注问题）甲、乙两个赌徒按照如下约定进行赌博：先胜t局者将赢得全部赌注，假定每局甲胜的概率为p，乙胜的概率为$q=1-p$，没有平局出现. 但进行到甲胜r局，乙胜s局（$r<t, s<t$）时，因故不得不中止比赛. 试问应如何分配这些赌注才公平合理？

历史上，这个问题曾引起不少人的兴趣，帕斯卡、费马及惠更斯分别给出了正确的答案. 帕斯卡、费马考虑甲乙两人最终获胜的概率$P_{甲}$及$P_{乙}$(等于$1-P_{甲}$)，并以$P_{甲}:P_{乙}$作为赌注的分配比例. 而惠更斯则引入了数学期望的概念.

记$n=t-r$及$m=t-s$，它们分别是甲、乙达到最终获胜所需再胜的局数.则分赌注问题归结为：在伯努利试验中，求在$\overline{A}$出现m次之前A出现n次的概率.

甲最终获胜$\longleftrightarrow$甲再胜第n局时，乙至多再胜$m-1$局.

不妨设乙再胜k局，$k=0,1,2,\ldots,m-1$. 此时，需要再比赛$n+k$局，甲在前$n+k-1$局中胜$n-1$局，最后一局甲胜. 利用帕斯卡分布得

$$p_{甲}=\sum_{k=0}^{m-1} f(k;n+k,p)=\sum_{k=0}^{m-1} C_{n+k-1}^{k}p^{n}q^{k}$$

甲最终获胜$\longleftrightarrow$乙再胜第m局时，甲再胜的局数不少于n局.

不妨设甲再胜k局,$k = n, n+1, n+2, \ldots$. 此时需要再比赛$m+k$局，乙在前$m+k-1$局中胜$m-1$局，最后一局乙胜. 利用帕斯卡分布得?

$$p_{甲} = \sum_{k=n}^{\infty} f(k; m+k, p) = \sum_{k=n}^{\infty} C_{m+k-1}^{k} p^k q^m$$

甲最终获胜$\longleftrightarrow$甲只需在接下来的$m+n-1$局比赛中至少胜n局.

不妨设甲再胜k局,$k = n, n+1, n+2, \ldots, m+n-1$. 此时需要再比赛$m+n-1$局，甲获胜的局数不少于$n$局，利用二项分布得

$$p_{甲}=\sum_{k=n}^{m+n-1} b(k;m+n-1,p)=\sum_{k=n}^{m+n-1} C_{m+n-1}^{k} p^{k} q^{m+n-1-k}$$

可以证明：上述三个答案是相等的（习题29）.

【例4】(巴拿赫火柴问题)数学家的左右衣袋中各放有一盒装有N根火柴的火柴盒，每次抽烟时任取一盒用一根，求他发现一盒用光时，另外一盒有r根的概率.

解　问题相当于抛掷硬币一直到正面或反面出现$N+1$次时，另一面只出现$N-r$次的概率.

当正面出现$N+1$次时，反面出现$N-r$次.该事件的概率为

$$f(2N-r+1;N+1,\frac{1}{2})=C_{2N-r}^{N}(\frac{1}{2})^{2N-r+1}.$$

因而所求事件的概率为

$$2f(2N-r+1;N+1,\frac{1}{2})=C_{2N-r}^{N}(\frac{1}{2})^{2N-r}.$$

三、直线上的随机游动

考虑x轴上的一个质点，假定它只能位于整数点，在时刻$t=0$时,它处在初始位置a(a是整数),以后每隔单位时间，它总是受到一个外力的随机作用，使位置发生变化，分别以概率p及概率$q=1-p$?向正的或负的方向移动一个单位，我们所关心的是质点在时

刻$t=n$时的位置. 用这种方式描述的质点运动称为随机游动.

若质点可以在整个数轴的整数点上游动，则称这种随机游动为无限制随机游动. 若在某点d设有一个吸收壁，质点一到达这点就被吸收而不再游动，因而整个游动也就结束了，这种随机游动称为在点d有吸收壁的随机游动. 此外还可以考虑带有反射壁及弹性壁的随机游动等类型.

当$p=q=0.5$时，随机游动称为是对称的，这时质点向右或向左的概率是一样的. 这里只介绍两种最简单的随机游动模型.

无限制的随机游动

假定质点在时刻 $t=0$ 从原点出发，以S_n表示它在时刻$t=n$时的位置. 为了使质点在时刻$t=n$时位于k (k也可以是负整数)，当且仅当在前n次游动中向右游动的次数(记为x)比向左游动的次数(记为y)多k次. 故有

$$\begin{cases} x+y=n \\ x-y=k \end{cases} \quad \text{解得} \quad x=\frac{n+k}{2},\quad y=\frac{n-k}{2}.$$

因为x , y为整数，所以k必须与n具有相同的奇偶性.

事件$\{S_n=k\}$发生相当于在前n次游动中有$\dfrac{n+k}{2}$

次向右，有$\dfrac{n-k}{2}$次向左，由二项分布得

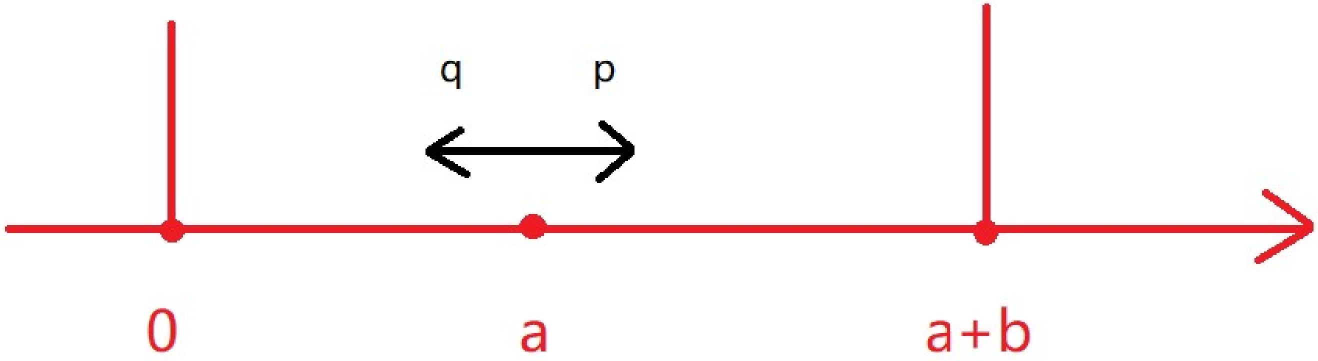
$$P\{S_n = k\} = C_n^{\frac{n+k}{2}} p^{\frac{n+k}{2}} q^{\frac{n-k}{2}}.$$

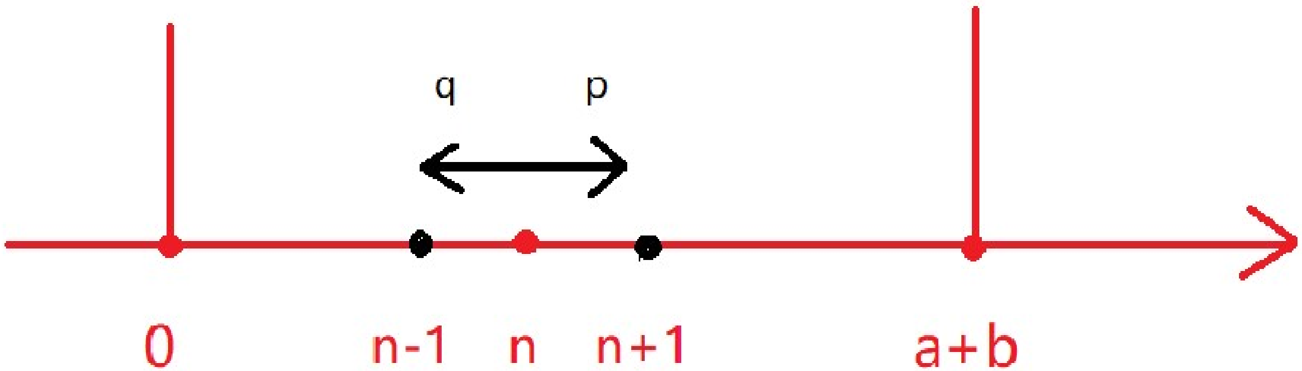
当k与n奇偶性相反时，概率为0.

两端带有吸收壁的随机游动

假定质点在$t = 0$时刻时，位于$x = a$，而在$x = 0$及$x = a + b$处各有一个吸收壁，下面求质点在$x = 0$被吸收或在$x = a + b$被吸收的概率. 用差分方程法.

以q_n表示事件“质点的初始位置为n而最终在





$$q_n = pq_{n+1} + qq_{n-1}, \qquad n = 1, 2, \cdots, a+b-1.$$

$$(p+q)q_n = pq_{n+1} + qq_{n-1}, \qquad n = 1, 2, \cdots, a+b-1.$$

$$p(q_{n+1} - q_n) = q(q_n - q_{n-1}), \qquad n = 1, 2, \cdots, a+b-1.$$

$x=a+b$点被吸收"的概率. 显然有

$$q_0=0,\qquad q_{a+b}=1.$$

如果某时刻质点位于$x=n$，这里$1\le n\le a+b-1$，则它要被$x=a+b$点吸收, 有两种方式来实现：一种是接下去一次移动是向右的, 即初始位置变为$n+1$而最终在$x=a+b$点被吸收. 根据全概率公式有

$$q_n=pq_{n+1}+qq_{n-1},\qquad n=1,2,\cdots,a+b-1.$$

将上述关于q_n的二阶差分方程改写为

$$p(q_{n+1}-q_n)=q(q_n-q_{n-1}).\quad n=1,2,\cdots,a+b-1.$$

下面利用边界条件$q_0=0,\quad q_{a+b}=1$求解q_n.

（1）当$p=q=1/2$时

$$q_{n+1}-q_n=q_n-q_{n-1}.$$

即$q_0,q_1,\cdots,q_{a+b}$成等差数列. 根据条件$q_0=0$, $q_{a+b}=1$. 易得公差$d=(a+b)^{-1}$. 所以

$$q_n=q_0+nd=n(a+b)^{-1}. \quad n=1,2,\cdots,a+b-1.$$

$$q_a=a(a+b)^{-1}.$$

（2）当 $p\neq q$ 时

$$q_{n+1}-q_n=\frac{q}{p}(q_n-q_{n-1})$$

所以

$$q_{n+1}-q_n=(\frac{q}{p})^n(q_1-q_0). \quad n=1,2,\cdots,a+b-1.$$

故 $$q_{a+b}-q_0=\sum_{k=0}^{a+b-1}(q_{k+1}-q_k)=\sum_{k=0}^{a+b-1}\left(\frac{q}{p}\right)^k(q_1-q_0)$$

$$q_{a+b}-q_n=\sum_{k=n}^{a+b-1}(q_{k+1}-q_k)=\sum_{k=n}^{a+b-1}\left(\frac{q}{p}\right)^k(q_1-q_0)$$

上面两式相比得

$$\frac{q_{a+b}-q_n}{q_{a+b}-q_0}=\frac{\sum_{k=n}^{a+b-1}\left(\frac{q}{p}\right)^k(q_1-q_0)}{\sum_{k=0}^{a+b-1}\left(\frac{q}{p}\right)^k(q_1-q_0)}=\frac{\left(\frac{q}{p}\right)^n-\left(\frac{q}{p}\right)^{a+b}}{1-\left(\frac{q}{p}\right)^{a+b}}$$

即 $$1-q_n=\frac{\left(\frac{q}{p}\right)^n-\left(\frac{q}{p}\right)^{a+b}}{1-\left(\frac{q}{p}\right)^{a+b}},\qquad n=1,2,\cdots,a+b-1.$$

所以 $$q_n = \frac{1-(\frac{q}{p})^n}{1-(\frac{q}{p})^{a+b}}, \quad n = 1, 2, \cdots, a+b-1.$$

特别地 $$q_a = \frac{1-(\frac{q}{p})^a}{1-(\frac{q}{p})^{a+b}}.$$

以p_n表示事件“质点的初始位置为n 而最终在$x=0$点被吸收”的概率. 显然有

$$p_n = pp_{n+1} + qp_{n-1}, \quad n = 1, 2, \cdots, a+b-1.$$

及边界条件 $p_0 = 1, \quad p_{a+b} = 0.$

类似可求得：

当 $p=q=1/2$ 时，$$p_a=\frac{b}{a+b}.$$

当 $p\neq q$时，$$p_a=\frac{\left(\frac{q}{p}\right)^a-\left(\frac{q}{p}\right)^{a+b}}{1-\left(\frac{q}{p}\right)^{a+b}}.$$

不论在什么场合，都有 $p_a+q_a=1$，也就是说，随机游动的质点最终一定被两个吸收壁之一所吸收.

赌徒输光问题

甲、乙两人进行赌博，其赌本分别为a及b，若每局赌注为1，而甲、乙在每局中赢的概率分别为p及$q=1-p$，试求乙（或甲）赌本输光的概率. 这个问题即是上面的两端带有吸收壁的随机游动. 根据计算结果我们知道：最后一定有一人会把赌本输光.

(1) 当$p=q=1/2$时，甲、乙两人输光的概率之比为

$$p_a : q_a = \frac{b}{a+b} : \frac{a}{a+b} = b : a$$

即：对赌技相同的两个人来说，输光的概率与赌本成反比，赌本越小，输光的可能性越大.

（2）当 $p < q$,且 $a \le b$ 时，　记 $x = q/p > 1.$

甲、乙两人输光的概率之比为

$$\frac{p_a}{q_a} = \frac{x^a - x^{a+b}}{1 - x^a} = \frac{x^a(x^b - 1)}{x^a - 1} \ge x^a > 1.$$

即：赌技较差，赌本较小的人，输光的可能性很大.

四、推广的伯努利试验与多项分布

二项分布可以容易地推广到n次重复独立试验且每次试验可能有若干结果的情形. 把每次试验的可能结果记为

$$A_1, A_2, \cdots, A_r$$

记 $p_i = P(A_i) \geq 0,\ i = 1,2,\cdots,r.$ 且$p_1 + p_2 + \cdots + p_r = 1.$

在这种推广的伯努利试验中，不难推得：在n 次试验中事件A_1出现k_1次，A_2出现k_2 次$,\cdots,$ A_r出现k_r次的概率为

$$\frac{n!}{k_1!k_2!\cdots k_r!} p_1^{k_1} p_2^{k_2} \cdots p_r^{k_r}$$

这里 $k_1 + k_2 + \cdots + k_r = n,\ k_i \geq 0.$ 上述公式称为多项分布. 它是$(p_1 + p_2 + \cdots + p_r)^n$的展开式的一般项，

而且显然有

$$\sum_{\substack{k_1+\cdots+k_r=n \\ k_i\geq 0}} \frac{n!}{k_1!k_2!\cdots k_r!} p_1^{k_1} p_2^{k_2} \cdots p_r^{k_r} = 1.$$

当$r=2$时，即是伯努利试验. 因此多项分布是二项分布的推广. 二项分布中的许多结果都能平行地推广到多项分布的场合.

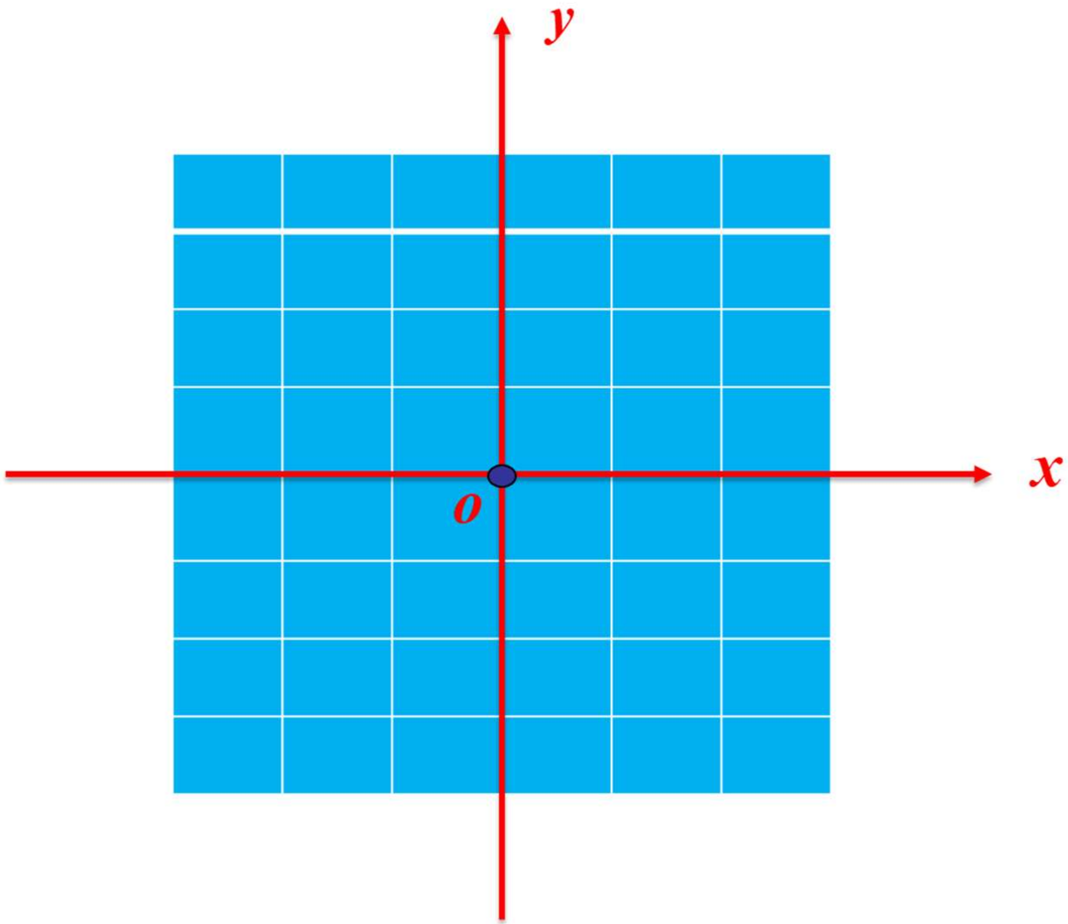
在产品检查中，若对产品质量所有的标准不只是合格品和不合格品，而是分得更细，例如有一等品，二等品，三等品，等外品四类，则从中取出n件，求一等品有k_1件，二等品有k_2件，三等品有k_3件，等外品有k_4件的概率就是多项分布.

【例5】人类的血型分为A,B,O,AB四种，假定某地区的居民中这四种血型的人的百分比分别为 0.4, 0.3, 0.25, 0.05，若从该地区的居民中随机地选出5人，求有两个为O型，其他三个分别为A,B,AB的概率.

解　利用多项分布得，所求的概率为

$$P=\frac{5!}{2!\cdot 1!\cdot 1!\cdot 1!}\times 0.4^2\times 0.3\times 0.25\times 0.05=0.036.$$

【例6】（平面上的随机游动）一质点平面上某点出发，等可能地向上, 下, 左, 右方向移动，每次移动的距离为1，求经过 $2n$ 次移动后回到出发点的概率.



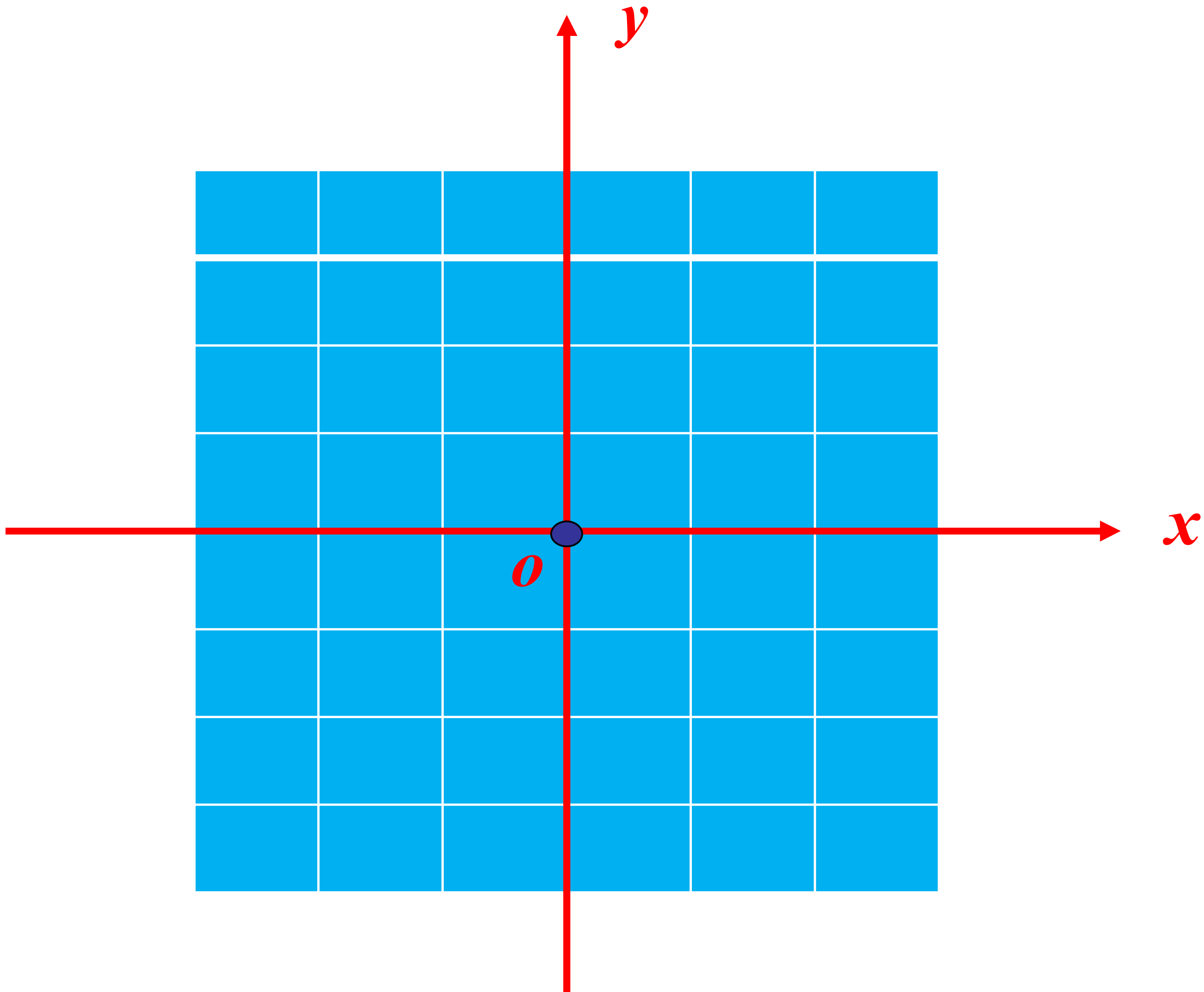
解 这可以归结为上述推广的伯努利试验的问题，分别以A_1，A_2，A_3，A_4表示质点向上, 下, 左，右移动一格，则

$$p_1 = p_2 = p_3 = p_4 = 1/4.$$

若要在$2n$次移动后回到出发点, 则向左移动的次数与向右移动的次数应该相等，向上移动的次数与向下移动的次数应该相等，而且总移动次数为$2n$次，假定质点向左、向右各移动k次，向上, 向下各移动m次. 因此所求的概率为

$$P=\sum_{k+m=n}\frac{(2n)!}{k!k!m!m!}\left(\frac{1}{4}\right)^{2n}=\left(\frac{1}{4}\right)^{2n}\frac{(2n)!}{(n!)^2}\sum_{k=0}^{n}\left[\frac{n!}{k!(n-k)!}\right]^2$$

$$=\sum_{k=0}^{n}\frac{(2n)!}{(k!)^2\left[(n-k)!\right]^2}\left(\frac{1}{4}\right)^{2n}=(C_{2n}^{n})^2\left(\frac{1}{4}\right)^{2n}$$


y
x
o