概　率　论

Probability

已学知识点

- 第一章 事件与概率
  - 随机现象与统计规律性

    ① 概率的频率解释依然是当今最通行的解释.

    ② 描述频率趋近于概率的大数定律总是概率论的第一大数定律.

    ③ 实际当中用频率作为概率的估计是十分自然的.
  - 样本空间与事件

符号	集合论含义	概率论含义
Ω	空间或全集	样本空间或必然事件
Φ	空集	不可能事件
ω	元素	样本点
A	子集	随机事件
$\omega \in A$	ω 是 A 的元素	事件 A 包含样本点 ω
$A \subset B$	A 是 B 的子集	A 发生则 B 发生
$AB = \Phi$	A, B 不相交	A, B 不可能同时发生
$A \cup B$	并集	A, B 至少有一个发生
$A \cap B$	交集	A, B 同时发生
$A - B$	差集	A 发生而 B 不发生
$\bar{A}$	余集	A 不发生

已学知识点

- 第一章 事件与概率
  - 古典概型 (等可能概率模型)：(1) 样本空间样本点有限；(2) 每个样本点等可能出现.
    - 计数方法：排列组合.
    - 三个基本性质：非负性、规范性、有限可加性.
  - 几何概率：以等可能性定义概率，处理无限场合，概率是几何体的测度之比.
    - 三个基本性质：非负性、规范性、可列可加性.
  - 概率空间：$(\Omega, \mathscr{F}, P)$
    - 难点和要点：事件域 $\mathscr{F}$ 的选择，太小不能满足需要，太大难以定义概率.
    - 选择包含我们关注的所有事件的 σ 域，保证事件对交、并、逆、差作可列次运算的封闭性.
    - 在这种 σ 域上，能定义满足非负、规范和可列可加性的概率测度.

已学知识点

- 第二章 条件概率与统计独立性

  - 条件概率的定义：$P(A\,|\,B) = \dfrac{P(AB)}{P(B)}.$

  - 条件概率也是概率：

    ① $P(A\,|\,B) \geq 0$

    ② $P(\Omega\,|\,B) = 1$

    A_i 两两互不相容

    ③ $P\left(\bigcup_{i=1}^{\infty} A_i \,\middle|\, B\right) = \sum_{i=1}^{\infty} P\left(A_i \,\middle|\, B\right)$

    ④ $P\left(\overline{A} \,\middle|\, B\right) = 1 - \left(A \,\middle|\, B\right)$

    ⑤ $P\left(A \cup B \,\middle|\, C\right) = P\left(A \,\middle|\, C\right) + P\left(B \,\middle|\, C\right) - P\left(AB \,\middle|\, C\right)$

已学知识点

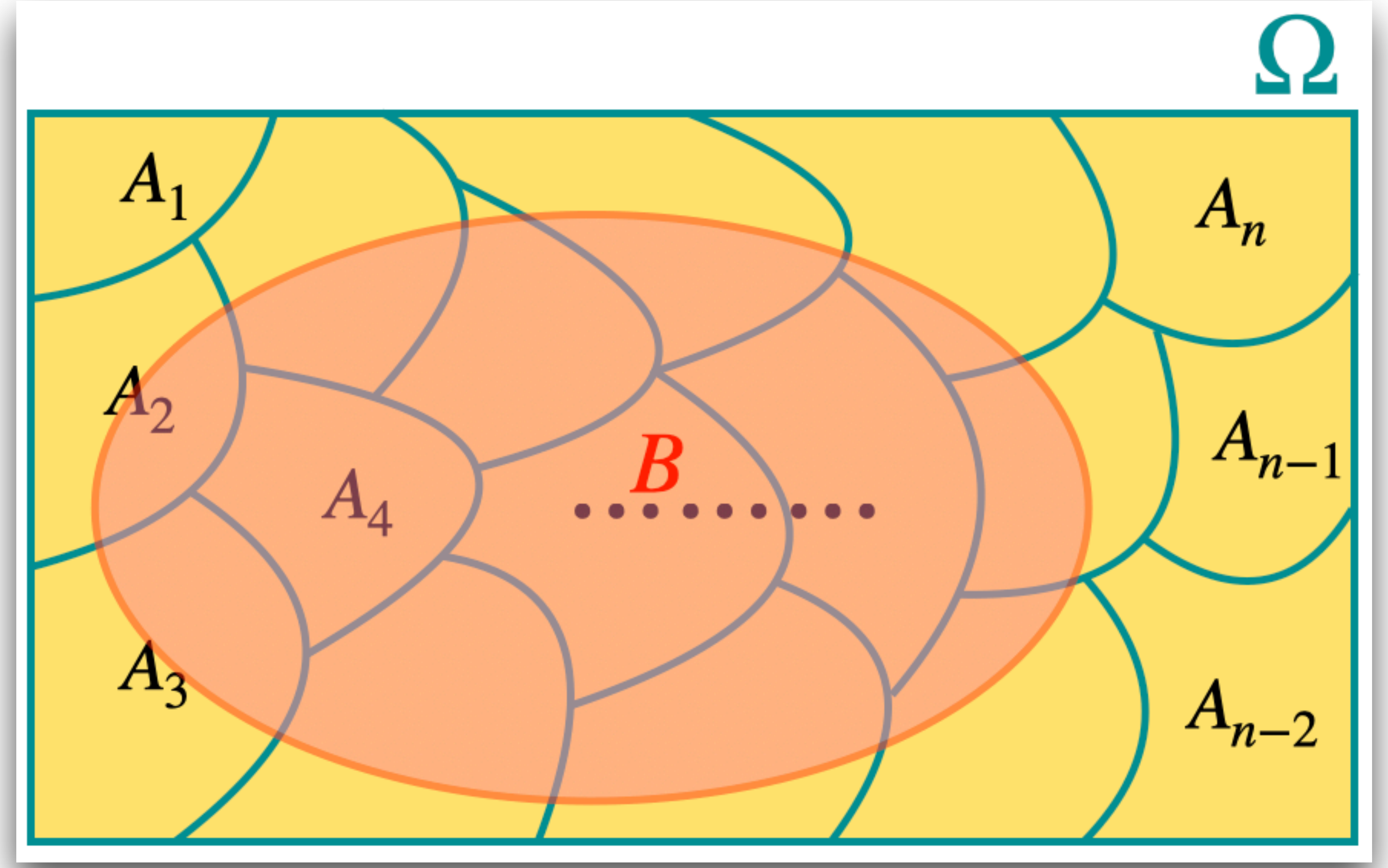
第二章 条件概率与统计独立性

- 乘法公式： $P(AB) = P(B) \cdot P(A \mid B) = P(A) \cdot P(B \mid A)$

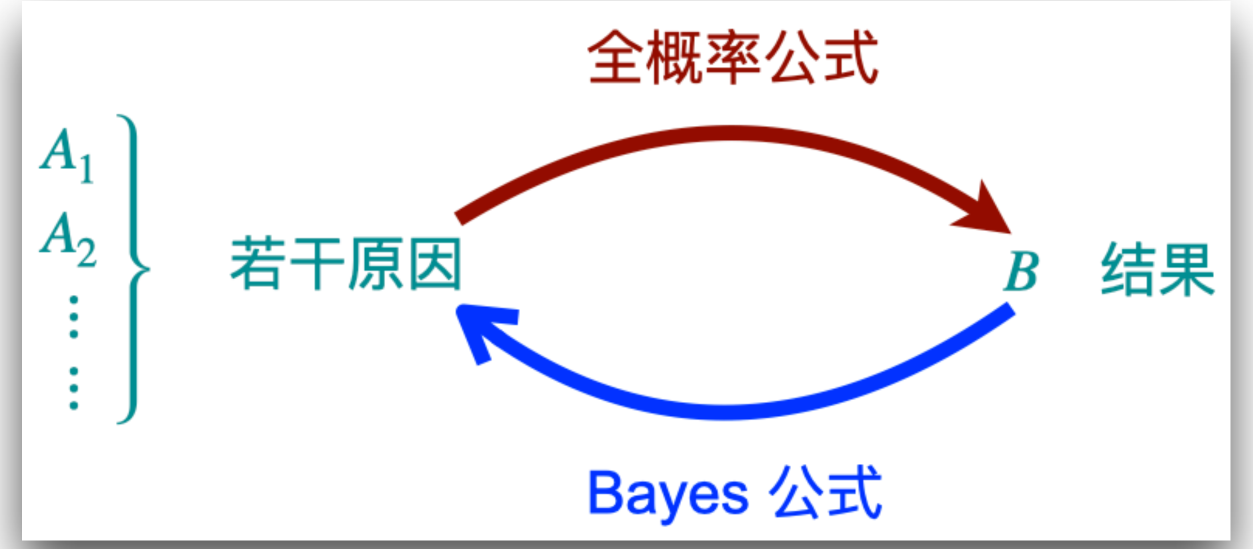
$$P\left(A_1A_2\cdots A_n\right) = P\left(A_1\right) \cdot P\left(A_2 \middle| A_1\right) \cdot P\left(A_3 \middle| A_1A_2\right) \cdot \cdots \cdot P\left(A_n \middle| A_1A_2\cdots A_{n-1}\right)$$

乘法公式主要用于计算若干个事件同时发生的概率.

- 全概率公式：$P(B) = \sum_{i=1}^{\infty} P\left(A_i\right) \cdot P\left(B \middle| A_i\right)$



- Bayes 公式：$P\left(A_k \middle| B\right) = \dfrac{P\left(A_k\right) \cdot P\left(B \middle| A_k\right)}{\sum_{i=1}^{\infty} P\left(A_i\right) \cdot P\left(B \middle| A_i\right)}$



已学知识点

- 第二章 条件概率与统计独立性
  - 两个事件独立：$P(AB) = P(A) \cdot P(B).$

$$\left.\begin{array}{l} \overline{A}\ ,\ B \\ A\ ,\ \overline{B} \\ \overline{A}\ ,\ \overline{B} \end{array}\right\} \text{均相互独立}$$

  - 三个事件独立：$\begin{cases} P(AB) = P(A) \cdot P(B) \\ P(AC) = P(A) \cdot P(C) \\ P(BC) = P(B) \cdot P(C) \\ P(ABC) = P(A) \cdot P(B) \cdot P(C) \end{cases}$
  - $A_1, A_2, \dots, A_n$ 相互独立，则

$$P\left(A_1 \cup A_2 \cup \cdots \cup A_n\right) = 1 - \left[1 - P\left(A_1\right)\right]\left[1 - P\left(A_2\right)\right] \cdot \cdots \cdot \left[1 - P\left(A_n\right)\right]$$

已学知识点

- 第二章 条件概率与统计独立性

  - 试验相互独立：一个试验的结果对其它各试验的可能结果的概率都无影响.
  - 试验独立性的定义：设 A_i 是试验 E_i 中的任一事件，$i=1,2,\ldots,n$，若

$$P\left(A_1A_2\cdots A_n\right)=P\left(A_1\right)\cdot P\left(A_2\right)\cdot\cdots\cdot P\left(A_n\right)$$

则称试验 $E_1, E_2,\ \ldots, E_n$ 是相互独立的.

  - 重复独立试验：研究在相同条件下重复进行的独立试验的数学模型.

2.3 Bernoulli 试验与直线上的随机游动

一、Bernoulli 概型

二、Bernoulli 概型中的一些分布

三、直线上的随机游动

四、推广的 Bernoulli 试验和多项分布

一、Bernoulli 概型

- 在许多问题中，人们往往关心实验中某一事件 A 是否发生. 例如

  ① 在产品质量抽样检测中是否抽到次品.

  ② 在掷硬币试验中是否出现正面.

  ③ 在股票市场中股票是涨还是跌等.

- 在这类问题中，我们可以把事件域取为

$$\mathscr{F}=\left\{\varnothing\ ,A\ ,\bar{A}\ ,\Omega\right\}$$

称出现事件 A 为成功，出现事件 $\bar{A}$ 为失败. 这种只有两个结果的试验称为伯努利 (Bernoulli) 试验.

- 具体而言，如果随机试验 E 只有两个结果：A 和 $\bar{A}$，其中 $P(A)=p,\ P\left(\bar{A}\right)=q,$ $\left(p>0,\ q>0,\ p+q=1\right)$，则称 E 为 伯努利 (Bernoulli) 试验.

一、Bernoulli 概型

- n 重伯努利试验： n 次独立重复的伯努利试验，记作 E_n，其特点是

  ① 每次试验最多出现两个可能结果.

  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 共重复进行了 n 次试验.

$$\omega_i = \begin{cases} A, & \text{第 } i \text{ 次试验中出现成功 } A \\ \overline{A}, & \text{第 } i \text{ 次试验中出现失败 } \overline{A} \end{cases}$$

- n 重伯努利试验的样本点： $\omega = (\omega_1, \omega_2, \ldots, \omega_n)$

这样的样本点共有 2^n 个. 相应的概率为

$$P\{(\underbrace{\omega_1, \omega_2, \ldots, \omega_n}_{k \text{ 个 } A})\} = p^k q^{n-k}$$

$$\overbrace{\overline{A}\ \overline{A}\cdots\cdots\overline{A}}^{n-k}\ \underbrace{A\ A\cdots\cdots A}_{k}$$

$$\overbrace{A\ A\cdots\cdots A}^{k}\ \underbrace{\overline{A}\ \overline{A}\cdots\cdots\overline{A}}_{n-k}$$

一、Bernoulli 概型

- Bernoulli 试验是一种非常重要的概率模型, 它是在同样条件下进行重复试验的一种数学模型, 特别在讨论某事件出现的频率时常用这种模型.
- 在历史上，Bernoulli 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，在理论上具有重要的意义.
- 另一方面，它有着广泛的应用，在我们这门课程中，一些较为深入的结果也是结合 Bernoulli 概型进行讨论的.

一、Bernoulli 概型

- 独立重复 Bernoulli 试验中的三个重要问题：

① n 次试验中 A 恰好发生 k 次的概率是多少?

② 到第 k 次试验 A 才首次发生的概率是多少?

③ 一直不停试验，A 最终发生的概率是多少?

二、Bernoulli 概型中的一些分布

- Bernoulli 分布
  - 只进行一次 Bernoulli 试验.
  - 概率： $P(A) = p,\ P\left(\bar{A}\right) = q = 1 - p, \quad \left(p + q = 1,\ p > 0,\ q > 0\right)$
  - Bernoulli 概型中最简单的情形.
  - 例：掷一枚硬币一次的随机试验.

样本点	H	T
概率	p	$1-p$

```
#### 模拟掷一枚硬币一次的随机试验
rm(list = ls(all = TRUE))
x = c("H", "T")
p = 0.5
?sample
sample(x, 1, replace = FALSE, prob = c(p, 1-p))
```

二、Bernoulli 概型中的一些分布

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

- 二项分布 (binomial distribution)

  - n 重 Bernoulli 试验中，成功 (事件 A) 出现的次数 (k 次)，相应的概率为

$$b(k;\ n, p) \triangleq P_n(k) = C_n^k p^k q^{n-k} = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, 2, \dots, n$$

定义 $\begin{cases} A_i : & \text{第 } i \text{ 次试验中出现 } A \\ B_k : & n \text{ 重 Bernoulli 试验中 } A \text{ 恰好出现 } k \text{ 次} \end{cases}$

$$\Longrightarrow \quad B_k = \underbrace{\overbrace{\left\{ \overbrace{A_1 A_2 \cdots A_k}^{k} \underbrace{\bar{A}_{k+1} \bar{A}_{k+2} \cdots \bar{A}_n}_{n-k} \right\}}^{p^k(1-p)^{n-k}} \bigcup \overset{\cdots\cdots\cdots}{\cdots\cdots\cdots} \bigcup \overbrace{\left\{ \overbrace{\bar{A}_1 \bar{A}_2 \cdots \bar{A}_{n-k}}^{n-k} \underbrace{A_{n-k+1} A_{n-k+2} \cdots A_n}_{k} \right\}}^{p^k(1-p)^{n-k}}}_{\binom{n}{k}}$$

二、Bernoulli 概型中的一些分布

- 二项分布 (binomial distribution)

- 例：设一批产品中有 a 件是次品，b 件是正品. 现有放回地从中抽取了 n 件产品. 求事件 A 的概率，其中 $A=\{n$ 件产品中恰有 k 件次品$\}$， $k=0,1,2,\ldots,n$.
  - 这是一个 n 重 Bernoulli 试验，成功代表取到一件次品，且成功的概率 $p=\dfrac{a}{a+b}$.

$$\Longrightarrow \quad P(A)=b(k;\ n,p)=C_n^k p^k q^{n-k}=\binom{n}{k}\left(\frac{a}{a+b}\right)^k\left(\frac{b}{a+b}\right)^{n-k}$$

```
#### 二项分布的模拟
rm(list = ls(all = TRUE))
n = 10
p = 0.5
?rbinom
rbinom(1, n, p)
```

二、Bernoulli 概型中的一些分布

- 几何分布 (geometric distribution)
  - n 重 Bernoulli 试验中，首次成功出现在第 k 次试验的概率，则

$$g(k;\ p) = q^{k-1}p = (1-p)^{k-1}p\ ,\quad k = 1, 2, \ldots$$

定义 $\begin{cases} A_i : & \text{第 } i \text{ 次试验中出现 } A \\ W_k : & n \text{ 重 Bernoulli 试验中首次出现 } A \text{ 恰好在第 } k \text{ 次} \end{cases}$

$$\Longrightarrow\quad W_1 = A_1\ ,\ W_k = \underbrace{\overline{A}_1\ \overline{A}_2\ \cdots\ \overline{A}_{k-1}}_{k-1}\ A_k\ ,\quad k = 1, 2, \ldots$$

$$(1-p)^{k-1}p$$

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

二、Bernoulli 概型中的一些分布

- 几何分布 (geometric distribution)

- 例：一个人要开门，他共有 n 把钥匙，其中仅有一把钥匙能打开这扇门. 他随机地选取一把钥匙开门，即在每次试开时每一把钥匙都以概率 $\frac{1}{n}$ 被使用. 问这个人在第 s 次试开时才首次成功打开门的概率是多少?

  - 这是一个 n 重 Bernoulli 试验，成功代表取到正确的钥匙打开了门，成功的概率 $p = \frac{1}{n}$.

$$\Longrightarrow \quad P\left(\text{第 } s \text{ 次试开时才首次成功打开门}\right) = g(k = s;\ p) = q^{s-1}p = \left(1 - \frac{1}{n}\right)^{s-1} \cdot \frac{1}{n}$$

```
#### 几何分布的模拟: R 中几何分布的定义为首次出现成功时失败发生的次数，故首次出现成功时试验的次数为 R 中的模拟结果 + 1
rm(list = ls(all = TRUE))
?Geometric
p = 0.3
rgeom(1, p) + 1
```

二、Bernoulli 概型中的一些分布

$$\sum_{k=r}^{\infty} f(k;\ r,p)$$

- 帕斯卡分布 (Pascal distribution)
  - n 重 Bernoulli 试验中，第 r 次成功出现在第 k 次试验的概率，则

$$f(k;\ r,p) = C_{k-1}^{r-1} p^r q^{k-r} = \binom{k-1}{r-1} p^r (1-p)^{k-r}\ ,\quad k = r,\ r+1,\ r+2,\ \ldots$$

定义 $\begin{cases} A_i & : \text{ 第 } i \text{ 次试验中出现 } A \\ C_k & : \ n \text{ 重 Bernoulli 试验中第 } r \text{ 成功出现在第 } k \text{ 次} \end{cases}$

$$\Longrightarrow \quad C_k = \{\text{前 } k-1 \text{ 次恰好出现 } r-1 \text{ 次成功}\} \bigcap \{\text{第 } r \text{ 次出现成功}\}$$

p

二项分布：$C_{k-1}^{r-1} p^{r-1} q^{k-r}$

$$P\left(C_k\right) = C_{k-1}^{r-1} p^{r-1} q^{k-r} \cdot p = \binom{k-1}{r-1} p^r (1-p)^{k-r}$$

二、Bernoulli 概型中的一些分布

- 帕斯卡分布 (Pascal distribution)
  - n 重 Bernoulli 试验中，第 r 次成功出现在第 k 次试验的概率，则

$$f(k;\ r, p) = C_{k-1}^{r-1} p^r q^{k-r} = \binom{k-1}{r-1} p^r (1-p)^{k-r}, \quad k = r, r+1, r+2, \ldots$$

$$l \triangleq k - r$$

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

$$= \boxed{\sum_{l=0}^{\infty} \binom{-r}{l} (-q)^l} \cdot p^r$$

$$= p^r (1-q)^{-r} = 1$$

牛顿二项式：$(1+x)^{\alpha} = \sum_{l=0}^{\infty} \binom{\alpha}{l} x^l$

$$\binom{-a}{k} = (-1)^k \binom{a+k-1}{k}$$

二、Bernoulli 概型中的一些分布

甲、乙两赌徒按某种方式下注赌博，先胜 t 局者将赢得全部赌注，但进行到甲胜 r 局、乙胜 s 局 $(r < t\ ,\ s < t)$ 时，因故不得不中止，问如何分配这些赌注才公平合理?

- 帕斯卡分布 (Pascal distribution)
- 分赌注问题：Pascal 分布的起源
  - 分配建议：用胜局比例 $r : s$ 分配！ 没有考虑到两人最终取胜的概率！
  - 分配建议：用甲、乙最终取胜的概率 $p_1 : p_2$ 分配.
  - 记每一局甲获胜为事件 A，设 $P(A) = p$，则每一局乙获胜的概率为 $P\left(\bar{A}\right) = 1 - p = q$.
  - 甲获胜还需要再胜 $n = t - r$ 局；乙获胜还需要再胜 $m = t - s$ 局.
  - 甲最终获胜：第 n 次出现 A 之前，$\bar{A}$ 出现的次数 k 小于 m 次 (即 $k = 0, 1, 2, \ldots, m - 1$).

    这种情形下，赌局还应进行的次数为 $n + k$ 次 $\left(k = 0, 1, 2, \ldots, m - 1\right)$.

Pascal 分布 $\Longrightarrow$ $P\left(\text{甲最终获胜}\right) = p_1 = \sum\limits_{k=0}^{m-1} C_{n+k-1}^{n-1} p^n q^k$

二、Bernoulli 概型中的一些分布

数学家衣服的左、右口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，求发现一盒用光时，另一盒尚有 r 根火柴的概率.

- 帕斯卡分布 (Pascal distribution)
- Banach 火柴盒问题
  - 每次任取一盒是 $p=\dfrac{1}{2}$ 的 Bernoulli 试验.
  - 考虑左边口袋空而右边尚有 r 根：左边取过 $N+1$ 次 (前 N 次用去 N 根火柴，第 $N+1$ 次发现空了) 而右边取过 $N-r$ 次.

Pascal 分布 $\Longrightarrow$ $P(\text{左空右剩 } r)=f\left(2N-r+1;\ N+1,\dfrac{1}{2}\right)=\dbinom{2N-r}{N}\left(\dfrac{1}{2}\right)^{2N-r+1}$

$\Longrightarrow$ $P(\text{右空左剩 } r)=f\left(2N-r+1;\ N+1,\dfrac{1}{2}\right)=\dbinom{2N-r}{N}\left(\dfrac{1}{2}\right)^{2N-r+1}$

$$P(\text{一盒用光另一盒剩 } r)=2\cdot f\left(2N-r+1;\ N+1,\frac{1}{2}\right)=\binom{2N-r}{N}2^{-2N+r}$$

二、Bernoulli 概型中的一些分布

- 小概率事件必然发生
  - n 重 Bernoulli 试验，$P(A) = p > 0$.
  - A_i 表示 A 在第 i 次试验时发生，$i = 1, 2, 3, \ldots$.

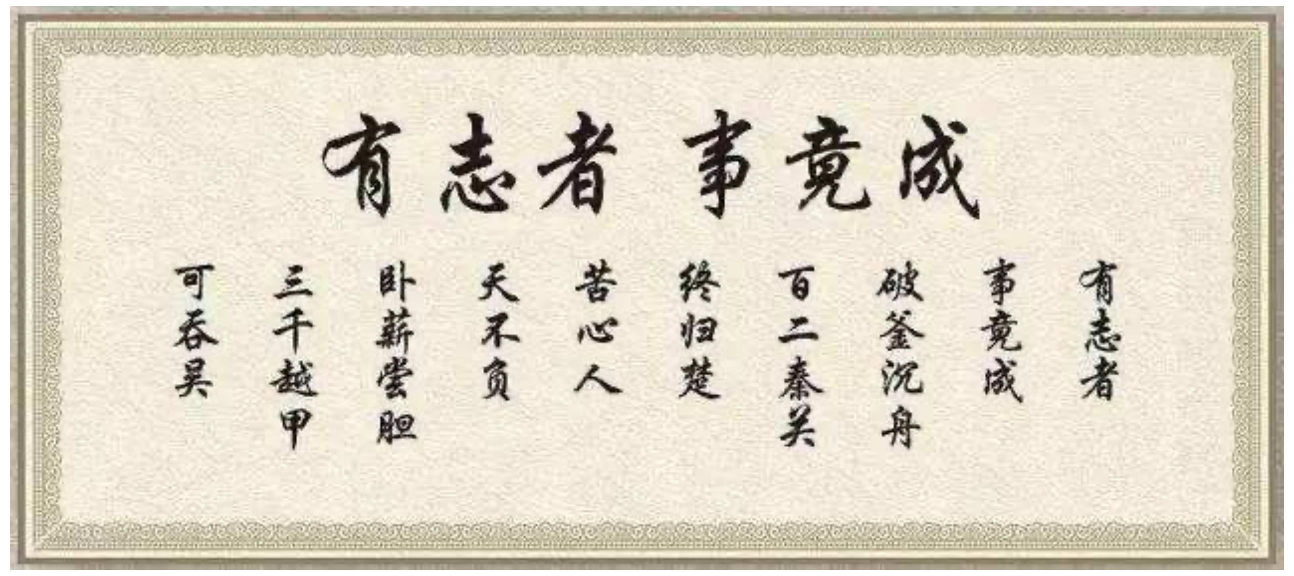
假定随机事件 A 在一次试验中发生的概率是 p (无论多小，但要 $p > 0$)，如果不停地独立重复进行该试验，那么 A 最终必然会发生.

A 最终必然发生 $\Longleftrightarrow$ $P\left(\bigcup_{i=1}^{\infty} A_i\right) = 1$

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

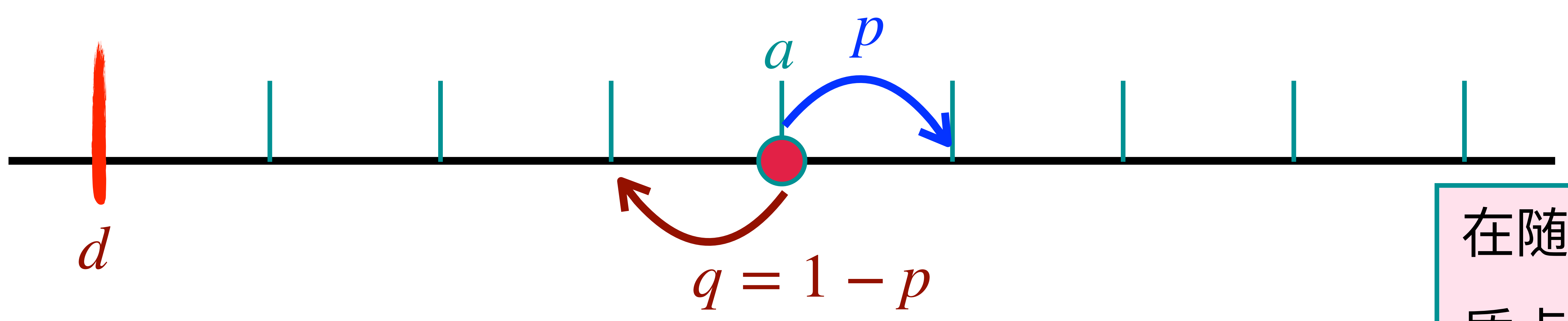
$$= 1 - \prod_{i=1}^{\infty} P\left(\overline{A_i}\right) = 1 - \prod_{i=1}^{\infty} (1 - p)$$

$$= 1 - \lim_{n \to \infty} (1 - p)^n = 1$$



三、直线上的随机游动

- 随机游动：考虑 x 轴上一个质点，在时刻 $t = 0$ 时它处于初始位置 a (a 是整数)，以后每隔单位时间，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.

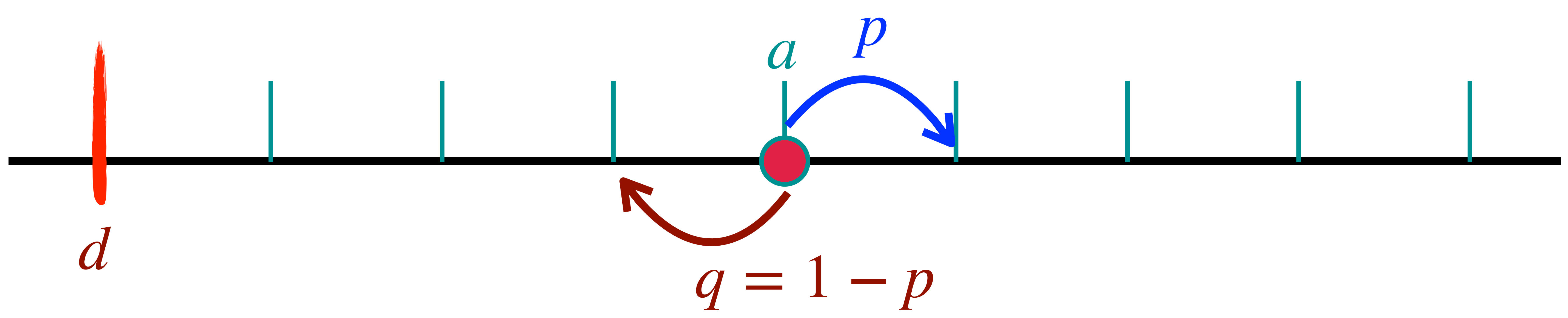


在随机游动模型中，我们关心质点在时刻 $t = n$ 时的位置.

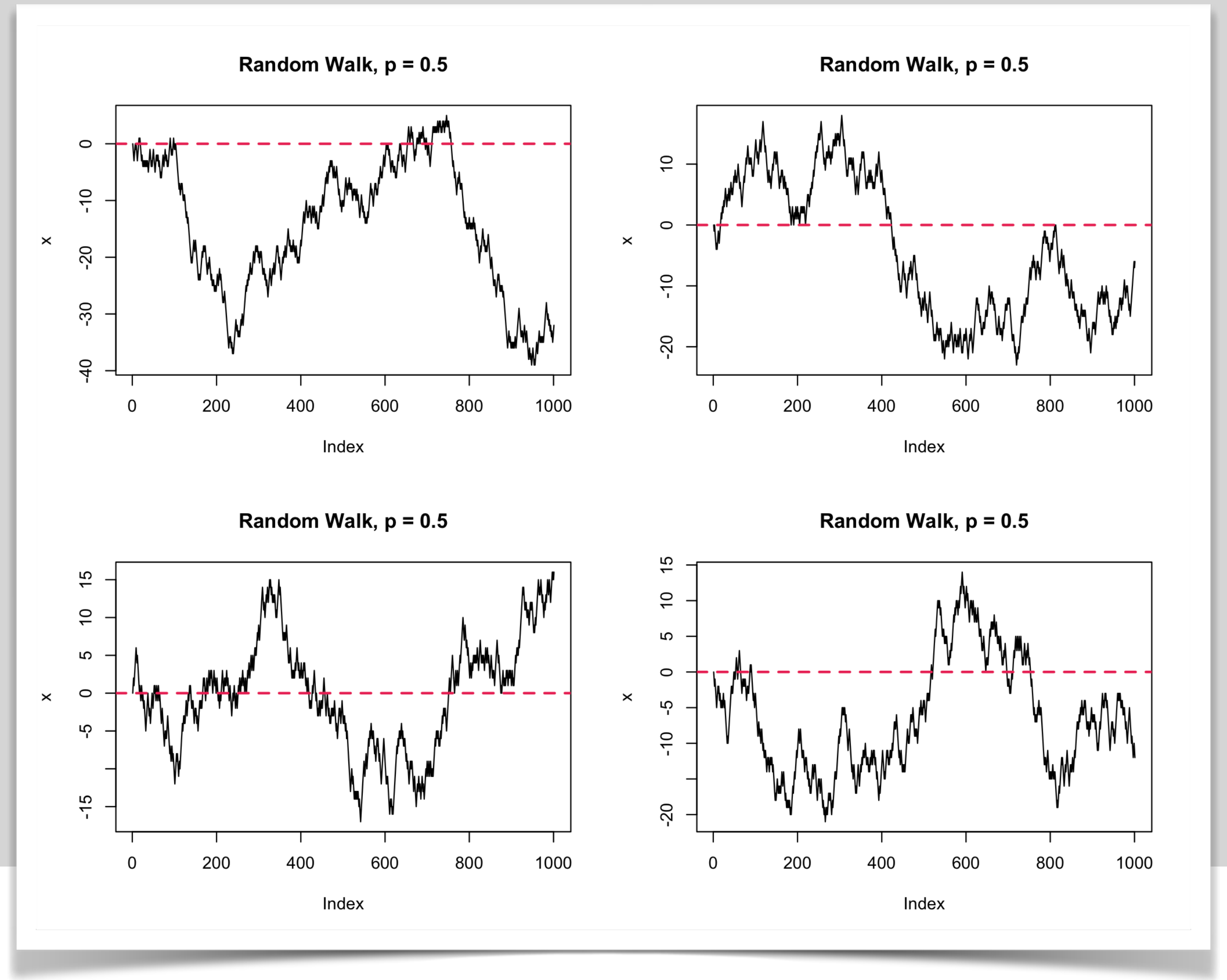
- 对称的随机游动：$p = q = \dfrac{1}{2}$ 时，这时质点向左或向右移动的可能性相等.
- 无限制随机游动：质点可以在整个数轴的整数点上游动.
- 在 d 点有吸收壁的随机游动：在 d 点设有一个吸收壁，质点一到达该点即被吸收而不再游动，因而整个游动就结束.

三、直线上的随机游动

- **随机游动**：考虑 x 轴上一个质点，在时刻 $t=0$ 时它处于初始位置 a (a 是整数)，以后每隔单位时间，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.

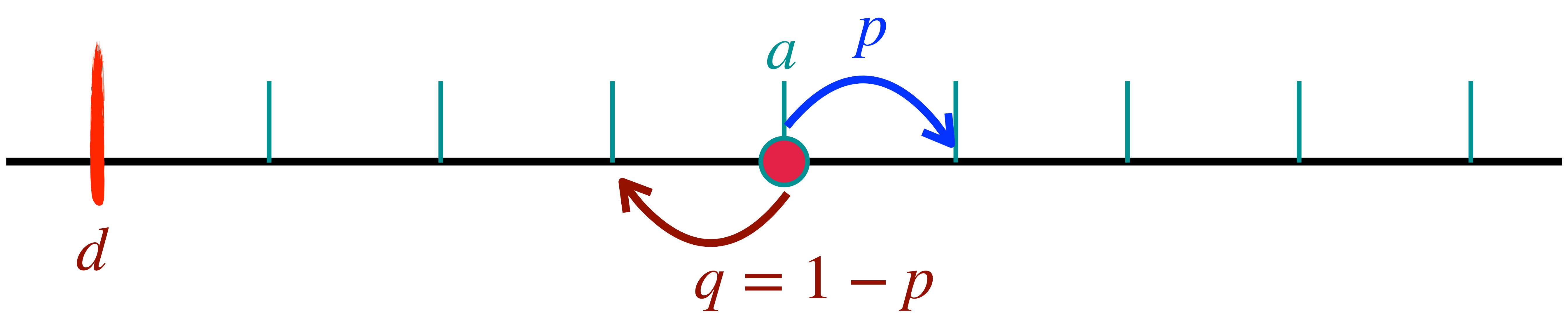



```
f = function(a = 0, N = 1000, p = 0.5){
  x = array(0, dim = N)
  for (i in 1:N) x[i+1] = x[i] + sample(c(-1, 1), 1, prob = c(1 - p, p))
  plot(x, type = 'l')
  abline(h=0, lty=2, col=2, lwd=2)
}
par(mfrow=c(2, 2))
for (i in 1:4) {
  f()
  title(main = 'Random Walk, p = 0.5')
}
```

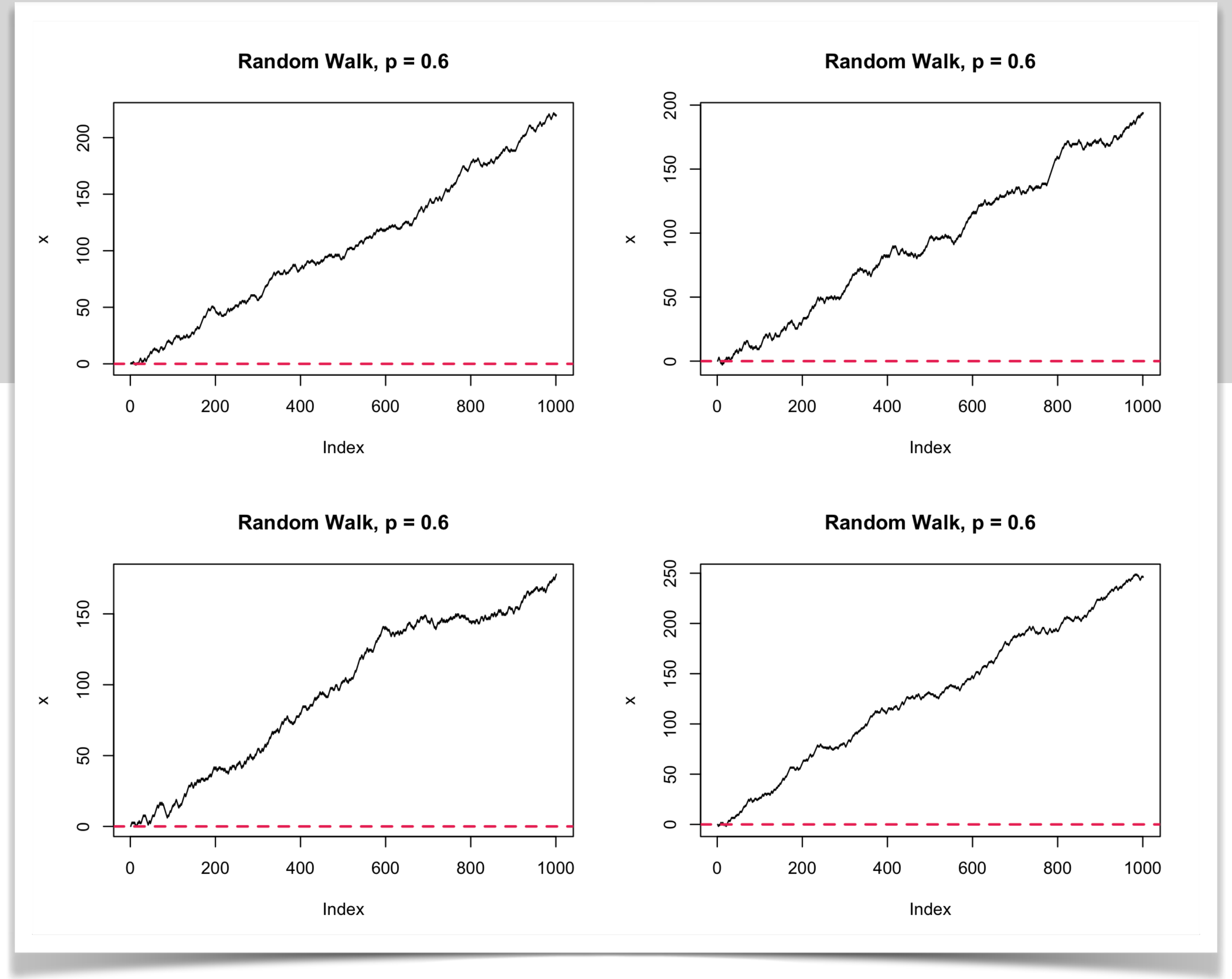


三、直线上的随机游动

- 随机游动：考虑 x 轴上一个质点，在时刻 $t = 0$ 时它处于初始位置 a (a 是整数)，以后每隔单位时间，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.

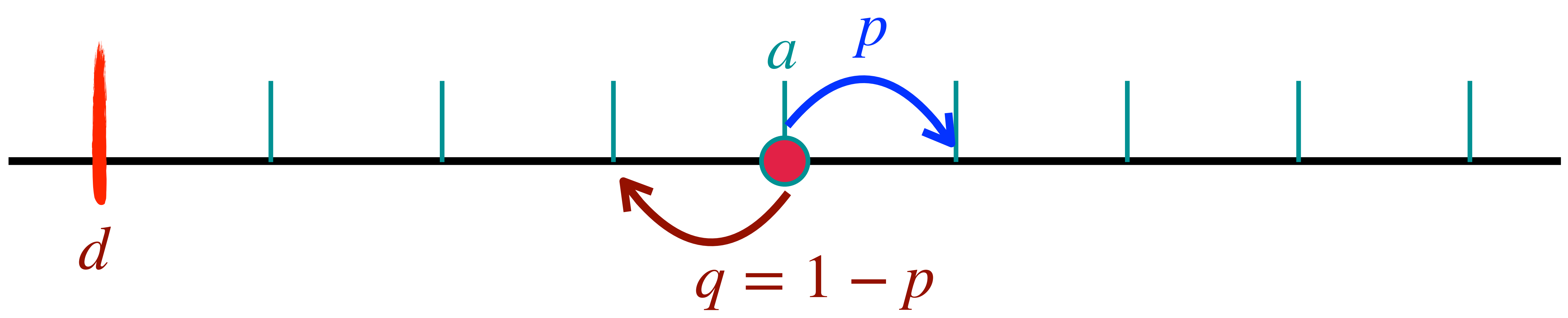



```
par(mfrow=c(2, 2))
for (i in 1:4) {
  f(p = 0.6)
  title(main = 'Random Walk, p = 0.6')
}
```

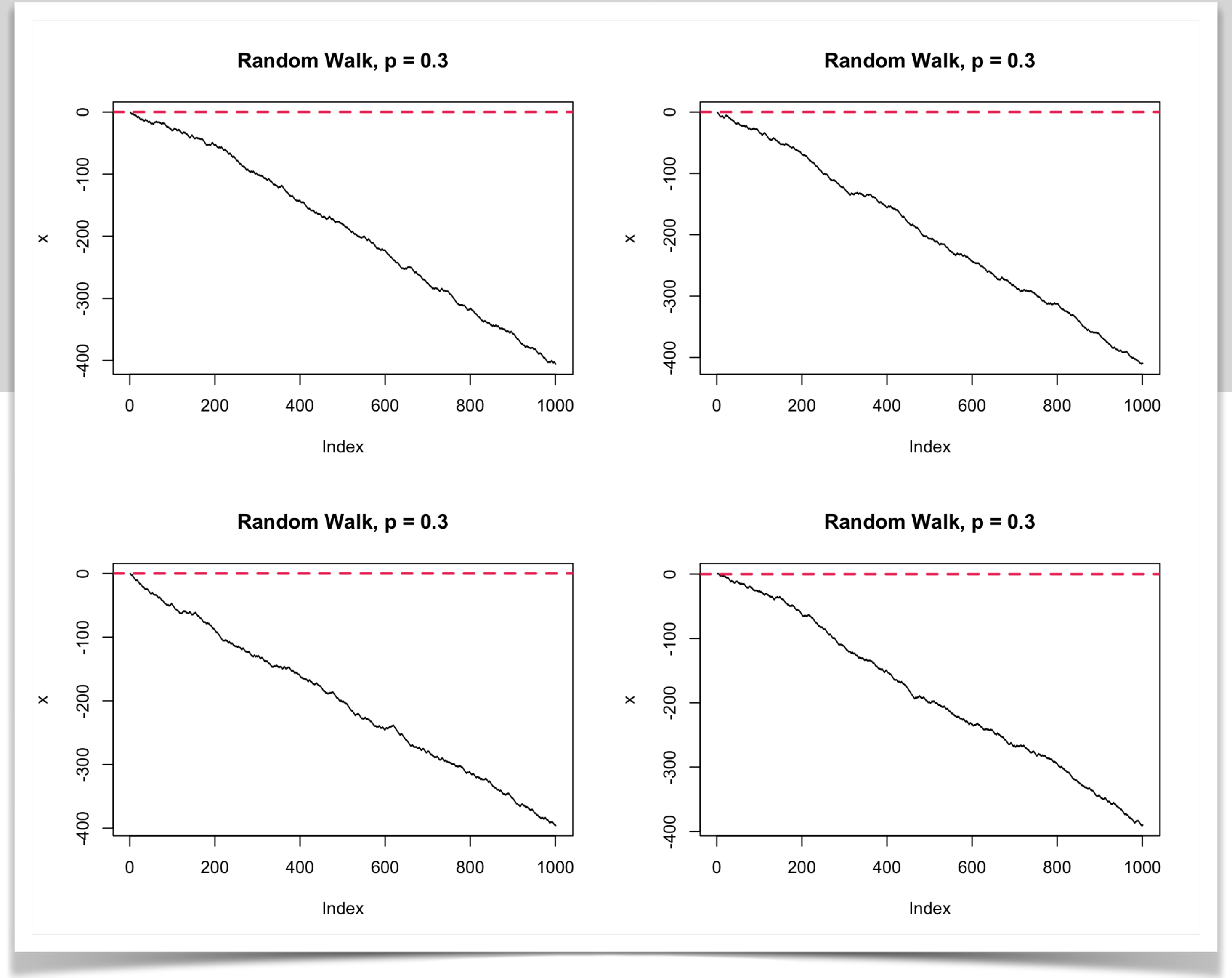


三、直线上的随机游动

- 随机游动：考虑 x 轴上一个质点，在时刻 $t = 0$ 时它处于初始位置 a (a 是整数)，以后每隔单位时间，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.




```
par(mfrow=c(2, 2))
for (i in 1:4) {
  f(p = 0.3)
  title(main = 'Random Walk, p = 0.3')
}
```



三、直线上的随机游动

- 无限制随机游动：有无穷赌资的赌徒在 n 局后的输赢
  - 假定质点在时刻 0 从原点出发，以 S_n 记它在时刻 n 的位置.
  - 为了使质点在时刻 $t = n$ 时位于 k (k 可以是负整数，$-n \le k \le n$)，必须且只须在前 n 次游动中向右移动的次数比向左移动的次数多 k 次.

定义 $\begin{cases} x : & \text{在 } n \text{ 次游动中向右移动的次数} \\ y : & \text{在 } n \text{ 次游动中向左移动的次数} \end{cases} \Longrightarrow \begin{cases} x + y = n \\ x - y = k \end{cases} \Longrightarrow \begin{cases} x = \dfrac{n+k}{2} \\ y = \dfrac{n-k}{2} \end{cases}$

  - 因为 x 是整数，所以 k 必须与 n 具有相同的奇偶性.

三、直线上的随机游动

- 无限制随机游动：有无穷赌资的赌徒在 n 局后的输赢

  - 事件 $\{S_n = k\}$ (赌徒在 n 局后赢 \$ k) 发生相当于要求在前 n 次游动中有 $x = \dfrac{n+k}{2}$ 次向右，$y = \dfrac{n-k}{2}$ 次向左.

$$\text{二项分布} \quad \Longrightarrow \quad P\{S_n = k\} = \binom{n}{\frac{n+k}{2}} p^{\frac{n+k}{2}} q^{\frac{n-k}{2}}$$

  - 当 n 与 k 奇偶性相反时，$P\{S_n = k\} = 0$.

三、直线上的随机游动

- 无限制随机游动：有无穷赌资的赌徒在 n 局后的输赢 $P\{S_n=k\}=\binom{n}{\frac{n+k}{2}}p^{\frac{n+k}{2}}q^{\frac{n-k}{2}}$

  - 赌徒在 n 局后赢的概率

① n 为偶数 $(n=2m)$ 时

$$\sum_{k=2}^{n}P\{S_n=k\}=\sum_{i=1}^{m}P\{S_n=2i\}=\sum_{i=1}^{m}\binom{2m}{m+i}p^{m+i}q^{m-i}=\sum_{l=m+1}^{2m}\binom{2m}{l}p^lq^{2m-l}=\sum_{l=m+1}^{2m}b(l;\ 2m,p)$$

$l\triangleq m+i$

② n 为奇数 $(n=2m+1)$ 时

$l\triangleq m+i+1$

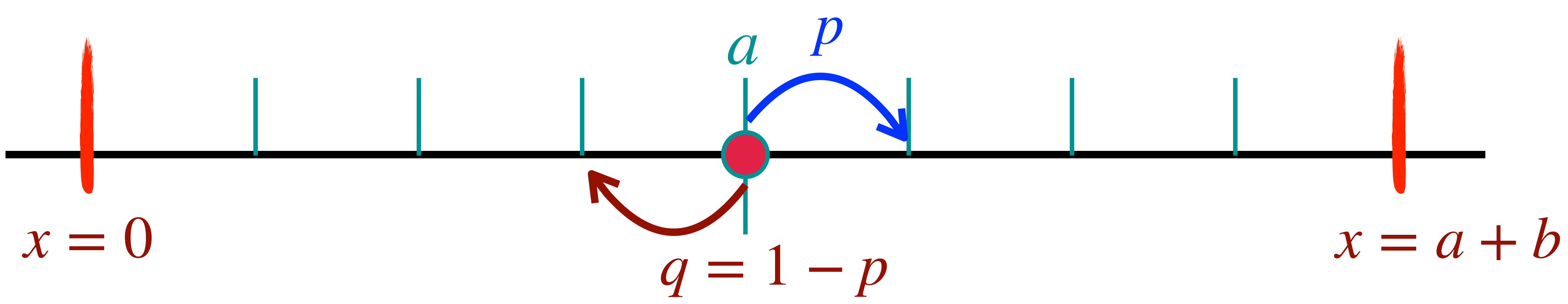
$$\sum_{k=1}^{n}P\{S_n=k\}=\sum_{i=1}^{m}P\{S_n=2i+1\}=\sum_{i=0}^{m}\binom{2m+1}{m+i+1}p^{m+i+1}q^{m-i}=\sum_{l=m+1}^{2m+1}\binom{2m+1}{l}p^lq^{2m+1-l}$$

$$=\sum_{l=m+1}^{2m+1}b(l;\ 2m+1,p)\ \Longrightarrow\ \sum_{k=1}^{n}P\{S_n=k\}=\sum_{l=\left[\frac{n}{2}\right]+1}^{n}b(l;\ n,p)\ \xrightarrow{p=q=0.5}\ \sum_{k=1}^{n}P\{S_n=k\}=\frac{1}{2}$$

三、直线上的随机游动

- 两端带有吸收壁的随机游动：有限赌资赌徒的输赢

  - 质点在 $t = 0$ 时位于 $x = a$，在 $x = 0$ 与 $x = a + b$ 处各有一个吸收壁.



  - 确定质点在 $x = 0$ 被吸收的概率？在 $x = a + b$ 处被吸收的概率？

  - 定义

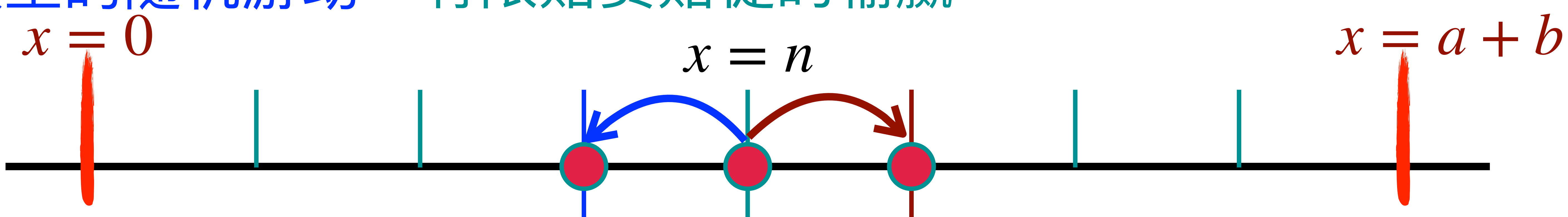
$$\begin{cases} q_0(n) = P\{\text{从 } x = n \text{ 出发被 } 0 \text{ 吸收}\} \\ p_{a+b}(n) = P\{\text{从 } x = n \text{ 出发被 } a+b \text{ 吸收}\} \end{cases}$$

$\Longrightarrow$ 边界条件： $q_0(0) = 1,\ q_0(a + b) = 0$

$\Longrightarrow$ 边界条件： $p_{a+b}(0) = 0,\ p_{a+b}(a + b) = 1$

三、直线上的随机游动

- 两端带有吸收壁的随机游动：有限赌资赌徒的输赢



全概率公式

$\{从\ x=n-1\ 出发被\ 0\ 吸收\}$ $\{从\ x=n+1\ 出发被\ 0\ 吸收\}$

下一步向左 $\{从\ x=n\ 出发被\ 0\ 吸收\}$ 下一步向右

$$P\{从\ x=n\ 出发被\ 0\ 吸收\}=p\cdot P\{从\ x=n+1\ 出发被\ 0\ 吸收\}+q\cdot P\{从\ x=n-1\ 出发被\ 0\ 吸收\}$$

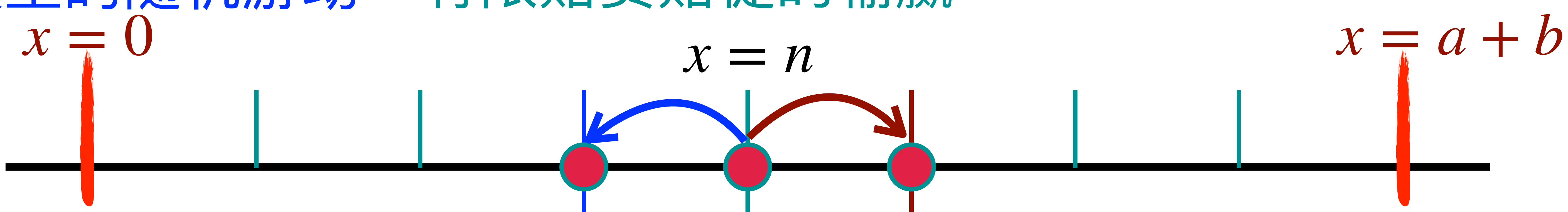
$\Longrightarrow$ 递推关系 $q_0(n)=p\cdot q_0(n+1)+q\cdot q_0(n-1)$

- 定义

$$\begin{cases} q_0(n)=P\{从\ x=n\ 出发被\ 0\ 吸收\} & \Longrightarrow\ 边界条件：\ q_0(0)=1,\ q_0(a+b)=0 \\ p_{a+b}(n)=P\{从\ x=n\ 出发被\ a+b\ 吸收\} & \Longrightarrow\ 边界条件：\ p_{a+b}(0)=0,\ p_{a+b}(a+b)=1 \end{cases}$$

三、直线上的随机游动

- 两端带有吸收壁的随机游动：有限赌资赌徒的输赢



$\{$从 $x = n-1$ 出发被 $a+b$ 吸收$\}$　$\{$从 $x = n+1$ 出发被 $a+b$ 吸收$\}$

全概率公式

下一步向左　$\{$从 $x = n$ 出发被 $a+b$ 吸收$\}$　下一步向右

$$P\{从\ x = n\ 出发被\ a+b\ 吸收\} = p \cdot P\{从\ x = n+1\ 出发被\ a+b\ 吸收\} + q \cdot P\{从\ x = n-1\ 出发被\ a+b\ 吸收\}$$

$\Longrightarrow$ 递推关系 $p_{a+b}(n) = p \cdot p_{a+b}(n+1) + q \cdot p_{a+b}(n-1)$

▸ 定义

$$\begin{cases} q_0(n) = P\{从\ x = n\ 出发被\ 0\ 吸收\} \\ p_{a+b}(n) = P\{从\ x = n\ 出发被\ a+b\ 吸收\} \end{cases}$$

$\Longrightarrow$ 边界条件： $q_0(0) = 1,\ q_0(a+b) = 0$

$\Longrightarrow$ 边界条件： $p_{a+b}(0) = 0,\ p_{a+b}(a+b) = 1$

三、直线上的随机游动

- 两端带有吸收壁的随机游动：有限赌资赌徒的输赢

  - 对于 $n=1,2,\ \dots,a+b-1$，我们来求解后者

$$(p+q)\cdot p_{a+b}(n)=p_{a+b}(n)$$

$$\begin{cases}\text{递推关系：} & p_{a+b}(n)=p\cdot p_{a+b}(n+1)+q\cdot p_{a+b}(n-1)\\ \text{边界条件：} & p_{a+b}(0)=0,\ p_{a+b}(a+b)=1\end{cases}$$

$$=(a+b)d \Longrightarrow d=\frac{1}{a+b}$$

$$c_n \triangleq p_{a+b}(n+1)-p_{a+b}(n)$$

$$p\cdot\left[p_{a+b}(n+1)-p_{a+b}(n)\right]=q\cdot\left[p_{a+b}(n)-p_{a+b}(n-1)\right]$$

$$r\triangleq\frac{q}{p}$$

$$\Longrightarrow\quad c_n=r\cdot c_{n-1}\ ,\quad n=1,2,\ \dots,a+b-1$$

① $r=1$ 时，$p=q=1/2$，即对称随机游动的场合， 此时有 $c_n=c_{n-1}=\cdots=c_1$，即

$$p_{a+b}(n+1)-p_{a+b}(n)=p_{a+b}(n)-p_{a+b}(n-1)=\cdots=p_{a+b}(2)-p_{a+b}(1)=p_{a+b}(1)-p_{a+b}(0)=p_{a+b}(1)\triangleq d$$

$$p_{a+b}(a)=\frac{a}{a+b}\ \xLeftarrow{n=a}\ p_{a+b}(n)=\frac{n}{a+b}\ \Longleftarrow\ p_{a+b}(n)=nd\ ,\ \cdots\ ,\ p_{a+b}(2)=2d\ ,\ p_{a+b}(1)=d$$

三、直线上的随机游动

- 两端带有吸收壁的随机游动：有限赌资赌徒的输赢

$$c_0 = \frac{1-r}{1-r^{a+b}}$$

▶ 对于 $n = 1, 2, \ldots, a+b-1$，我们来求解后者

$$\frac{1-r^{a+b}}{1-r} \cdot c_0 = 1$$

$$\begin{cases} \text{递推关系：} & p_{a+b}(n) = p \cdot p_{a+b}(n+1) + q \cdot p_{a+b}(n-1) \\ \text{边界条件：} & p_{a+b}(0) = 0,\ p_{a+b}(a+b) = 1 \end{cases}$$

$$c_n \triangleq p_{a+b}(n+1) - p_{a+b}(n)$$

$$p \cdot \left[p_{a+b}(n+1) - p_{a+b}(n)\right] = q \cdot \left[p_{a+b}(n) - p_{a+b}(n-1)\right]$$

$$r \triangleq \frac{q}{p}$$

$$\Longrightarrow \quad c_n = r \cdot c_{n-1}, \quad n = 1, 2, \ldots, a+b-1$$

② $r \neq 1$ 时，$p \neq q$，此时有 $c_1 = r \cdot c_0$, $c_2 = rc_1 = r^2 c_0$, $c_3 = rc_2 = r^3 c_0$, $\cdots$, $c_n = r^n c_0$

$$p_{a+b}(n) = p_{a+b}(n) - p_{a+b}(0) = \sum_{k=0}^{n-1} \left[p_{a+b}(k+1) - p_{a+b}(k)\right] = \sum_{k=0}^{n-1} c_k = \sum_{k=0}^{n-1} r^k c_0 = \frac{1-r^n}{1-r} \cdot c_0$$

$$\Longrightarrow \quad p_{a+b}(n) = \frac{1-r^n}{1-r^{a+b}} \xrightarrow{n=a} p_{a+b}(a) = \frac{1-r^a}{1-r^{a+b}} = \frac{1-\left(\frac{q}{p}\right)^a}{1-\left(\frac{q}{p}\right)^{a+b}}$$

三、直线上的随机游动

- 两端带有吸收壁的随机游动：有限赌资赌徒的输赢
  - 对于 $n = 1, 2, \ldots, a+b-1$，我们来求解后者

$$\begin{cases} \text{递推关系：} & p_{a+b}(n) = p \cdot p_{a+b}(n+1) + q \cdot p_{a+b}(n-1) \\ \text{边界条件：} & p_{a+b}(0) = 0,\ p_{a+b}(a+b) = 1 \end{cases}$$

① $r = 1$ 时，即 $p = q = 1/2$ 时，$p_{a+b}(n) = \dfrac{n}{a+b}$，$p_{a+b}(a) = \dfrac{a}{a+b}$

② $r \neq 1$ 时，$p \neq q$，$p_{a+b}(n) = \dfrac{1-\left(\frac{q}{p}\right)^{n}}{1-\left(\frac{q}{p}\right)^{a+b}}$，$p_{a+b}(a) = \dfrac{1-\left(\frac{q}{p}\right)^{a}}{1-\left(\frac{q}{p}\right)^{a+b}}$

三、直线上的随机游动

- 两端带有吸收壁的随机游动：有限赌资赌徒的输赢

  - 对于 $n = 1, 2, \ldots, a+b-1$，我们同理可以求解前者

$$\begin{cases} \text{递推关系：} & q_0(n) = p \cdot q_0(n+1) + q \cdot q_0(n-1) \\ \text{边界条件：} & q_0(0) = 1,\ q_0(a+b) = 0 \end{cases}$$

① $r = 1$ 时，即 $p = q = 1/2$ 时，$q_0(n) = \dfrac{a+b-n}{a+b}$，$q_0(a) = \dfrac{b}{a+b}$

② $r \neq 1$ 时，$p \neq q$，$q_0(n) = \dfrac{\left(\frac{q}{p}\right)^n - \left(\frac{q}{p}\right)^{a+b}}{1 - \left(\frac{q}{p}\right)^{a+b}}$，$q_0(a) = \dfrac{\left(\frac{q}{p}\right)^a - \left(\frac{q}{p}\right)^{a+b}}{1 - \left(\frac{q}{p}\right)^{a+b}}$

三、直线上的随机游动

- 两端带有吸收壁的随机游动：有限赌资赌徒的输赢
  - 对于 $n = 1, 2, \ldots, a+b-1$，我们看到

    ① $r = 1$ 时，即 $p = q = 1/2$ 时，$q_0(a) = \dfrac{b}{a+b}$，$p_{a+b}(a) = \dfrac{a}{a+b}$.

    甲有赌本 a，乙有赌本 b，公平赌博中甲输光的概率为 $q_0(a) = \dfrac{b}{a+b}$，甲赢光乙的概率为 $p_{a+b}(a) = \dfrac{a}{a+b}$，两者必居其一.

三、直线上的随机游动

- 两端带有吸收壁的随机游动：有限赌资赌徒的输赢

  - 对于 $n = 1, 2, \ldots, a+b-1$，我们看到

② $r \neq 1$ 时，即 $p \neq q$ 时，$q_0(a) = \dfrac{\left(\frac{q}{p}\right)^a - \left(\frac{q}{p}\right)^{a+b}}{1-\left(\frac{q}{p}\right)^{a+b}}$，$p_{a+b}(a) = \dfrac{1-\left(\frac{q}{p}\right)^a}{1-\left(\frac{q}{p}\right)^{a+b}}$.

甲有赌本 a，乙有赌本 b，不公平赌博中甲输光的概率为 $q_0(a) = \dfrac{\left(\frac{q}{p}\right)^a - \left(\frac{q}{p}\right)^{a+b}}{1-\left(\frac{q}{p}\right)^{a+b}}$，甲

赢光乙的概率为 $p_{a+b}(a) = \dfrac{1-\left(\frac{q}{p}\right)^a}{1-\left(\frac{q}{p}\right)^{a+b}}$，两者必居其一.

四、推广的 Bernoulli 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个结果的情形.
  - 设每次试验的可能结果为 $A_1, A_2, \ldots, A_r$，且 $\begin{cases} P(A_i) = p_i, & i = 1, 2, \ldots, r \\ p_i \geq 0, & \text{且} \quad p_1 + p_2 + \cdots + p_r = 1 \end{cases}$

$$\Longrightarrow \quad P(A_1 = k_1, A_2 = k_2, \ldots, A_r = k_r) = \frac{n!}{k_1!\, k_2! \cdots k_r!} p_1^{k_1} p_2^{k_2} \cdot \cdots \cdot p_r^{k_r}, \quad k_i \geq 0 \text{ 且 } k_1 + k_2 + \cdots + k_r = n$$

  - 多项分布：$(p_1 + p_2 + \cdots + p_r)^n$ 展开式的一般项.

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

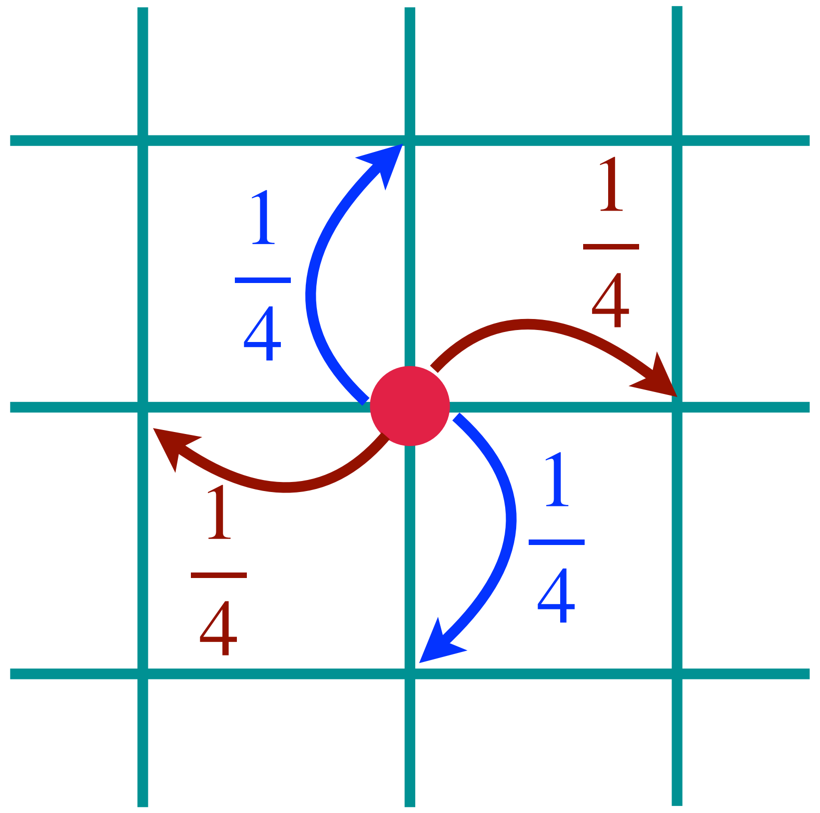
  - $r = 2$ 时即为 Bernoulli 试验，对应于二项分布. 二项分布很多性质可平行推广到多项分布.

四、推广的 Bernoulli 试验与多项分布

- 例：平面上的随机游动
  - 一质点从平面上某一点出发，等可能的向上、下、左、右四个方向移动，每次的移动距离为 1，求经过 $2n$ 次移动后回到出发点的概率.

定义 $\begin{cases} A_1 : \text{质点向上移动一格} \\ A_2 : \text{质点向下移动一格} \\ A_3 : \text{质点向左移动一格} \\ A_4 : \text{质点向右移动一格} \end{cases} \Longrightarrow \begin{cases} P(A_1) = \frac{1}{4} \\ P(A_2) = \frac{1}{4} \\ P(A_3) = \frac{1}{4} \\ P(A_4) = \frac{1}{4} \end{cases}$

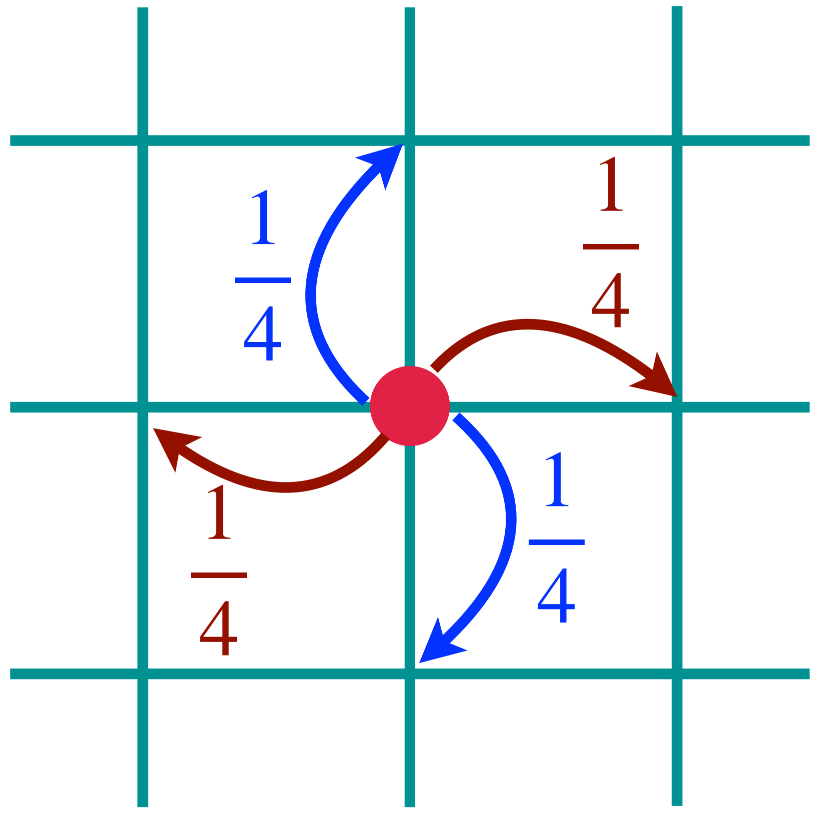
多项分布



四、推广的 Bernoulli 试验与多项分布

- 例：平面上的随机游动

  - 一质点从平面上某一点出发，等可能的向上、下、左、右四个方向移动，每次的移动距离为 1，求经过 $2n$ 次移动后回到出发点的概率.



  - 若要在 $2n$ 次移动后回到原来的出发点，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也应该相等.

  - 总移动次数为 $2n$，故所求概率 $= \sum\limits_{k+m=n} \dfrac{(2n)!}{k!\cdot k!\cdot m!\cdot m!}\left(\dfrac{1}{4}\right)^{2n} = \sum\limits_{k=0}^{n} \dfrac{(2n)!}{(k!)^2\cdot\left[(n-k)!\right]^2}\left(\dfrac{1}{4}\right)^{2n}$

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

四、推广的 Bernoulli 试验与多项分布

- 例：平面上的随机游动

```
random_walk_2d = function(x0 = 0, y0 = 0, N = 100) {
  rand = sample.int(4, N, replace=TRUE)
  x = x_temp = x0
  y = y_temp = y0
  for (i in 1:N) {
    if (rand[i] == 1L)
      x_temp = x_temp + 1L
    else if (rand[i] == 2L)
      x_temp = x_temp - 1L
    else if (rand[i] == 3L)
      y_temp = y_temp + 1L
    else
      y_temp = y_temp - 1L
    x = c(x, x_temp)
    y = c(y, y_temp)
  }
  result = data.frame(step = 0:N, x = x, y = y)
  return(result)
}
```

```
n = 500
result = random_walk_2d(N = n)
x = result$x
y = result$y
max_xy = max(abs(range(c(x, y))))
plot(x, y, type = "o", asp = 1, cex = 1, lwd = 2, pch = 16,
    xlab = 'x', ylab = 'y', cex.lab = 1.5, cex.axis = 1.5,
    xlim = c(-max_xy-1, max_xy+1), ylim = c(-max_xy-1, max_xy+1), col = "cyan4")
points(0, 0, pch = 21, col='red', bg = 'blue', cex = 2)
points(x[n], y[n], pch = 21, col = 'blue', bg = 'red', cex = 2)
text(x[n], -max_xy-1, paste0('(', x[n], ', ', y[n], ')'), pos = 3, col = 'red', cex = 1.5)
abline(h = seq(-max_xy-1, max_xy+1), col = "gray50", lty = 3)
abline(v = seq(-max_xy-1, max_xy+1), col = "gray50", lty = 3)
```