

概 率 论

Probability

肖磊, 2024年11月26日

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: (Ω, \mathcal{F}, P)
 - 难点和要点: 事件域 \mathcal{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: $P(A|B) = \frac{P(AB)}{P(B)}$.

- 乘法公式: $P(A_1A_2\cdots A_n) = P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_1A_2) \cdot \cdots \cdot P(A_n|A_1A_2\cdots A_{n-1})$

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

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

$$\left. \begin{array}{l} A_i \cap A_j = \Phi \quad (i \neq j) \\ \bigcup_{i=1}^{\infty} A_i = \Omega \end{array} \right\}$$

已学知识点

- 第二章 条件概率与统计独立性

- ▶ 事件独立性：两个事件 $P(AB) = P(A) \cdot P(B)$. 三个事件
$$\begin{cases} P(AB) = P(A) \cdot P(B) \\ P(AC) = P(A) \cdot P(C) \\ P(BC) = P(B) \cdot P(C) \\ P(ABC) = P(A) \cdot P(B) \cdot P(C) \end{cases}$$

- ▶ 试验独立性：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.

- ▶ Bernoulli 试验 E : 概率空间 (Ω, \mathcal{F}, P) , 其中

$$A \subset \Omega, \quad \mathcal{F} = \{\Phi, A, \bar{A}, \Omega\}, \quad P(A) = p, \quad P(\bar{A}) = q, \quad (p > 0, q > 0, p + q = 1)$$

- Bernoulli 分布

- 二项分布

- 几何分布

- Pascal 分布

- 多项分布

- Poisson 分布

已学知识点

● 第三章 随机变量与分布函数

▶ 随机变量 (r.v.) 与分布函数 (c.d.f.):

- 随机变量 ξ 是概率空间 (Ω, \mathcal{F}, P) 中定义域为 Ω 、取值于 \mathbb{R} 的一个单值 Borel 函数.
- 分布函数 $F(x) = P\{\xi(\omega) \leq x\}$, $-\infty < x < \infty$ 是单调不降、取值 $[0, 1]$ 的右连续函数. 它完整描述了随机变量, 是研究的主要对象.
- 随机变量依取值 $\begin{cases} \text{离散型:} & \text{分布律 (mass function)、分布列} \\ \text{连续型:} & \text{概率密度 (p.d.f.)} \end{cases}$
- 主要分布: $\begin{cases} \text{离散型:} & \text{Bernoulli, binomial, Poisson, hyper-geometric, geometric} \\ \text{连续型:} & \text{uniform, exponential, normal, } \Gamma \end{cases}$
- 概率计算: $\begin{cases} \text{离散型:} & P\{x \in D\} = \sum_{x_i \in D} p_i, \quad P\{(x, y) \in D\} = \sum_{(x_i, y_j) \in D} p_{ij} \\ \text{连续型:} & P\{x \in D\} = \int_D p(x) dx, \quad P\{(x, y) \in D\} = \iint_D p(x, y) dx dy \end{cases}$

已学知识点

● 第三章 随机变量与分布函数

▶ 随机向量，随机变量的独立性：

○ 随机向量即多元随机变量

| | | |
|---|-------|-------------------|
| { | 联合分布： | 联合分布函数、联合分布律、联合密度 |
| | 边际分布： | 边际分布函数、边际分布律、边际密度 |
| | 条件分布： | 条件分布函数、条件分布律、条件密度 |
| | 独立性： | 与事件独立性几乎完全相同 |

○ 主要分布：

| | | |
|---|------|---------------|
| { | 离散型： | 多项分布、多元超几何分布 |
| | 连续型： | 二元均匀分布、二元正态分布 |

$$(\xi, \eta) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho) \sim N\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_1\sigma_2\rho & \sigma_2^2 \end{pmatrix}\right) \implies \left\{ \begin{array}{l} \xi \sim N(\mu_1, \sigma_1^2), \quad \eta \sim N(\mu_2, \sigma_2^2) \\ (\eta | \xi = x) \sim N\left(\mu_2 + \rho \frac{\sigma_2}{\sigma_1} (x - \mu_1), \sigma_2^2 (1 - \rho^2)\right) \\ (\xi | \eta = y) \sim N\left(\mu_1 + \rho \frac{\sigma_1}{\sigma_2} (y - \mu_2), \sigma_1^2 (1 - \rho^2)\right) \\ \xi \text{ 与 } \eta \text{ 相互独立} \iff \rho = 0 \end{array} \right.$$

已学知识点

● 第三章 随机变量与分布函数

▶ 随机变量的函数及其分布:

○ 随机变量的函数什么情况下还是随机变量

○ 离散型易: $\begin{cases} \eta = g(\xi) : & \text{对应法} \\ \zeta = g(\xi, \eta) : & \text{表上作业法, 独立情形和的卷积公式} \end{cases}$

○ 连续型难: $\begin{cases} \eta = g(\xi) : \text{直接法 } F_{\eta}(y) = P\{\eta < y\} = \int_{g(x) < y} p_{\xi}(x) dx \\ \zeta = g(\xi, \eta) : \text{直接法 } F_{\zeta}(z) = P\{\zeta < z\} = \iint_{g(x, y) < z} p(x, y) dx dy \\ \text{和 (卷积公式)、差、商、积、max, min 的公式} \end{cases}$

$$\begin{cases} \zeta_1 = g_1(\xi, \eta) \\ \zeta_2 = g_2(\xi, \eta) \end{cases} : \text{变换法} \quad \begin{cases} u = g_1(x, y) \\ v = g_2(x, y) \end{cases} \implies \begin{cases} x = h_1^{-1}(u, v) \\ y = h_2^{-1}(u, v) \end{cases} \implies J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix}$$

$$q_{\zeta_1, \zeta_2}(u, v) = p(h_1^{-1}(u, v), h_2^{-1}(u, v)) \cdot |J|$$

已学知识点

● 第四章 数字特征与特征函数

▶ 数学期望

○ 定义: $E(\xi) = \begin{cases} \sum_i x_i \cdot p_i, & \text{离散型} \\ \int_{-\infty}^{\infty} x \cdot p_{\xi}(x) dx, & \text{连续型} \end{cases}$, 绝对收敛.

○ 随机变量函数的数学期望: $E[g(\xi)] = \begin{cases} \sum_i g(x_i) \cdot p_i, & \text{离散型} \\ \int_{-\infty}^{\infty} g(x) \cdot p_{\xi}(x) dx, & \text{连续型} \end{cases}$

$E[g(\xi, \eta)] = \begin{cases} \sum_i \sum_j g(x_i, y_j) \cdot p_{ij}, & \text{离散型} \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) \cdot p_{(\xi, \eta)}(x, y) dx dy, & \text{连续型} \end{cases}$

数学期望的性质:

- ① $E(c) = c$
- ② $a \leq \xi \leq b \implies a \leq E(\xi) \leq b$
- ③ $\xi \leq \eta \implies E(\xi) \leq E(\eta)$
- ④ $E(a\xi + b\eta) = a \cdot E(\xi) + b \cdot E(\eta)$
- ⑤ ξ, η 独立 $\implies E(\xi\eta) = E(\xi) \cdot E(\eta)$

已学知识点

● 第四章 数字特征与特征函数

▶ **方差**: $D(\xi) = E\left\{\left[\xi - E(\xi)\right]^2\right\}$ 标准差: $\sqrt{D(\xi)}$ 计算公式: $D(\xi) = E(\xi^2) - [E(\xi)]^2$

▶ **Chebyshev 不等式**:

$$P\left\{\left|\xi - E(\xi)\right| \geq \varepsilon\right\} \leq \frac{D(\xi)}{\varepsilon^2} \iff P\left\{\left|\xi - E(\xi)\right| < \varepsilon\right\} \geq 1 - \frac{D(\xi)}{\varepsilon^2}$$

▶ **协方差**:

$$\text{cov}(\xi, \eta) \triangleq E\left\{\left[\xi - E(\xi)\right]\left[\eta - E(\eta)\right]\right\} = E(\xi\eta) - E(\xi) \cdot E(\eta)$$

▶ **和的方差**: $D(X \pm Y) = D(X) + D(Y) \pm 2\text{cov}(X, Y)$

方差的性质:

- ① 设 C 是常数, 则 $D(C) = 0$.
- ② 设 ξ 是随机变量, c 为常数, 则 $D(\xi + c) = D(\xi)$.
- ③ 设 ξ 是随机变量, c 为常数, 则 $D(c\xi) = c^2D(\xi)$.
- ④ 若 $c \neq E(\xi)$, 则 $D(\xi) < E[(\xi - c)^2]$.
- ⑤ $D(\xi) = 0 \iff P\{\xi = C\} = 1$.
- ⑥ 若 ξ 与 η 独立, 则 $D(\xi \pm \eta) = D(\xi) + D(\eta)$.

协方差的性质:

- ① $\text{cov}(X, X) = D(X)$.
- ② $\text{cov}(X, Y) = \text{cov}(Y, X)$
- ③ $\text{cov}(aX, bY) = ab \cdot \text{cov}(X, Y)$, a, b 为常数.
- ④ $\text{cov}(X_1 + X_2, Y) = \text{cov}(X_1, Y) + \text{cov}(X_2, Y)$.
- ⑤ 若 X, Y 相互独立, 则 $\text{cov}(X, Y) = 0$.
- ⑥ 若 C 为常数, 则 $\text{cov}(X, C) = 0$.

已学知识点

● 第四章 数字特征与特征函数

▶ 常见分布的数学期望与方差:

| 分布 | 符号表示 | 数学期望 | 方差 |
|--------------|--------------------------------|------------------------------|--------------------------------|
| 二项分布 | $\xi \sim b(n, p)$ | $E(\xi) = np$ | $D(\xi) = np(1-p)$ |
| Bernoulli 分布 | $\xi \sim b(1, p)$ | $E(\xi) = p$ | $D(\xi) = p(1-p)$ |
| Poisson 分布 | $\xi \sim P(\lambda)$ | $E(\xi) = \lambda$ | $D(\xi) = \lambda$ |
| 均匀分布 | $\xi \sim U[a, b]$ | $E(\xi) = \frac{a+b}{2}$ | $D(\xi) = \frac{(b-a)^2}{12}$ |
| Gamma 分布 | $X \sim \Gamma(r, \lambda)$ | $E(\xi) = \frac{r}{\lambda}$ | $D(\xi) = \frac{r}{\lambda^2}$ |
| 指数分布 | $\xi \sim \text{Exp}(\lambda)$ | $E(\xi) = \frac{1}{\lambda}$ | $D(\xi) = \frac{1}{\lambda^2}$ |
| χ^2 分布 | $X \sim \chi_n^2$ | $E(\xi) = n$ | $D(\xi) = 2n$ |
| 正态分布 | $\xi \sim N(\mu, \sigma^2)$ | $E(\xi) = \mu$ | $D(\xi) = \sigma^2$ |

已学知识点

● 第四章 数字特征与特征函数

▶ **相关系数:** $\rho_{\xi\eta} \triangleq \frac{\text{cov}(\xi, \eta)}{\sqrt{D(\xi)}\sqrt{D(\eta)}}$

▶ **Cauchy-Schwarz 不等式:** $|E(\xi\eta)|^2 \leq E(\xi^2) \cdot E(\eta^2)$

▶ **矩:** $\left\{ \begin{array}{l} k \text{ 阶原点矩: } m_k = E(\xi^k), \quad k = 1, 2, 3, \dots \\ k \text{ 阶中心矩: } c_k = E\{[\xi - E(\xi)]^k\}, \quad k = 2, 3, \dots \\ k+l \text{ 阶混合原点矩: } E(\xi^k\eta^l) \\ k+l \text{ 阶混合中心矩: } E\{[\xi - E(\xi)]^k[\eta - E(\eta)]^l\} \end{array} \right.$

▶ **定理:** 中心矩和原点矩可以相互表示. $\left\{ \begin{array}{l} c_k = \sum_{i=0}^k \binom{k}{i} m_i (-m_1)^{k-i} \\ m_k = \sum_{i=0}^k \binom{k}{i} c_{k-i} m_1^i \end{array} \right.$

相关系数的性质:

- ① $|\rho_{\xi\eta}| \leq 1.$
- ② 等价命题: $\text{cov}(\xi, \eta) = 0$; ξ 与 η 不相关;
 $E(\xi \cdot \eta) = E(\xi) \cdot E(\eta);$
 $D(\xi + \eta) = D(\xi) + D(\eta).$
- ③ 若随机变量 ξ 与 η 独立, 则 ξ 与 η 不相关.
- ④ 二元正态分布的独立性和不相关性是等价的.
- ⑤ 二值随机变量的不相关性与独立性等价.

已学知识点

● 第四章 数字特征与特征函数

▶ **条件期望：**

$$\begin{cases} E(\eta | \xi = x_i) = \sum_j y_j \cdot P\{\eta = y_j | \xi = x_i\} = \sum_j y_j \cdot \frac{p(x_i, y_j)}{p_1(x_i)} \\ E(\xi | \eta = y_j) = \sum_i x_i \cdot P\{\xi = x_i | \eta = y_j\} = \sum_i x_i \cdot \frac{p(x_i, y_j)}{p_2(y_j)} \end{cases}$$

| | | | | | | |
|-----------------------|---------------|---------------|----------|---------------|----------|--------------|
| $\xi \backslash \eta$ | y_1 | y_2 | \dots | y_j | \dots | $p_1(\cdot)$ |
| x_1 | $p(x_1, y_1)$ | $p(x_1, y_2)$ | \dots | $p(x_1, y_j)$ | \dots | $p_1(x_1)$ |
| x_2 | $p(x_2, y_1)$ | $p(x_2, y_2)$ | \dots | $p(x_2, y_j)$ | \dots | $p_1(x_2)$ |
| \vdots | \vdots | \vdots | \ddots | \vdots | \vdots | \vdots |
| x_i | $p(x_i, y_1)$ | $p(x_i, y_2)$ | \dots | $p(x_i, y_j)$ | \dots | $p_1(x_i)$ |
| \vdots | \vdots | \vdots | \ddots | \vdots | \vdots | \vdots |
| $p_2(\cdot)$ | $p_2(y_1)$ | $p_2(y_2)$ | \dots | $p_2(y_j)$ | \dots | 1 |

▶ **连续型随机变量的条件期望：** $E(\eta | \xi = x) = \int_{-\infty}^{+\infty} y \cdot p(y | x) dy$

○ η 关于 ξ 的条件期望 $E(\eta | \xi)$ 是一个随机变量，它在 $E(\eta | \xi = x)$ 处的密度函数为 $p_\xi(x)$.

▶ **重期望公式：** $E(\xi) = E\left[E(\xi | \eta)\right]$.

已学知识点

● 第四章 数字特征与特征函数

▶ **复随机变量**: 若 ξ 与 η 是概率空间 (Ω, \mathcal{F}, P) 上的实值随机变量, 称 $\zeta = \xi + i\eta$ 为复随机变量.

○ 可平行定义或建立一系列结果: 如 $E(\zeta) = E(\xi) + iE(\eta)$ 、独立性等.

○ Euler 公式: $e^{ix} = \cos x + i \sin x$.

▶ **特征函数**: $f_{\xi}(t) = E(e^{it\xi}) = \int_{-\infty}^{+\infty} e^{itx} dF_{\xi}(x) = \begin{cases} \sum_{j=1}^{\infty} p_j \cdot e^{itx_j}, & \text{离散型} \\ \int_{-\infty}^{+\infty} e^{itx} \cdot p_{\xi}(x) dx, & \text{连续型} \end{cases}$

特征函数的性质:

- ① 特征函数 $f(t)$ 有如下性质: $|f(t)| \leq f(0) = 1$, $f(-t) = \overline{f(t)}$.
- ② 特征函数 $f(t)$ 在 $(-\infty, +\infty)$ 一致连续.
- ③ 对于任意的正整数 n 以及任意实数 t_1, t_2, \dots, t_n 、任意复数 $\lambda_1, \lambda_2, \dots, \lambda_n$, 均有 $\sum_{k=1}^n \sum_{j=1}^n f(t_k - t_j) \lambda_k \bar{\lambda}_j \geq 0$
- ④ 两个相互独立的随机变量之和的特征函数等于它们各自特征函数的乘积.
- ⑤ 设随机变量 ξ 的 n 阶矩存在, 则它的特征函数可微分 n 次, 且当 $k \leq n$ 时: $f^{(k)}(0) = i^k E(\xi^k)$.
- ⑥ 设 $\eta = a\xi + b$, 这里 a, b 为常数, 则 $f_{\eta}(t) = e^{ibt} \cdot f_{\xi}(at)$.

常见分布的特征函数:

- ① 退化分布 $I_c(x)$: $f(t) = e^{itc}$
- ② 二项分布 $b(n, p)$: $f(t) = (pe^{it} + 1 - p)^n$
- ③ Poisson 分布: $f(t) = e^{\lambda(e^{it} - 1)}$
- ④ Γ 分布: $f(t) = \left(1 - \frac{it}{\lambda}\right)^{-r}$
- ⑤ 正态分布: $f(t) = e^{i\mu t - \frac{1}{2}\sigma^2 t^2}$

已学知识点

● 第四章 数字特征与特征函数

▶ 特征函数和分布函数相互唯一确定.

○ 分布函数可以确定特征函数: $f_{\xi}(t) \triangleq E(e^{it\xi}) = \int_{-\infty}^{+\infty} e^{itx} dF_{\xi}(x) = \begin{cases} \sum_{j=1}^{\infty} p_j \cdot e^{itx_j}, & \text{离散型} \\ \int_{-\infty}^{+\infty} e^{itx} \cdot p_{\xi}(x) dx, & \text{连续型} \end{cases}$

○ (唯一性定理) 分布函数由特征函数唯一确定.

○ 若特征函数 $f(t)$ 绝对可积, 则相应分布函数 $F(x)$ 的导数存在并连续, 并且 $F'(x) = p(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} f(t) dt$.

▶ 分布函数的再生性: 两个具有同一类型分布的独立随机变量之和的分布仍是这种类型的分布, 且对应的参数等于两个随机变量相应参数之和.

○ 二项分布、Poisson 分布、正态分布、 Γ 分布具有再生性.

○ 分解问题: 对于正态分布、Poisson 分布, 分解问题成立.

第 5 章 极限定理

目 录

5.1 Bernoulli 试验场合的极限定理

5.2 收敛性

5.3 独立同分布场合的极限定理

5.1 Bernoulli 试验场合的极限定理

- 一、问题的提出
- 二、大数定律
- 三、De Moivre - Laplace 极限定理
- 四、De Moivre - Laplace 极限定理的应用

一、问题的提出

- **Bernoulli 试验**: 随机试验 E 只有两个结果: A 和 \bar{A} , 其中

$$P(A) = p, P(\bar{A}) = q, \quad (p > 0, q > 0, p + q = 1)$$

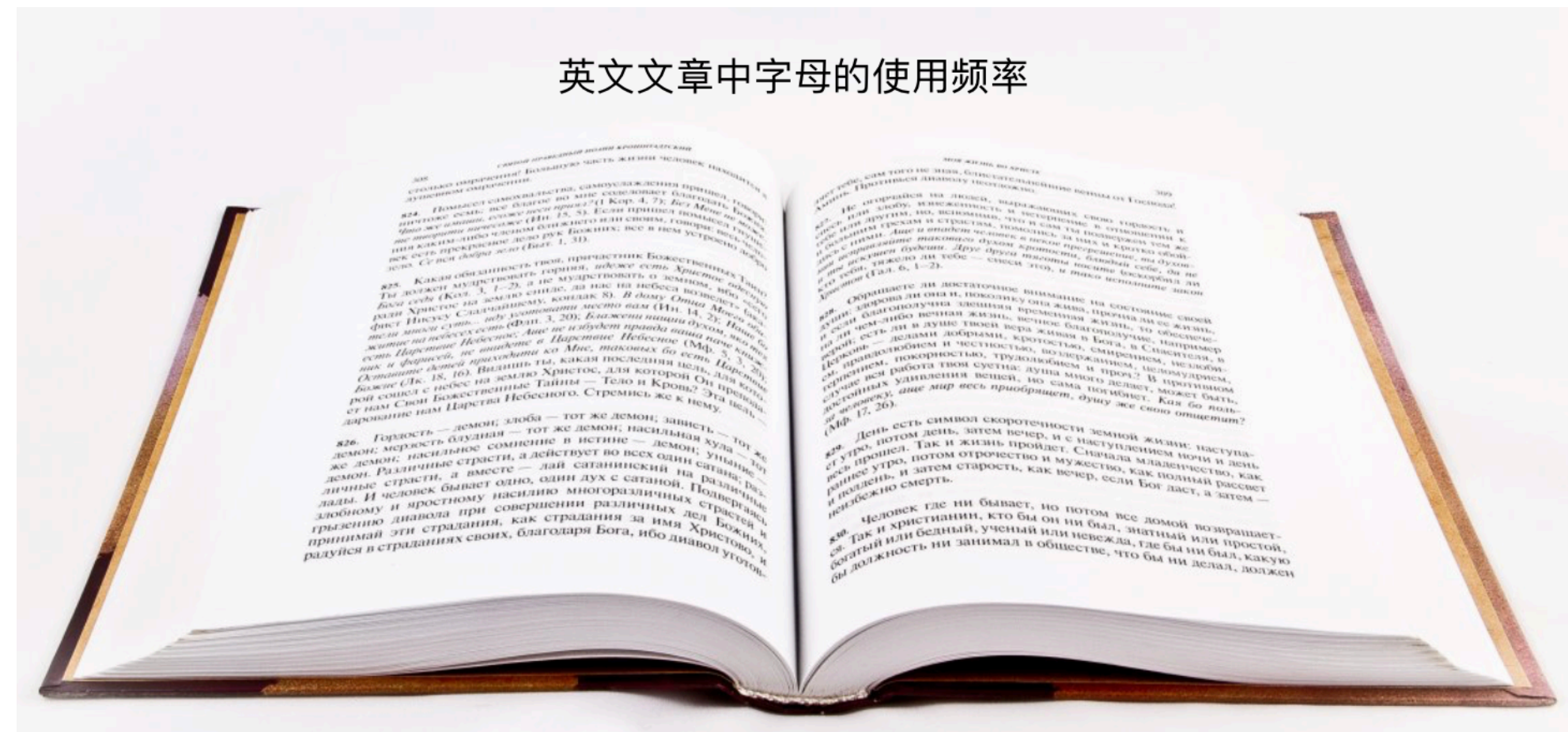
- **n 重 Bernoulli 试验**: n 次独立重复的 Bernoulli 试验, 记作 E_n



掷硬币



生产灯泡的废品率



英文文章中字母的使用频率

- **频率的稳定性**: n 次独立重复 Bernoulli 试验中事件 A 出现的次数为 μ_n , 则事件 A 出现的频率

$\frac{\mu_n}{n}$ 当 n 增大时接近于某个固定的常数

研究 $\lim_{n \rightarrow \infty} \frac{\mu_n}{n}$ 的极限行为

一、问题的提出

- 频率 $\frac{\mu_n}{n}$ 是随机变量，研究它的极限该如何进行分析？

- ▶ 当 n 很大时，频率 $\frac{\mu_n}{n}$ 与概率 p 有较大差异的概率很小. 写成数学的语言就是要证明

$$\lim_{n \rightarrow \infty} P \left\{ \left| \frac{\mu_n}{n} - p \right| \geq \varepsilon \right\} = 0, \text{ 或者 } \lim_{n \rightarrow \infty} P \left\{ \left| \frac{\mu_n}{n} - p \right| < \varepsilon \right\} = 1$$

$$\xi_k \triangleq \begin{cases} 1, & \text{第 } k \text{ 次试验出现事件 } A \\ 0, & \text{第 } k \text{ 次试验不出现事件 } A \end{cases}$$

$$\Rightarrow \mu_n = \xi_1 + \xi_2 + \cdots + \xi_n$$

$\xi_1, \xi_2, \dots, \xi_n$ 相互独立

$$\lim_{n \rightarrow \infty} P \left\{ \left| \frac{1}{n} \sum_{i=1}^n \xi_i - p \right| \geq \varepsilon \right\} = 0, \text{ 或者 } \lim_{n \rightarrow \infty} P \left\{ \left| \frac{1}{n} \sum_{i=1}^n \xi_i - p \right| < \varepsilon \right\} = 1$$

一、问题的提出

- 频率 $\frac{\mu_n}{n}$ 是随机变量，研究它的极限该如何进行分析？
 - ▶ 当 n 很大时，频率 $\frac{\mu_n}{n}$ 与概率 p 有较大差异的概率很小。

定义：设 $\xi_1, \xi_2, \dots, \xi_n, \dots$ 是随机变量序列，令

$$\eta_n = \frac{\xi_1 + \xi_2 + \dots + \xi_n}{n}$$

如果存在常数 a ，使得 $\forall \varepsilon > 0$ ，恒有

$$\lim_{n \rightarrow \infty} P \left\{ \left| \eta_n - a \right| < \varepsilon \right\} = 1$$

则称随机变量序列 $\{\eta_n\}$ 服从**大数定律** (law of large numbers)。

一、问题的提出

- 频率 $\frac{\mu_n}{n}$ 是随机变量，研究它的极限该如何进行分析？

| | | |
|---------------------|-----|---------|
| ξ_i | 1 | 0 |
| $P \{ \xi_i = x \}$ | p | $1 - p$ |

- ▶ 当 n 很大时，考虑频率 $\frac{\mu_n}{n} = \frac{1}{n} \sum_{i=1}^n \xi_i$ 的概率分布. 通常研究其标准化的随机变量

$$\zeta_n = \frac{\mu_n - np}{\sqrt{np(1-p)}} \sim N(0, 1), \quad n \longrightarrow \infty$$

可以证明 \longrightarrow

$$\lim_{n \rightarrow \infty} P \{ \zeta_n < x \} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt$$

一、问题的提出

- 频率 $\frac{\mu_n}{n}$ 是随机变量，研究它的极限该如何进行分析？

- ▶ 当 n 很大时，考虑频率 $\frac{\mu_n}{n} = \frac{1}{n} \sum_{i=1}^n \xi_i$ 的概率分布.

定义： 设 $\xi_1, \xi_2, \dots, \xi_n, \dots$ 是相互独立的随机变量序列，且 $E(\xi_i)$ 与 $D(\xi_i)$ 存在，令

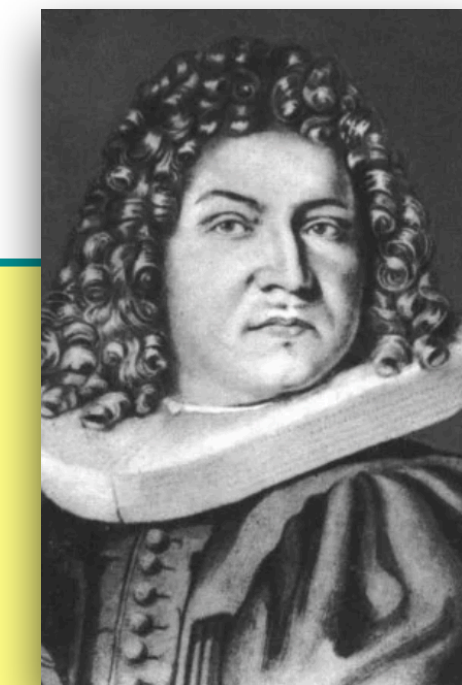
$$\zeta_n = \frac{\sum_{i=1}^n \xi_i - \sum_{i=1}^n E(\xi_i)}{\sqrt{\sum_{i=1}^n D(\xi_i)}}$$

如果

$$\lim_{n \rightarrow \infty} P \{ \zeta_n < x \} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt$$

则称随机变量序列 $\{\zeta_n\}$ 服从 **中心极限定理** (central limit theorem).

二、大数定律



定理: (Bernoulli 大数定律)

设 μ_n 是 n 重 Bernoulli 试验中事件 A 发生的次数, $p = P\{A\}$, 则 $\forall \varepsilon > 0$,

$$\lim_{n \rightarrow \infty} P \left\{ \left| \frac{\mu_n}{n} - p \right| < \varepsilon \right\} = 1 \quad \text{或} \quad \lim_{n \rightarrow \infty} P \left\{ \left| \frac{\mu_n}{n} - p \right| \geq \varepsilon \right\} = 0$$

Chebyshev 不等式: 设 ξ 为具有有限均值 $E(\xi)$ 与有限方差 $D(\xi)$ 的随机变量, 则 $\forall \varepsilon > 0$ 有

$$P \left\{ |\xi - E(\xi)| \geq \varepsilon \right\} \leq \frac{D(\xi)}{\varepsilon^2}$$

$\mu_n \sim b(n, p)$

$$\Rightarrow P \left\{ \left| \frac{\mu_n}{n} - p \right| \geq \varepsilon \right\} \leq \frac{1}{\varepsilon^2} D \left(\frac{\mu_n}{n} \right) = \frac{1}{\varepsilon^2} \cdot \frac{p(1-p)}{n} \leq \frac{1}{4n\varepsilon^2} \rightarrow 0 \quad (n \rightarrow \infty)$$

二、大数定律

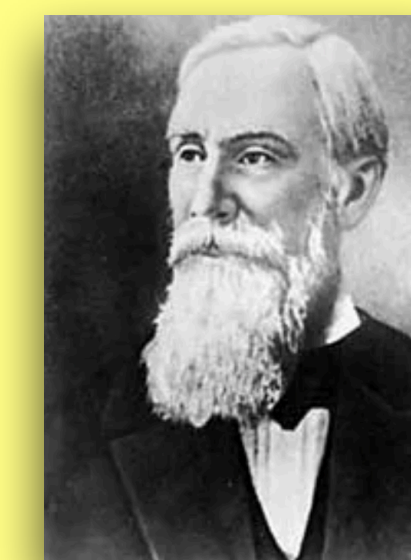
定理: (Chebyshev 大数定律)

设 $\xi_1, \xi_2, \dots, \xi_n, \dots$ 是两两不相关的随机变量序列, 每一随机变量都有有限的方差, 并且它们有公共的上界

$$D(\xi_1) \leq C, D(\xi_2) \leq C, \dots, D(\xi_n) \leq C, \dots$$

则对任意的 $\varepsilon > 0$, 皆有

$$\lim_{n \rightarrow \infty} P \left\{ \left| \frac{1}{n} \sum_{k=1}^n \xi_k - \frac{1}{n} \sum_{k=1}^n E(\xi_k) \right| < \varepsilon \right\} = 1$$



Пафну́тий Льво́вич Чебышёв

二、大数定律

$$E\left(\frac{1}{n}\sum_{k=1}^n \xi_k\right) = \frac{1}{n}\sum_{k=1}^n E(\xi_k)$$

$$D\left(\frac{1}{n}\sum_{k=1}^n \xi_k\right) = \frac{1}{n^2}\sum_{k=1}^n D(\xi_k) \leq \frac{1}{n^2}\sum_{k=1}^n C = \frac{C}{n}$$

$$0 \leq P\left\{\left|\frac{1}{n}\sum_{k=1}^n \xi_k - \frac{1}{n}\sum_{k=1}^n E(\xi_k)\right| \geq \varepsilon\right\} \leq \frac{D\left(\frac{1}{n}\sum_{k=1}^n \xi_k\right)}{\varepsilon^2} = \frac{C}{n\varepsilon^2} \longrightarrow 0 \quad (n \rightarrow \infty)$$

$$\implies \lim_{n \rightarrow \infty} P\left\{\left|\frac{1}{n}\sum_{k=1}^n \xi_k - \frac{1}{n}\sum_{k=1}^n E(\xi_k)\right| \geq \varepsilon\right\} = 0$$

定理: (Chebyshev 大数定律)

设 $\xi_1, \xi_2, \dots, \xi_n, \dots$ 是两两不相关的随机变量序列, 每一随机变量都有有限的方差, 并且它们有公共的上界

$$D(\xi_1) \leq C, D(\xi_2) \leq C, \dots, D(\xi_n) \leq C, \dots$$

则对任意的 $\varepsilon > 0$, 皆有

$$\lim_{n \rightarrow \infty} P\left\{\left|\frac{1}{n}\sum_{k=1}^n \xi_k - \frac{1}{n}\sum_{k=1}^n E(\xi_k)\right| < \varepsilon\right\} = 1$$



Пафнутий Львович Чебышев

Chebyshev 不等式: 设 ξ 为具有有限均值 $E(\xi)$ 与有限方差 $D(\xi)$ 的随机变量, 则 $\forall \varepsilon > 0$ 有

$$P\left\{|\xi - E(\xi)| \geq \varepsilon\right\} \leq \frac{D(\xi)}{\varepsilon^2}$$

二、大数定律

- Markov 注意到在 Chebyshev 的论证中, 只要

$$\frac{1}{n^2} \cdot D \left(\sum_{k=1}^n \xi_k \right) \longrightarrow 0$$

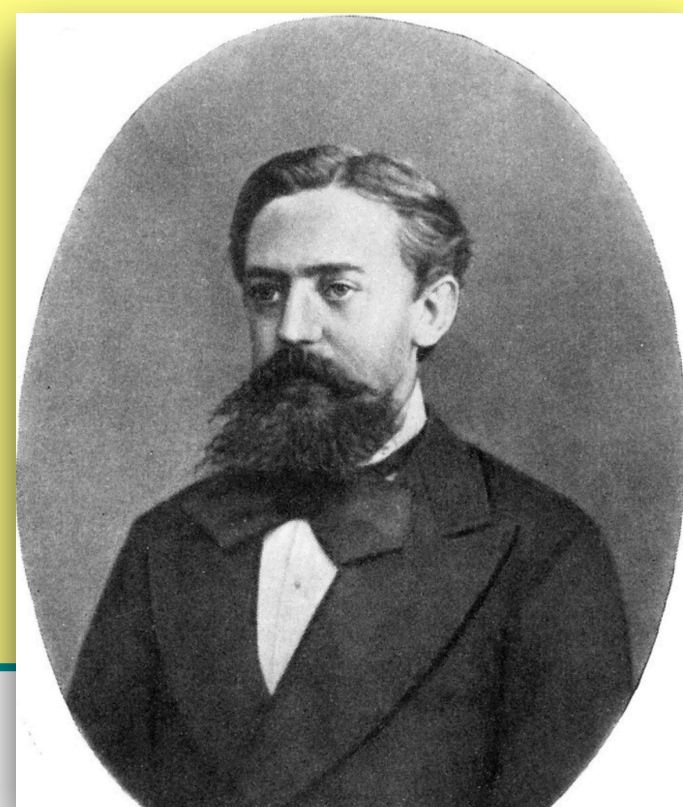
Chebyshev 大数定律显然可由 Markov 大数定律推出; 更重要的是 Markov 大数定律已经没有任何关于独立性的假定.

则大数定律就能成立, 通常称这个条件为 **Markov 条件**, 由此得下述 **Markov 大数定律**.

定理: (Markov 大数定律)

对于随机变量序列 $\xi_1, \xi_2, \dots, \xi_n, \dots$, 若 $\frac{1}{n^2} \cdot D \left(\sum_{k=1}^n \xi_k \right) \longrightarrow 0$, 则对任意的 $\varepsilon > 0$, 均有

$$\lim_{n \rightarrow \infty} P \left\{ \left| \frac{1}{n} \sum_{k=1}^n \xi_k - \frac{1}{n} \sum_{k=1}^n E(\xi_k) \right| < \varepsilon \right\} = 1$$



二、大数定律

- 作为 Chebyshev 大数定律的特殊情况，我们有下面的推论。

推论：(独立同分布下的大数定律)

设 $\xi_1, \xi_2, \dots, \xi_n, \dots$ 是独立同分布的随机变量序列，且 $E(\xi_i) = \mu, D(\xi_i) = \sigma^2, i = 1, 2, \dots$

则对任意给定的 $\varepsilon > 0$,

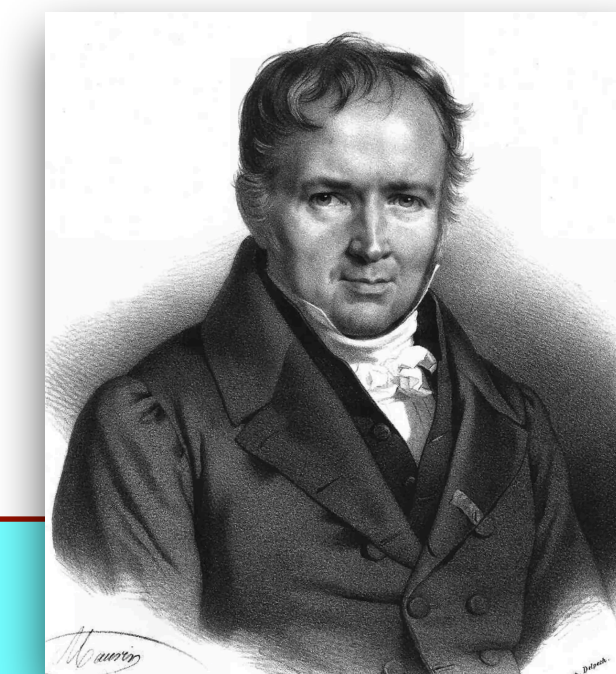
$$\lim_{n \rightarrow \infty} P \left\{ \left| \frac{1}{n} \sum_{i=1}^n \xi_i - \mu \right| < \varepsilon \right\} = 1.$$

Chebyshev 大数定律 $\implies \lim_{n \rightarrow \infty} P \left\{ \left| \frac{1}{n} \sum_{k=1}^n \xi_k - \frac{1}{n} \sum_{k=1}^n E(\xi_k) \right| < \varepsilon \right\} = 1$

$\implies \lim_{n \rightarrow \infty} P \left\{ \left| \frac{1}{n} \sum_{k=1}^n \xi_k - \mu \right| < \varepsilon \right\} = 1$

二、大数定律

- 作为 Chebyshev 大数定律的特殊情况，我们有下面的推论。



推论: (Poisson 大数定律)

在一个独立试验序列中，如果事件 A 在第 k 次试验中出现的概率等于 p_k ，以 μ_n 表示在前 n 次试验中事件 A 出现的次数，则对任意 $\varepsilon > 0$ ，都有

$$\lim_{n \rightarrow \infty} P \left\{ \left| \frac{\mu_n}{n} - \frac{p_1 + p_2 + \dots + p_n}{n} \right| < \varepsilon \right\} = 1.$$

$\xi_k \triangleq$ 第 k 次试验中事件 A 出现的次数

$$\implies E(\xi_k) = p_k, \quad D(\xi_k) = p_k(1 - p_k) \leq \frac{1}{4}$$

| | | |
|------------------|-------|-----------|
| ξ_k | 1 | 0 |
| $P\{\xi_k = x\}$ | p_k | $1 - p_k$ |

Chebyshev 大数定律 \implies

$$\lim_{n \rightarrow \infty} P \left\{ \left| \frac{1}{n} \sum_{k=1}^n \xi_k - \frac{1}{n} \sum_{k=1}^n E(\xi_k) \right| < \varepsilon \right\} = 1 \implies \lim_{n \rightarrow \infty} P \left\{ \left| \frac{\mu_n}{n} - \frac{p_1 + p_2 + \dots + p_n}{n} \right| < \varepsilon \right\} = 1$$

二、大数定律

- 大数定律的重要意义

- ▶ 大数定律建立了大量重复独立试验中事件出现频率具有稳定性，使得**概率**有了客观意义.
- ▶ 大数定律在**偶然性**与**必然性**之间架起了桥梁，是概率论学科最有特色的命题.
 - 要求 n 很大，即它是一类极限定理.
 - 讨论的是平均值，这是概率论的特色.
 - 建立概率接近于 0 或 1 的规律，这是概率论研究中特别要强调的.
 - 规律的产生是大量独立或弱相关因素积累的结果 —— 统计独立性的概念.
- ▶ 参数估计：大数定律是重要理论基础之一.

二、大数定律

- 例：掷一枚均匀的骰子，以 ξ 记向上一面的点数，则出现 1, 2, 3, 4, 5, 6 点向上的概率等可能。

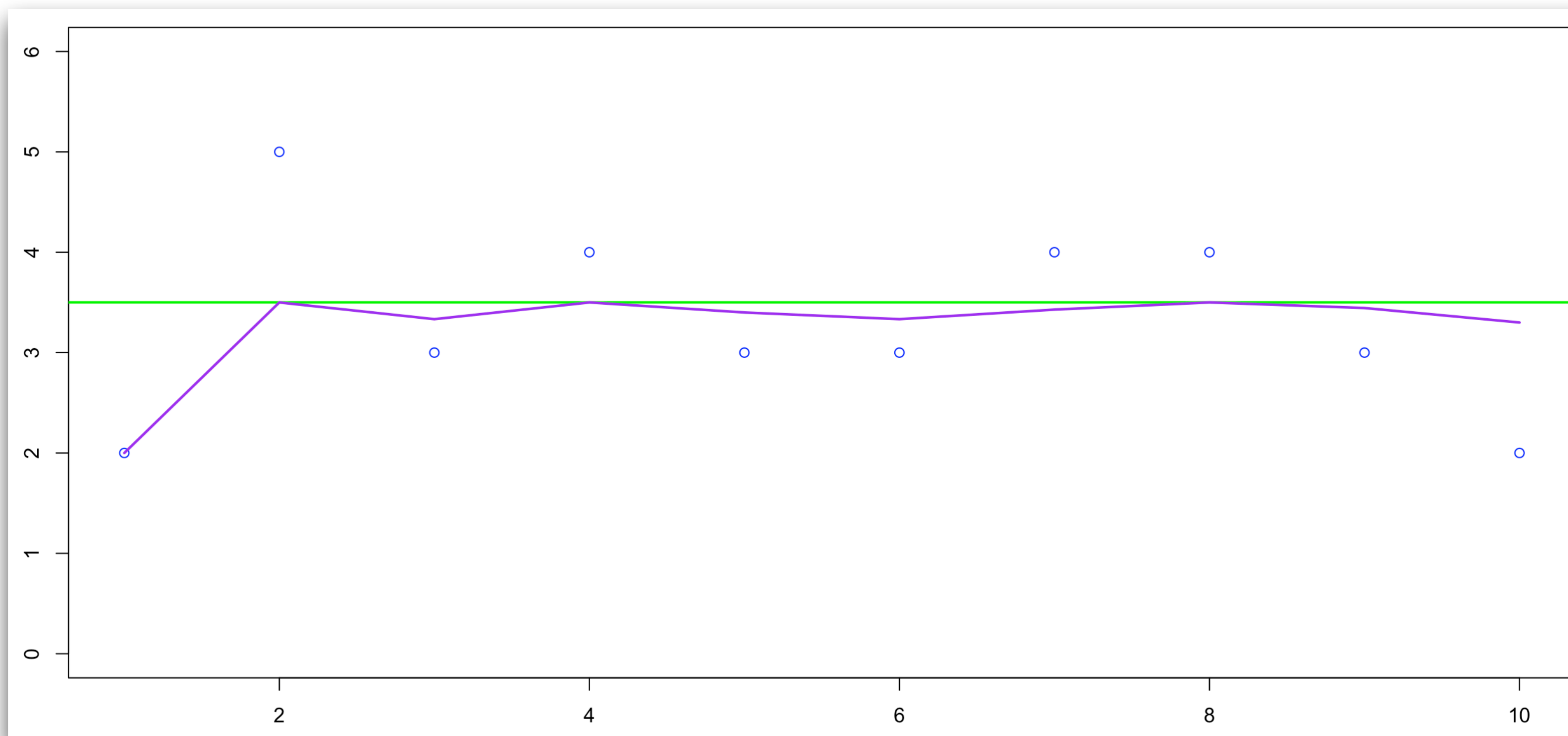
$$E(\xi) = 1 \times \frac{1}{6} + 2 \times \frac{1}{6} + 3 \times \frac{1}{6} + 4 \times \frac{1}{6} + 5 \times \frac{1}{6} + 6 \times \frac{1}{6} = 3.5$$

```
f = function(N=1000) {  
  x = sample(1:6, N, replace=TRUE)  
  plot(x, col="blue", ylim=c(0, 6))  
  abline(h=3.5, col="green", lwd=2)  
  a = array(0, dim=N)  
  b = array(0, dim=N)  
  a = cumsum(x)  
  for (i in 1:N) { b[i] = a[i] / i }  
  lines(b, lwd=2, col="purple")  
}  
f(10)  
f(100)  
f(1000)
```

二、大数定律

- 例：掷一枚均匀的骰子，以 ξ 记向上一面的点数，则出现 1, 2, 3, 4, 5, 6 点向上的概率等可能。

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

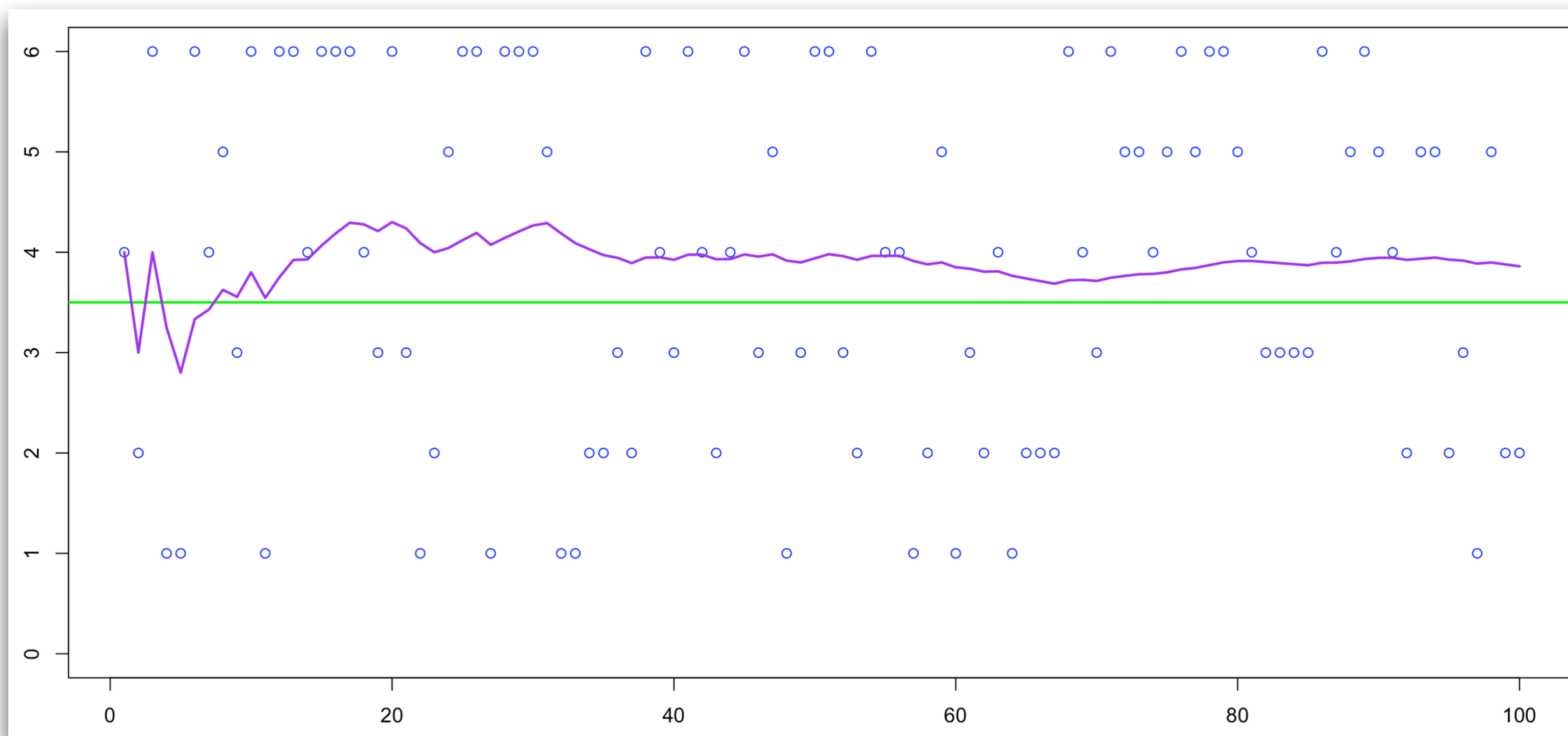


$N = 10$

二、大数定律

- 例：掷一枚均匀的骰子，以 ξ 记向上一面的点数，则出现 1, 2, 3, 4, 5, 6 点向上的概率等可能。

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

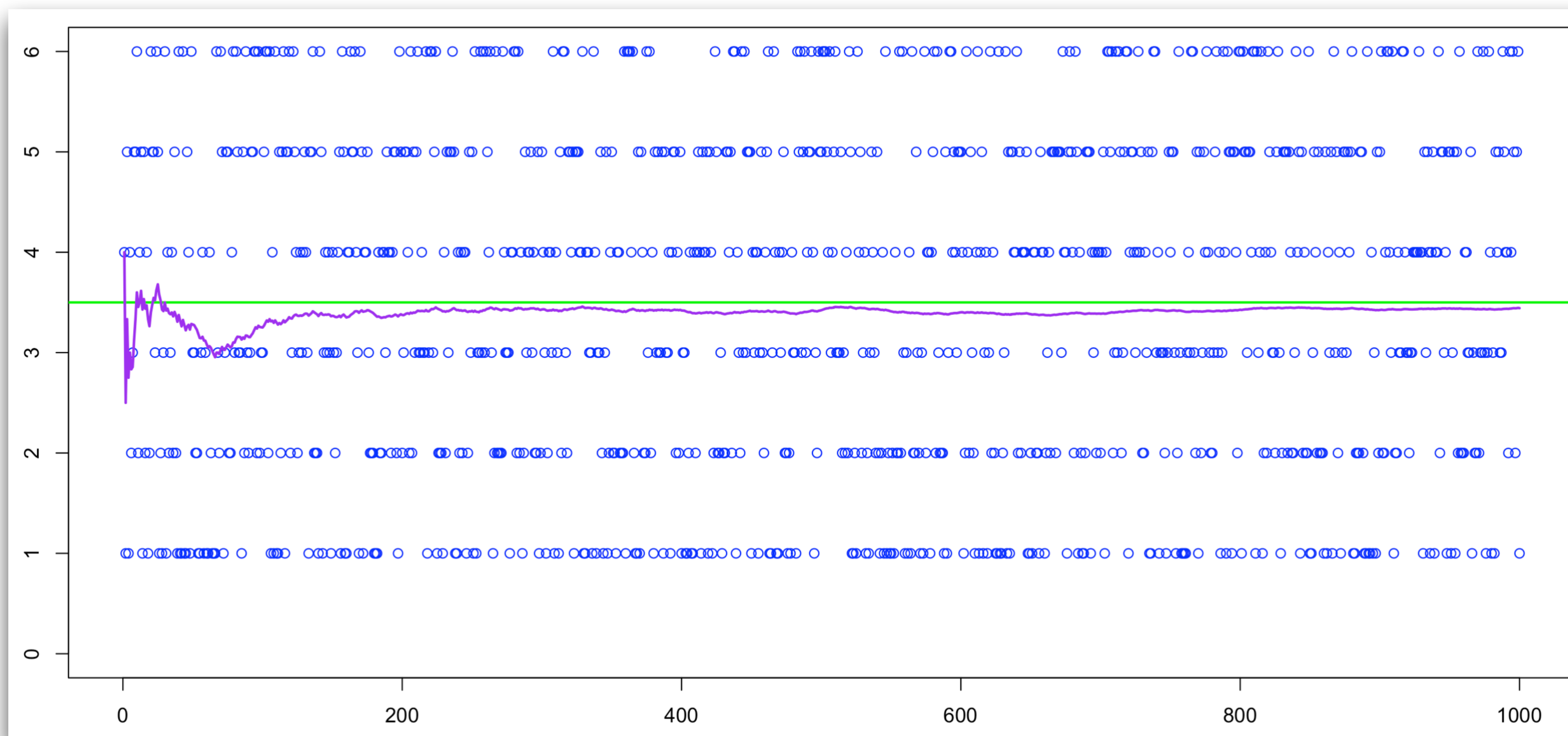


$N = 100$

二、大数定律

- 例：掷一枚均匀的骰子，以 ξ 记向上一面的点数，则出现 1, 2, 3, 4, 5, 6 点向上的概率等可能。

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



$N = 1000$

二、大数定律

- 例：掷一枚硬币， p 为正面向上的概率， ξ 表示掷一次硬币出现正面向上的次数.

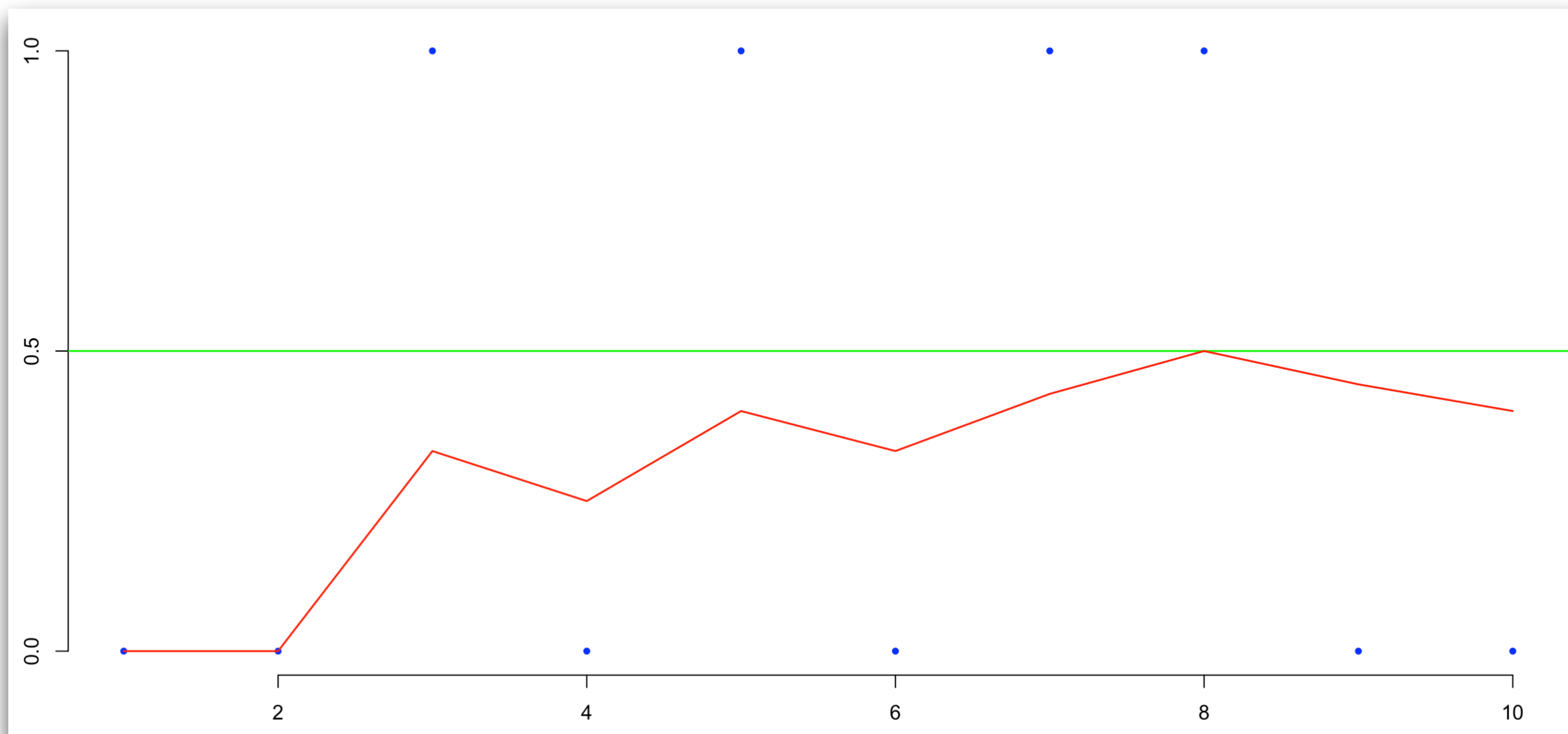
$$E(\xi) = 1 \times p + 0 \times (1 - p) = p$$

```
f = function(N=500, p=1/2){  
  x = sample(0:1, N, replace=TRUE, prob=c(1-p, p))  
  plot(x, col="blue", pch=16, cex=0.75, axes=F, xlab="", ylab="")  
  axis(1); axis(2, at=c(0, 0.5, 1))  
  abline(h=p, col="green", lwd=1.5)  
  a = array(0, dim=N)  
  b = array(0, dim=N)  
  a[1] = x[1]  
  for ( i in 1:(N-1)) { a[i+1] = a[i] + x[i+1] }  
  for (i in 1:N) { b[i] = a[i] / i }  
  lines(b, lwd=1.5, col="red")  
}  
f(10)  
f(100)  
f(500)
```

二、大数定律

- 例：掷一枚硬币， p 为正面向上的概率， ξ 表示掷一次硬币出现正面向上的次数。

$$E(\xi) = 1 \times p + 0 \times (1 - p) = p$$

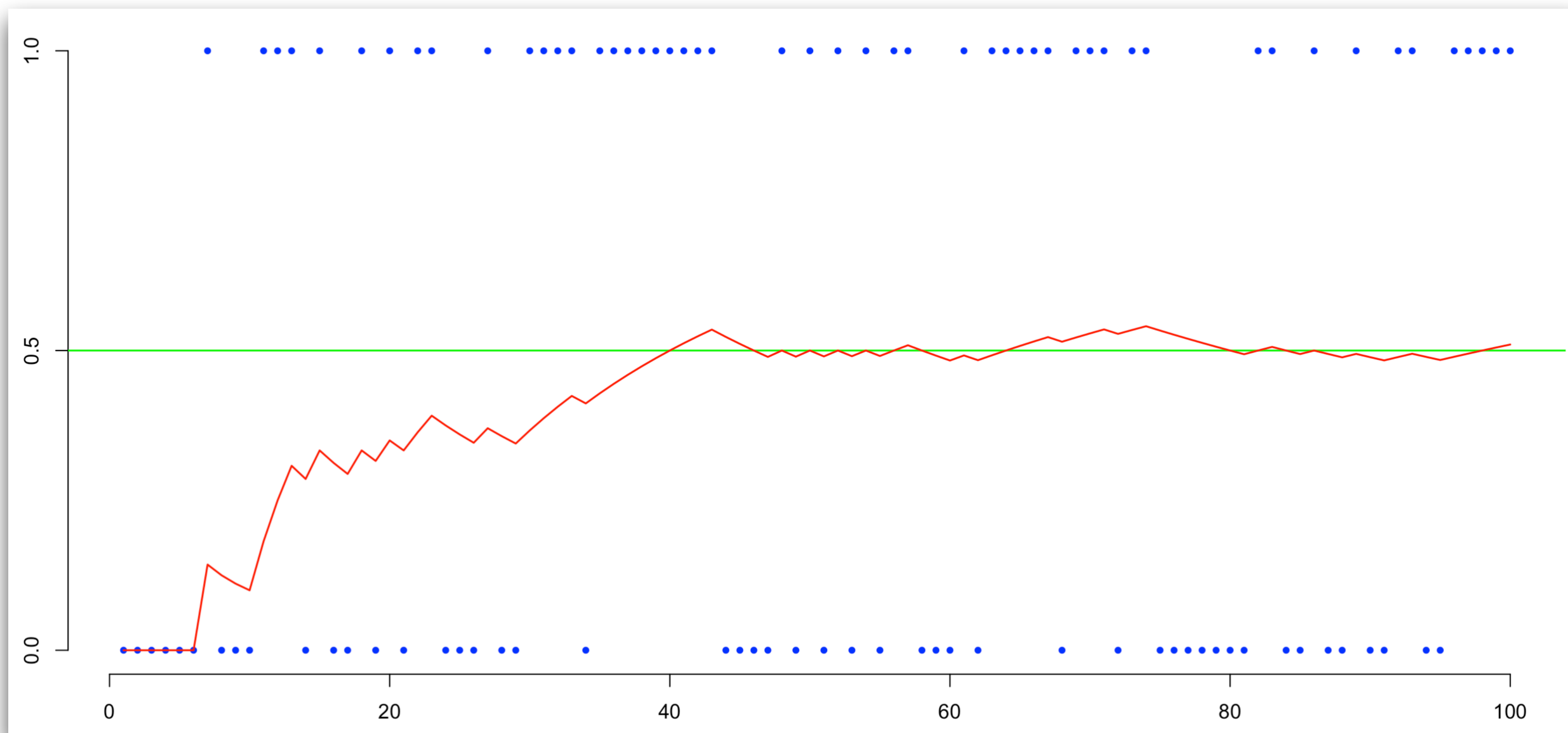


$N = 10$

二、大数定律

- 例：掷一枚硬币， p 为正面向上的概率， ξ 表示掷一次硬币出现正面向上的次数.

$$E(\xi) = 1 \times p + 0 \times (1 - p) = p$$

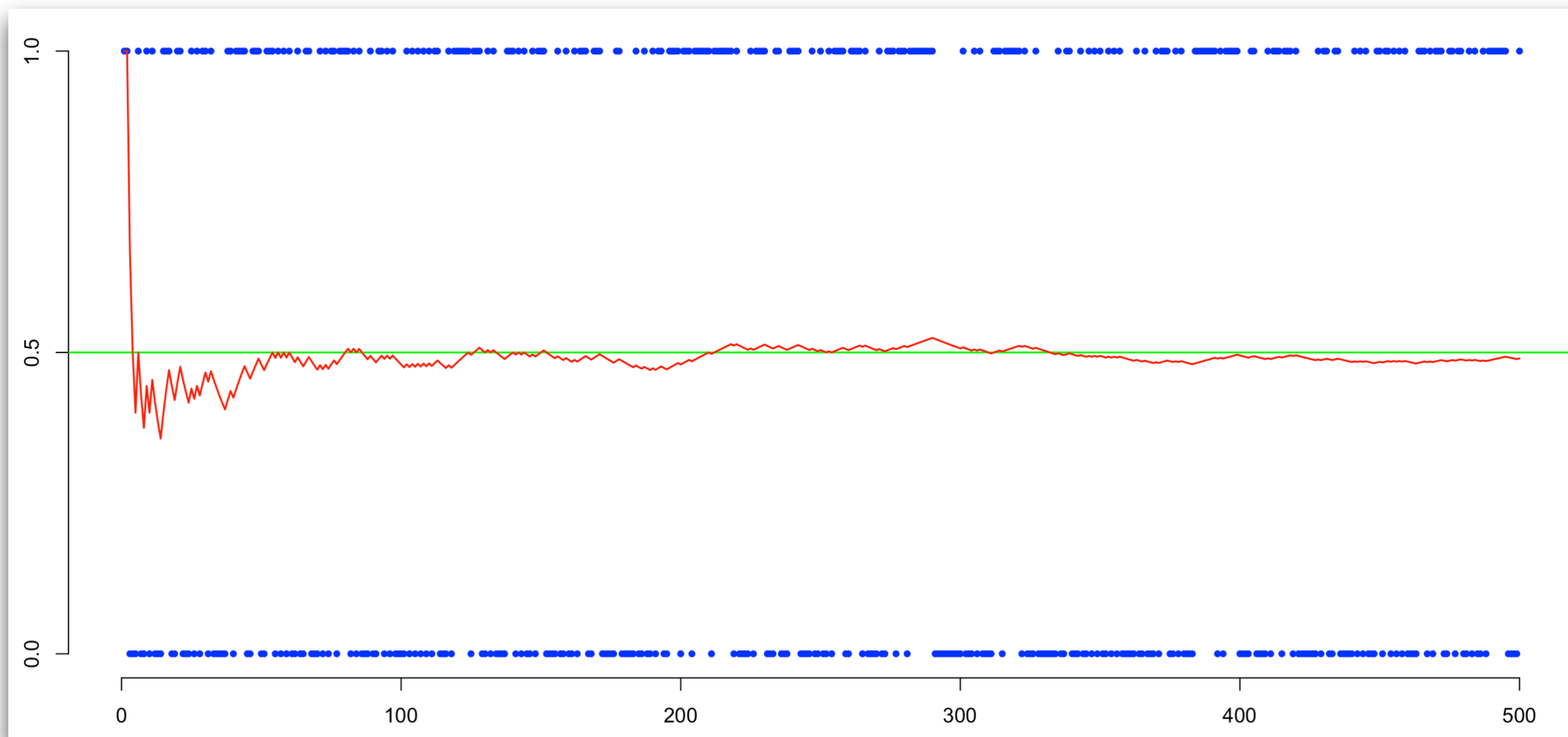


$N = 100$

二、大数定律

- 例：掷一枚硬币， p 为正面向上的概率， ξ 表示掷一次硬币出现正面向上的次数.

$$E(\xi) = 1 \times p + 0 \times (1 - p) = p$$



$N = 500$

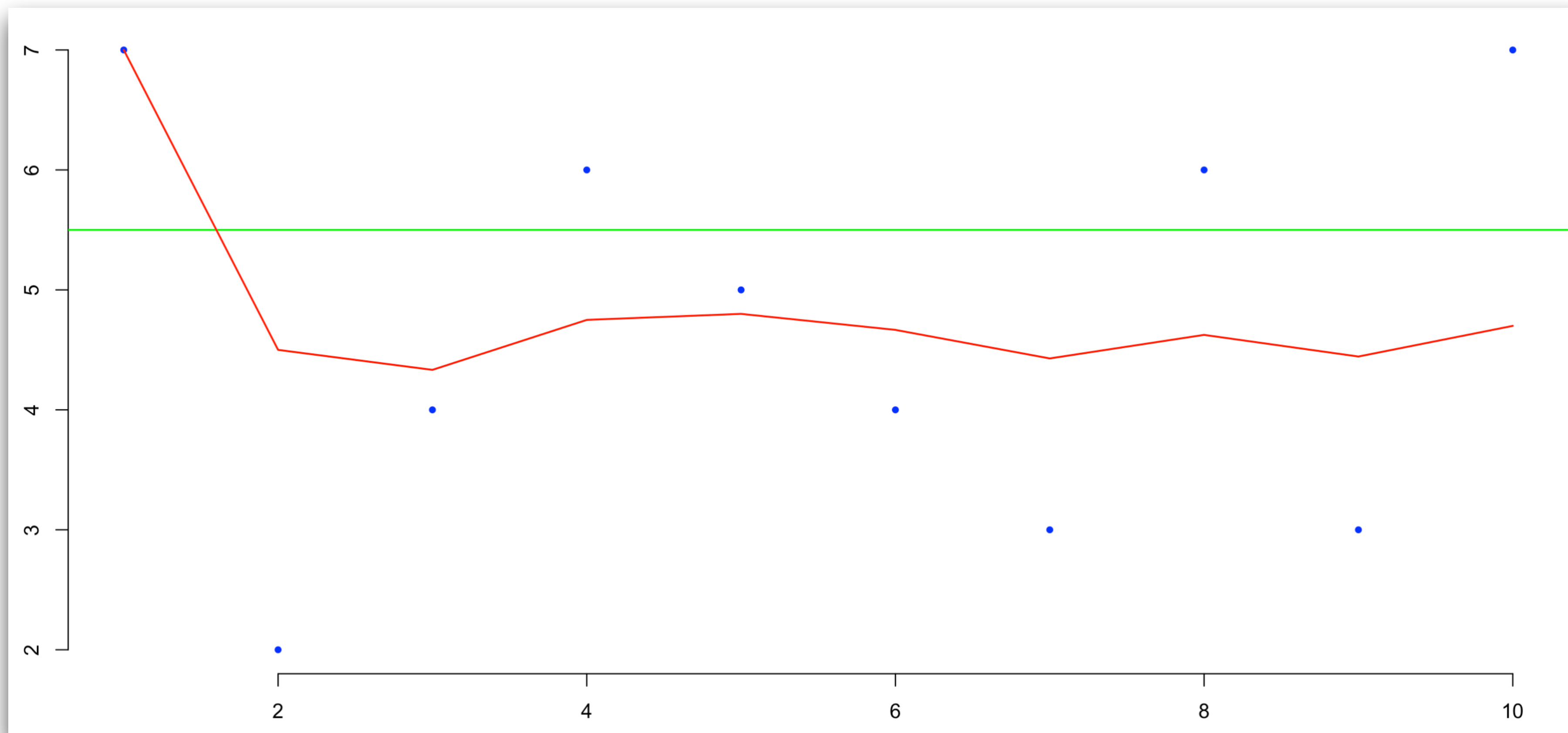
二、大数定律

- 例：设 $\xi \sim P(\lambda)$. 则 $E(\xi) = \lambda$.

```
f = function(N=300, lambda=5.5){  
  x = rpois(N, lambda)  
  plot(x, col="blue", pch=16, cex=0.75, axes=F, xlab="", ylab="")  
  axis(1); axis(2)  
  abline(h=lambda, col="green", lwd=1.5)  
  a = array(0, dim=N)  
  b = array(0, dim=N)  
  a[1] = x[1]  
  for ( i in 1:(N-1)) { a[i+1] = a[i] + x[i+1] }  
  for (i in 1:N) { b[i] = a[i] / i }  
  lines(b, lwd=1.5, col="red")  
}  
f(10)  
f(100)  
f(500)
```

二、大数定律

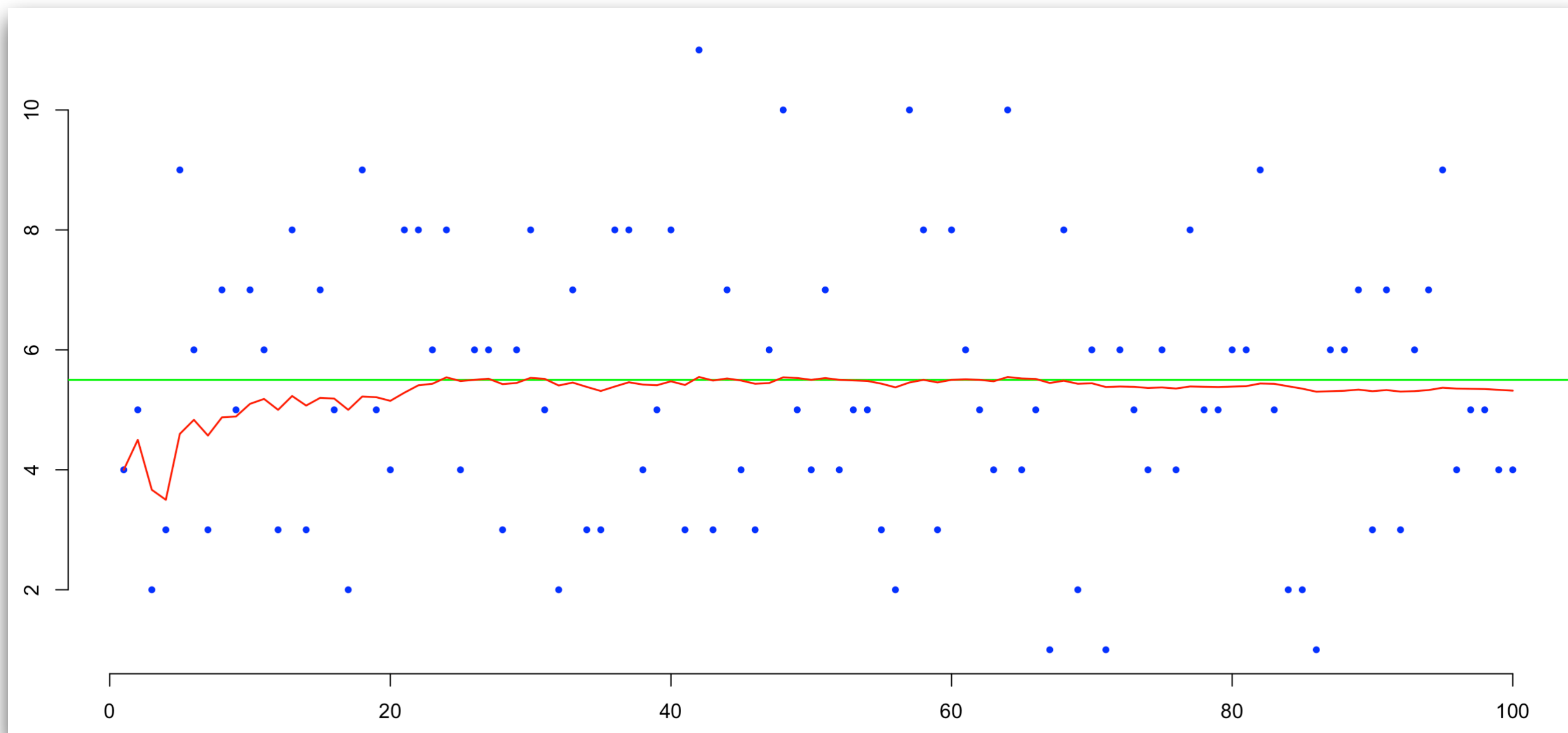
- 例：设 $\xi \sim P(\lambda)$. 则 $E(\xi) = \lambda$.



$N = 10$

二、大数定律

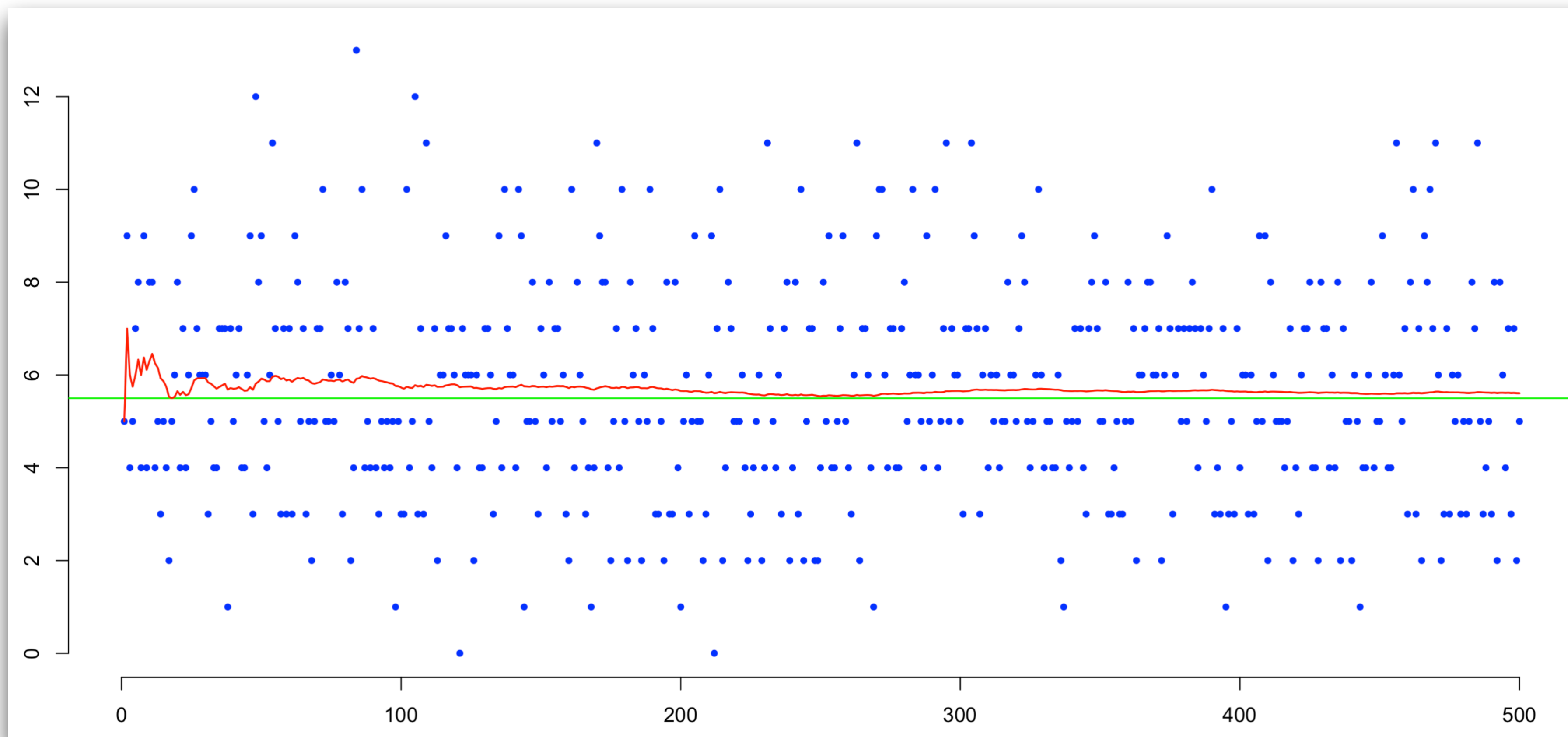
- 例：设 $\xi \sim P(\lambda)$. 则 $E(\xi) = \lambda$.



$N = 100$

二、大数定律

- 例：设 $\xi \sim P(\lambda)$. 则 $E(\xi) = \lambda$.



$N = 500$

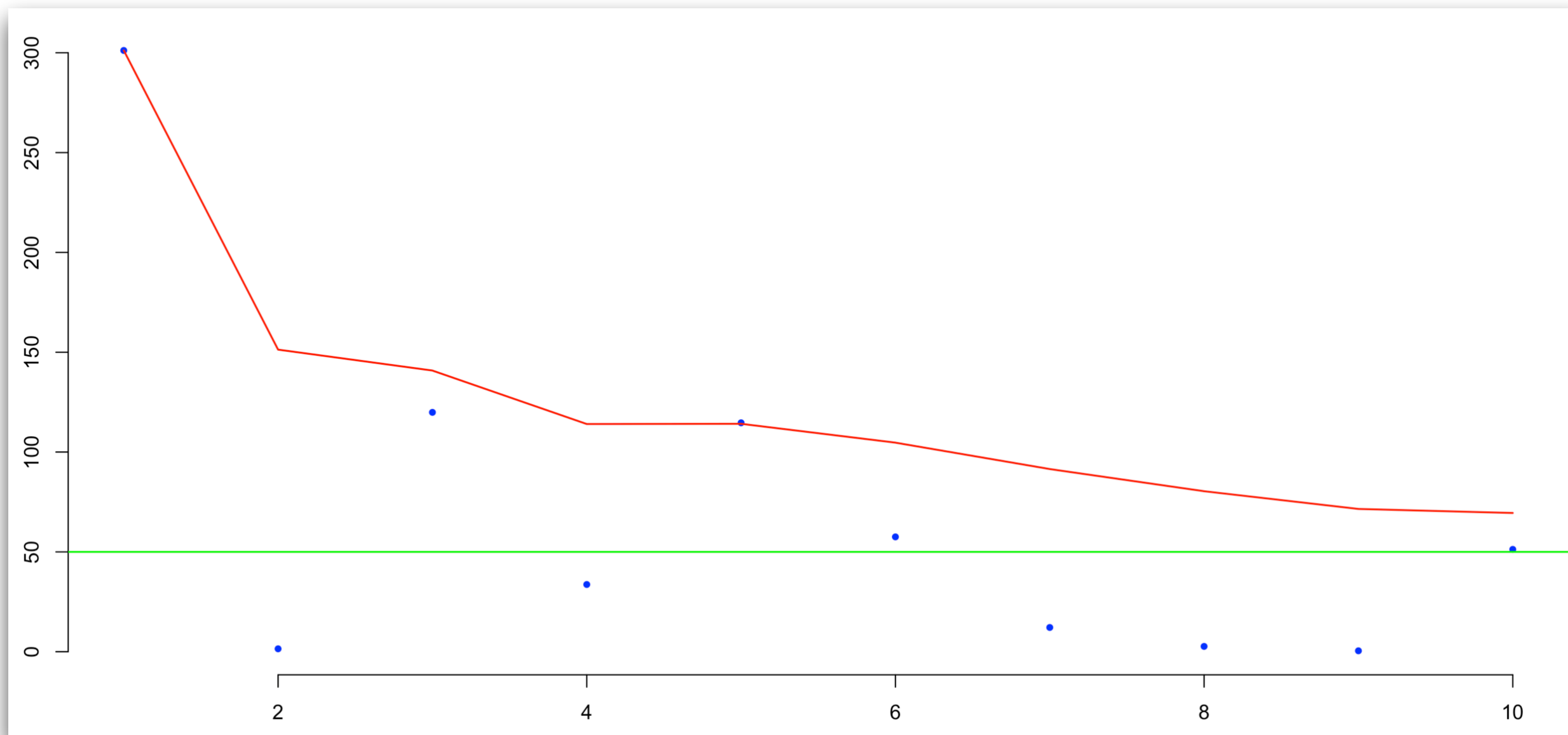
二、大数定律

- 例：设 $\xi \sim \text{Exp}(\theta)$. 则 $E(\xi) = \frac{1}{\theta}$.

```
f = function(N=300, theta=1/50){  
  x = rexp(N, theta)  
  plot(x, col="blue", pch=16, cex=0.75, axes=F, xlab="", ylab="")  
  axis(1); axis(2)  
  abline(h=(1/theta), col="green", lwd=1.5)  
  a = array(0, dim=N)  
  b = array(0, dim=N)  
  a[1] = x[1]  
  for ( i in 1:(N-1)) { a[i+1] = a[i] + x[i+1] }  
  for (i in 1:N) { b[i] = a[i] / i }  
  lines(b, lwd=1.5, col="red")  
}  
f(10)  
f(100)  
f(1000)
```

二、大数定律

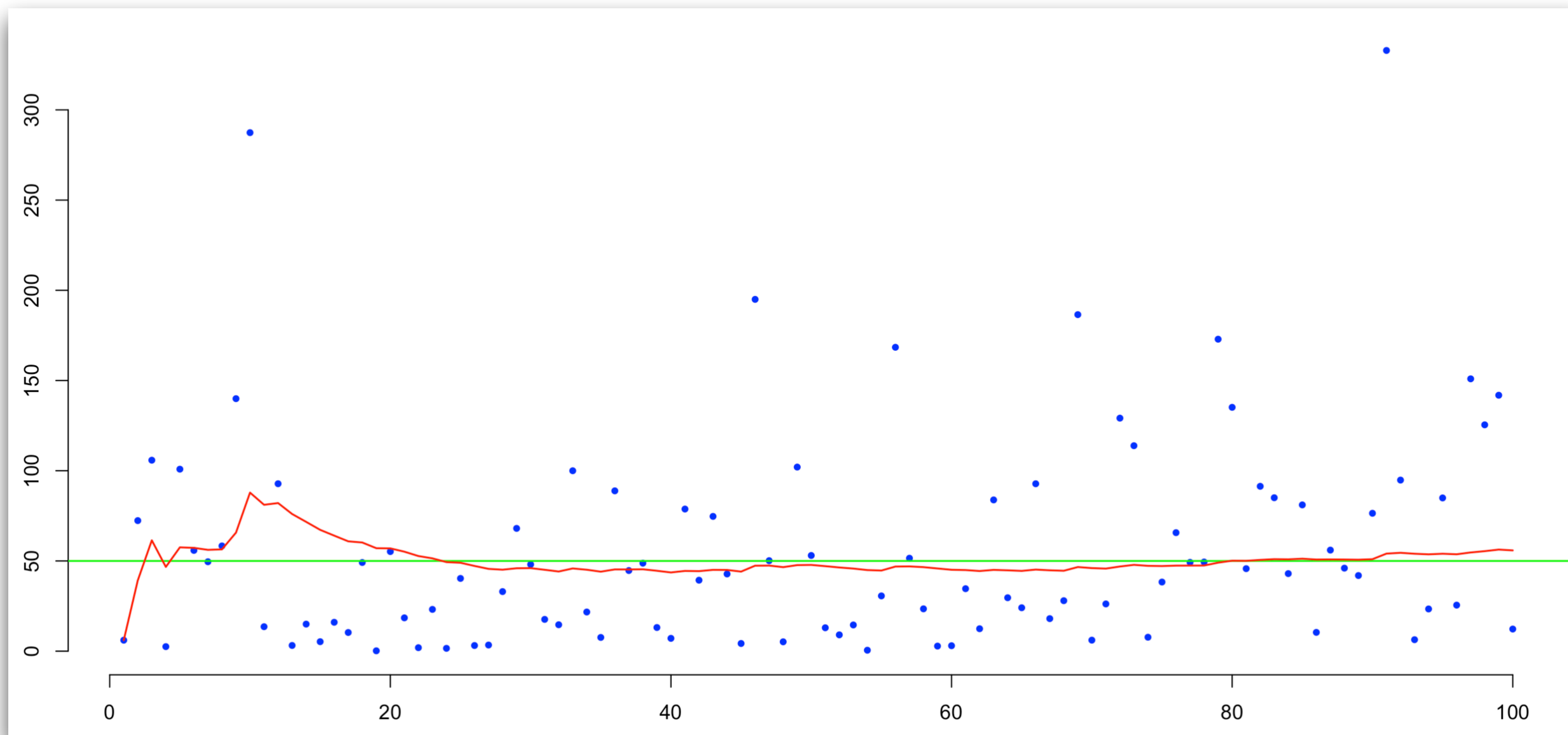
- 例：设 $\xi \sim \text{Exp}(\theta)$. 则 $E(\xi) = \frac{1}{\theta}$.



$N = 10$

二、大数定律

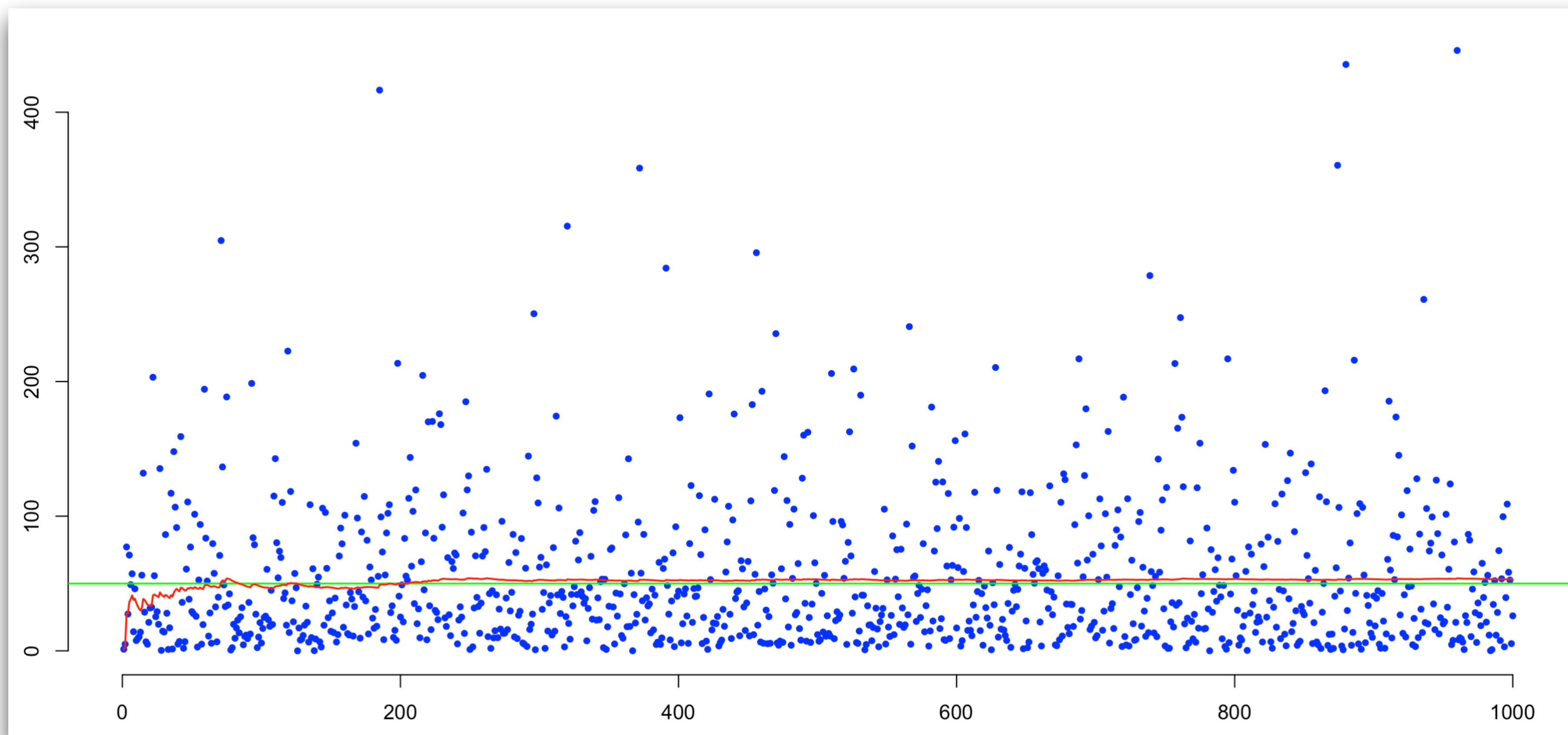
- 例：设 $\xi \sim \text{Exp}(\theta)$. 则 $E(\xi) = \frac{1}{\theta}$.



$N = 100$

二、大数定律

- 例：设 $\xi \sim \text{Exp}(\theta)$. 则 $E(\xi) = \frac{1}{\theta}$.



$N = 1000$

三、De Moivre - Laplace 极限定理

定理： 若 μ_n 是 n 次 Bernoulli 试验中事件 A 出现的次数， $0 < p < 1$ ，则对任意有限区间 $[a, b]$ ：

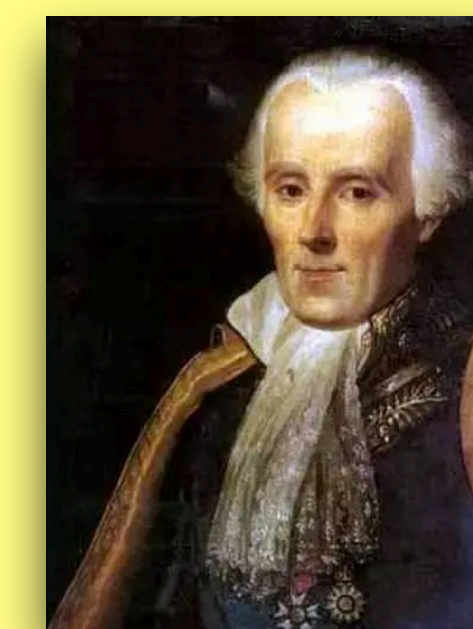
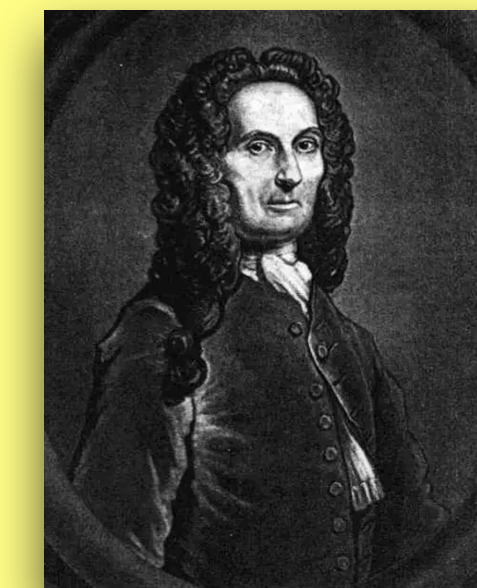
① 当 $a \leq x_k \equiv \frac{k - np}{\sqrt{np(1-p)}} \leq b$ 及 $n \rightarrow \infty$ 时，一致地有

局部极限定理

$$P\{\mu_n = k\} \div \left(\frac{1}{\sqrt{np(1-p)}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_k^2} \right) \longrightarrow 1$$

② 当 $n \rightarrow \infty$ 时一致地有

积分极限定理

$$P\left\{ a \leq \frac{\mu_n - np}{\sqrt{np(1-p)}} < b \right\} \longrightarrow \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \quad (-\infty < x < +\infty)$$


三、De Moivre - Laplace 极限定理

局部极限定理的证明:

$$x_k \in [a, b], \quad n \rightarrow \infty$$

$$\Rightarrow \begin{cases} k = np + x_k \sqrt{np(1-p)} \rightarrow \infty \\ j = n - k = n(1-p) - x_k \sqrt{np(1-p)} \rightarrow \infty \end{cases}$$

Stirling 公式 $\rightarrow m! = \sqrt{2\pi m} \cdot m^m \cdot e^{-m} \cdot e^{\theta_m}, \quad \left(0 < \theta_m < \frac{1}{12m}\right)$

$$\Rightarrow P\{\mu_n = k\} = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} = \frac{n!}{k!j!} p^k q^j = \frac{\sqrt{2\pi n} \cdot n^n \cdot e^{-n} \cdot e^{\theta_n}}{\left(\sqrt{2\pi k} \cdot k^k \cdot e^{-k} \cdot e^{\theta_k}\right) \left(\sqrt{2\pi j} \cdot j^j \cdot e^{-j} \cdot e^{\theta_j}\right)} p^k q^j$$

$j = n - k, q = 1 - p$

$\mu_n \sim b(n, p)$

$$= \frac{\sqrt{2\pi n} \cdot n^n \cdot e^{-n}}{\left(\sqrt{2\pi k} \cdot k^k \cdot e^{-k}\right) \left(\sqrt{2\pi j} \cdot j^j \cdot e^{-j}\right)} p^k q^j \cdot e^{\theta_n - \theta_k - \theta_j} = \frac{\sqrt{2\pi} n^{n+\frac{1}{2}}}{2\pi k^{k+\frac{1}{2}} j^{j+\frac{1}{2}}} p^k q^j \cdot e^{\theta}$$

$k + j = n \qquad \theta = \theta_n - \theta_k - \theta_j$

定理: 若 μ_n 是 n 次 Bernoulli 试验中事件 A 出现的次数, $0 < p < 1$, 则对任意有限区间 $[a, b]$:

① 当 $a \leq x_k \equiv \frac{k - np}{\sqrt{np(1-p)}} \leq b$ 及 $n \rightarrow \infty$ 时, 一致地有
局部极限定理

$$P\{\mu_n = k\} \div \left(\frac{1}{\sqrt{np(1-p)}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_k^2} \right) \rightarrow 1$$

② 当 $n \rightarrow \infty$ 时一致地有

积分极限定理 $P\left\{a \leq \frac{\mu_n - np}{\sqrt{np(1-p)}} < b\right\} \rightarrow \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \quad (-\infty < x < +\infty)$



三、De Moivre - Laplace 极限定理

局部极限定理的证明:

$$\Rightarrow P\{\mu_n = k\} = \frac{\sqrt{2\pi} n^{n+\frac{1}{2}}}{2\pi k^{k+\frac{1}{2}} j^{j+\frac{1}{2}}} p^k q^j \cdot e^\theta$$

$$= \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\sqrt{npq}} \left(\frac{np}{k}\right)^{k+\frac{1}{2}} \left(\frac{nq}{n-k}\right)^{n-k+\frac{1}{2}} \cdot e^\theta$$

$$\Rightarrow \sqrt{2\pi npq} \cdot P\{\mu_n = k\} = \left(\frac{k}{np}\right)^{-(k+\frac{1}{2})} \left(\frac{n-k}{nq}\right)^{-(n-k+\frac{1}{2})} \cdot e^\theta = \left(1 + x_k \sqrt{\frac{q}{np}}\right)^{-(k+\frac{1}{2})} \left(1 - x_k \sqrt{\frac{p}{nq}}\right)^{-(n-k+\frac{1}{2})} \cdot e^\theta$$

$$\begin{cases} k = np + x_k \sqrt{np(1-p)} \rightarrow \infty \\ j = n - k = n(1-p) - x_k \sqrt{np(1-p)} \rightarrow \infty \end{cases} \Rightarrow \begin{cases} \frac{k}{np} = 1 + x_k \sqrt{\frac{q}{np}} \\ \frac{n-k}{nq} = 1 - x_k \sqrt{\frac{p}{nq}} \end{cases}$$

$$\Rightarrow \ln\left(\sqrt{2\pi npq} \cdot P\{\mu_n = k\}\right) = \theta - \left(np + x_k \sqrt{npq} + \frac{1}{2}\right) \ln\left(1 + x_k \sqrt{\frac{q}{np}}\right) - \left(nq - x_k \sqrt{npq} + \frac{1}{2}\right) \ln\left(1 - x_k \sqrt{\frac{p}{nq}}\right)$$

定理: 若 μ_n 是 n 次 Bernoulli 试验中事件 A 出现的次数, $0 < p < 1$, 则对任意有限区间 $[a, b]$:

① 当 $a \leq x_k \equiv \frac{k - np}{\sqrt{np(1-p)}} \leq b$ 及 $n \rightarrow \infty$ 时, 一致地有
局部极限定理

$$P\{\mu_n = k\} \div \left(\frac{1}{\sqrt{np(1-p)}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_k^2} \right) \rightarrow 1$$

② 当 $n \rightarrow \infty$ 时一致地有
积分极限定理

$$P\left\{a \leq \frac{\mu_n - np}{\sqrt{np(1-p)}} < b\right\} \rightarrow \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \quad (-\infty < x < +\infty)$$



三、De Moivre - Laplace 极限定理

局部极限定理的证明:

$$\theta = \theta_n - \theta_k - \theta_j, \quad 0 < \theta_m < \frac{1}{12m}$$

$$|\theta| = |\theta_n - \theta_k - \theta_{n-k}| \leq |\theta_n| + |\theta_k| + |\theta_{n-k}| < \frac{1}{12} \left(\frac{1}{n} + \frac{1}{k} + \frac{1}{n-k} \right) \rightarrow 0$$

展开式 $\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots$ 收敛域 $x \in (-1, 1]$, 且在 $|x|$ 很小时收敛的快

$$\begin{aligned} \Rightarrow \ln \left(\sqrt{2\pi npq} \cdot P \{ \mu_n = k \} \right) &= \theta - \left(np + x_k \sqrt{npq} + \frac{1}{2} \right) \cdot \left(x_k \sqrt{\frac{q}{np}} - \frac{x_k^2 q}{2np} + \frac{x_k^3 q \sqrt{npq}}{3n^2 p^2} - \frac{x_k^4 q^2}{4n^2 p^2} + \dots \right) \\ &\quad - \left(nq - x_k \sqrt{npq} + \frac{1}{2} \right) \cdot \left(-x_k \sqrt{\frac{p}{nq}} - \frac{x_k^2 p}{2nq} - \frac{x_k^3 p \sqrt{npq}}{3n^2 q^2} - \frac{x_k^4 p^2}{4n^2 q^2} + \dots \right) \end{aligned}$$

$$\Rightarrow \ln \left(\sqrt{2\pi npq} \cdot P \{ \mu_n = k \} \right) = \theta - \left(np + x_k \sqrt{npq} + \frac{1}{2} \right) \ln \left(1 + x_k \sqrt{\frac{q}{np}} \right) - \left(nq - x_k \sqrt{npq} + \frac{1}{2} \right) \ln \left(1 - x_k \sqrt{\frac{p}{nq}} \right) \rightarrow 0 \quad (n \rightarrow \infty)$$

定理: 若 μ_n 是 n 次 Bernoulli 试验中事件 A 出现的次数, $0 < p < 1$, 则对任意有限区间 $[a, b]$:

① 当 $a \leq x_k \equiv \frac{k - np}{\sqrt{np(1-p)}} \leq b$ 及 $n \rightarrow \infty$ 时, 一致地有
局部极限定理

$$P \{ \mu_n = k \} \div \left(\frac{1}{\sqrt{np(1-p)}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_k^2} \right) \rightarrow 1$$

② 当 $n \rightarrow \infty$ 时一致地有
积分极限定理

$$P \left\{ a \leq \frac{\mu_n - np}{\sqrt{np(1-p)}} < b \right\} \rightarrow \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \quad (-\infty < x < +\infty)$$



三、De Moivre - Laplace 极限定理

局部极限定理的证明:

$$= \theta - \frac{1}{2}x_k^2 + \frac{x_k^3 q \sqrt{npq}}{2np} - \frac{x_k^3 q \sqrt{npq}}{3np} - \frac{x_k^3 p \sqrt{npq}}{2nq} + \frac{x_k^3 p \sqrt{npq}}{3nq} - \frac{1}{2}x_k \sqrt{\frac{q}{np}} + \frac{1}{2}x_k \sqrt{\frac{p}{nq}}$$

$$+ \frac{x_k^2 q}{4np} + \frac{x_k^2 p}{4nq} - \frac{x_k^4 q^2}{3np} + \frac{x_k^4 q^2}{4np} - \frac{x_k^4 p^2}{3nq} + \frac{x_k^4 p^2}{4nq} + o\left(\frac{1}{n}\right)$$

$$\Rightarrow \ln\left(\sqrt{2\pi npq} \cdot P\{\mu_n = k\}\right) = \theta - \left(np + x_k \sqrt{npq} + \frac{1}{2}\right) \cdot \left(x_k \sqrt{\frac{q}{np}} - \frac{x_k^2 q}{2np} + \frac{x_k^3 q \sqrt{npq}}{3n^2 p^2} - \frac{x_k^4 q^2}{4n^2 p^2} + \dots\right)$$

$$- \left(nq - x_k \sqrt{npq} + \frac{1}{2}\right) \cdot \left(-x_k \sqrt{\frac{p}{nq}} - \frac{x_k^2 p}{2nq} - \frac{x_k^3 p \sqrt{npq}}{3n^2 q^2} - \frac{x_k^4 p^2}{4n^2 q^2} + \dots\right)$$

$$= \theta - x_k \sqrt{npq} - x_k^2 q - \frac{1}{2}x_k \sqrt{\frac{q}{np}} + \frac{1}{2}x_k^2 q + \frac{x_k^3 q \sqrt{npq}}{2np} + \frac{x_k^2 q}{4np} - \frac{x_k^3 q \sqrt{npq}}{3np} - \frac{x_k^4 npq^2}{3n^2 p^2} - \frac{x_k^3 q \sqrt{npq}}{6n^2 p^2} + \frac{x_k^4 q^2}{4np} + \frac{x_k^5 q^2 \sqrt{npq}}{4n^2 p^2} + \frac{x_k^4 q^2}{8n^2 p^2} - \dots$$

$$+ x_k \sqrt{npq} - x_k^2 p + \frac{1}{2}x_k \sqrt{\frac{p}{nq}} + \frac{1}{2}x_k^2 p - \frac{x_k^3 p \sqrt{npq}}{2nq} + \frac{x_k^2 p}{4nq} + \frac{x_k^3 p \sqrt{npq}}{3nq} - \frac{x_k^4 np^2 q}{3n^2 q^2} + \frac{x_k^3 p \sqrt{npq}}{6n^2 q^2} + \frac{x_k^4 p^2}{4nq} - \frac{x_k^5 p^2 \sqrt{npq}}{4n^2 q^2} + \frac{x_k^4 p^2}{8n^2 q^2} + \dots$$

三、De Moivre - Laplace 极限定理

局部极限定理的证明:

$$\begin{aligned}
 &= \theta - \frac{1}{2}x_k^2 + \frac{x_k^3 q \sqrt{npq}}{2np} - \frac{x_k^3 q \sqrt{npq}}{3np} - \frac{x_k^3 p \sqrt{npq}}{2nq} + \frac{x_k^3 p \sqrt{npq}}{3nq} - \frac{1}{2}x_k \sqrt{\frac{q}{np}} + \frac{1}{2}x_k \sqrt{\frac{p}{nq}} \\
 &\quad + \frac{x_k^2 q}{4np} + \frac{x_k^2 p}{4nq} - \frac{x_k^4 q^2}{3np} + \frac{x_k^4 q^2}{4np} - \frac{x_k^4 p^2}{3nq} + \frac{x_k^4 p^2}{4nq} + o\left(\frac{1}{n}\right)
 \end{aligned}$$

$$\begin{aligned}
 \Rightarrow \ln\left(\sqrt{2\pi npq} \cdot P\{\mu_n = k\}\right) &= \theta - \left(np + x_k \sqrt{npq} + \frac{1}{2}\right) \cdot \left(x_k \sqrt{\frac{q}{np}} - \frac{x_k^2 q}{2np} + \frac{x_k^3 q \sqrt{npq}}{3n^2 p^2} - \frac{x_k^4 q^2}{4n^2 p^2} + \dots\right) \\
 &\quad - \left(nq - x_k \sqrt{npq} + \frac{1}{2}\right) \cdot \left(-x_k \sqrt{\frac{p}{nq}} - \frac{x_k^2 p}{2nq} - \frac{x_k^3 p \sqrt{npq}}{3n^2 q^2} - \frac{x_k^4 p^2}{4n^2 q^2} + \dots\right)
 \end{aligned}$$

$$= \theta - \frac{1}{2}x_k^2 + \frac{x_k^3}{6\sqrt{npq}}(3q^2 - 2q^2 - 3p^2 + 2p^2) - \frac{3x_k(q-p)}{6\sqrt{npq}} + \frac{x_k^2}{4npq}(q^2 + p^2) - \frac{x_k^4}{12npq}(4q^3 - 3q^3 + 4p^3 - 3p^3) + o\left(\frac{1}{n}\right)$$

$$= \theta - \frac{1}{2}x_k^2 + \frac{x_k^3}{6\sqrt{npq}}(q^2 - p^2) - \frac{3x_k(q-p)}{6\sqrt{npq}} + \frac{x_k^2}{4npq}(q^2 + p^2) - \frac{x_k^4}{12npq}(q^3 + p^3) + o\left(\frac{1}{n}\right)$$

三、De Moivre - Laplace 极限定理

局部极限定理的证明:

$$\Rightarrow \ln \left(\sqrt{2\pi npq} \cdot P \{ \mu_n = k \} \right) = \theta - \frac{1}{2} x_k^2 + \frac{q-p}{6\sqrt{npq}} (x_k^3 - 3x_k) + \frac{1}{12npq} \left[3(q^2 + p^2) x_k^2 - (q^3 + p^3) x_k^4 \right] + o \left(\frac{1}{n} \right)$$

$$\Rightarrow P \{ \mu_n = k \} = \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\sqrt{npq}} \cdot \exp \left\{ \theta - \frac{1}{2} x_k^2 + \frac{q-p}{6\sqrt{npq}} (x_k^3 - 3x_k) + \frac{1}{12npq} \left[3(q^2 + p^2) x_k^2 - (q^3 + p^3) x_k^4 \right] + o \left(\frac{1}{n} \right) \right\}$$

$$\Rightarrow \ln \left(\sqrt{2\pi npq} \cdot P \{ \mu_n = k \} \right) = \theta - \left(np + x_k \sqrt{npq} + \frac{1}{2} \right) \cdot \left(x_k \sqrt{\frac{q}{np}} - \frac{x_k^2 q}{2np} + \frac{x_k^3 q \sqrt{npq}}{3n^2 p^2} - \frac{x_k^4 q^2}{4n^2 p^2} + \dots \right) - \left(nq - x_k \sqrt{npq} + \frac{1}{2} \right) \cdot \left(-x_k \sqrt{\frac{p}{nq}} - \frac{x_k^2 p}{2nq} - \frac{x_k^3 p \sqrt{npq}}{3n^2 q^2} - \frac{x_k^4 p^2}{4n^2 q^2} + \dots \right)$$

$$= \theta - \frac{1}{2} x_k^2 + \frac{x_k^3}{6\sqrt{npq}} (3q^2 - 2q^2 - 3p^2 + 2p^2) - \frac{3x_k (q-p)}{6\sqrt{npq}} + \frac{x_k^2}{4npq} (q^2 + p^2) - \frac{x_k^4}{12npq} (4q^3 - 3q^3 + 4p^3 - 3p^3) + o \left(\frac{1}{n} \right)$$

$$= \theta - \frac{1}{2} x_k^2 + \frac{x_k^3}{6\sqrt{npq}} (q^2 - p^2) - \frac{3x_k (q-p)}{6\sqrt{npq}} + \frac{x_k^2}{4npq} (q^2 + p^2) - \frac{x_k^4}{12npq} (q^3 + p^3) + o \left(\frac{1}{n} \right)$$

三、De Moivre - Laplace 极限定理

局部极限定理的证明:

$$\Rightarrow \ln \left(\sqrt{2\pi npq} \cdot P \{ \mu_n = k \} \right) = \theta - \frac{1}{2} x_k^2 + \frac{q-p}{6\sqrt{npq}} (x_k^3 - 3x_k) + \frac{1}{12npq} \left[3(q^2 + p^2) x_k^2 - (q^3 + p^3) x_k^4 \right] + o \left(\frac{1}{n} \right)$$

$$\Rightarrow P \{ \mu_n = k \} = \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\sqrt{npq}} \cdot \exp \left\{ \theta - \frac{1}{2} x_k^2 + \frac{q-p}{6\sqrt{npq}} (x_k^3 - 3x_k) + \frac{1}{12npq} \left[3(q^2 + p^2) x_k^2 - (q^3 + p^3) x_k^4 \right] + o \left(\frac{1}{n} \right) \right\}$$

$$= \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\sqrt{npq}} \cdot e^\theta \cdot e^{-\frac{1}{2} x_k^2} \cdot \exp \left[\frac{q-p}{6\sqrt{npq}} (x_k^3 - 3x_k) \right] \cdot \exp \left\{ \frac{1}{12npq} \left[3(q^2 + p^2) x_k^2 - (q^3 + p^3) x_k^4 \right] + o \left(\frac{1}{n} \right) \right\}$$

展开式

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$

$$= \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\sqrt{npq}} \cdot e^{-\frac{1}{2} x_k^2} \cdot \left[1 + \frac{q-p}{6\sqrt{npq}} (x_k^3 - 3x_k) + o \left(\frac{1}{n} \right) \right]$$

$$\Rightarrow P \{ \mu_n = k \} \div \left(\frac{1}{\sqrt{npq}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} x_k^2} \right) = 1 + \frac{q-p}{6\sqrt{npq}} (x_k^3 - 3x_k) + o \left(\frac{1}{n} \right) \longrightarrow 1 \quad (n \rightarrow \infty)$$

三、De Moivre - Laplace 极限定理

积分极限定理的证明:

$$P \left\{ a \leq \frac{\mu_n - np}{\sqrt{npq}} < b \right\} = P \left\{ np + a\sqrt{npq} \leq \mu_n < np + b\sqrt{npq} \right\}$$

k_2 为小于 $np + b\sqrt{npq}$ 的最大整数

k_1 为不小于 $np + a\sqrt{npq}$ 的最小整数

$$= \sum_{k=k_1}^{k_2} P \{ \mu_n = k \}$$

局部极限定理

$$P \{ \mu_n = k \} \div \left(\frac{1}{\sqrt{npq}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_k^2} \right) = 1 + \frac{q-p}{6\sqrt{npq}} (x_k^3 - 3x_k) + o\left(\frac{1}{n}\right) \rightarrow 1 \quad (n \rightarrow \infty)$$

$\forall \varepsilon > 0$

$$P \{ \mu_n = k \} = \frac{1}{\sqrt{npq}} \cdot \left[\varphi(x_k) + \varepsilon_k \right], \quad |\varepsilon_k| < \varepsilon, \quad k = k_1, k_1 + 1, \dots, k_2$$

$$\Rightarrow P \left\{ a \leq \frac{\mu_n - np}{\sqrt{npq}} < b \right\} = \sum_{k=k_1}^{k_2} P \{ \mu_n = k \} = \sum_{k=k_1}^{k_2} \left[\varphi(x_k) \cdot \frac{1}{\sqrt{npq}} \right] + \sum_{k=k_1}^{k_2} \frac{\varepsilon_k}{\sqrt{npq}}$$

定理: 若 μ_n 是 n 次 Bernoulli 试验中事件 A 出现的次数, $0 < p < 1$, 则对任意有限区间 $[a, b]$:

① 当 $a \leq x_k \equiv \frac{k - np}{\sqrt{np(1-p)}} \leq b$ 及 $n \rightarrow \infty$ 时, 一致地有

局部极限定理

$$P \{ \mu_n = k \} \div \left(\frac{1}{\sqrt{np(1-p)}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_k^2} \right) \rightarrow 1$$

② 当 $n \rightarrow \infty$ 时一致地有

积分极限定理

$$P \left\{ a \leq \frac{\mu_n - np}{\sqrt{np(1-p)}} < b \right\} \rightarrow \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \quad (-\infty < x < +\infty)$$



三、De Moivre - Laplace 极限定理

积分极限定理的证明:

定理: 若 μ_n 是 n 次 Bernoulli 试验中事件 A 出现的次数, $0 < p < 1$, 则对任意有限区间 $[a, b]$:

① 当 $a \leq x_k \equiv \frac{k - np}{\sqrt{np(1-p)}} \leq b$ 及 $n \rightarrow \infty$ 时, 一致地有
局部极限定理

$$P\{\mu_n = k\} \div \left(\frac{1}{\sqrt{np(1-p)}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_k^2} \right) \rightarrow 1$$

② 当 $n \rightarrow \infty$ 时一致地有
积分极限定理

$$P\left\{a \leq \frac{\mu_n - np}{\sqrt{np(1-p)}} < b\right\} \rightarrow \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \quad (-\infty < x < +\infty)$$



$$\left| \sum_{k=k_1}^{k_2} \frac{\varepsilon_k}{\sqrt{npq}} \right| \leq \frac{1}{\sqrt{npq}} \sum_{k=k_1}^{k_2} |\varepsilon_k| \leq \frac{1}{\sqrt{npq}} \sum_{k=k_1}^{k_2} \varepsilon = \frac{(k_2 - k_1 + 1) \varepsilon}{\sqrt{npq}}$$

$\rightarrow 0 \quad (n \rightarrow \infty)$

$$\left. \begin{array}{l} k_1 \text{ 为不小于 } np + a\sqrt{npq} \text{ 的最小整数} \\ k_2 \text{ 为小于 } np + b\sqrt{npq} \text{ 的最大整数} \end{array} \right\} \Rightarrow a \leq x_{k_1} \equiv \frac{k_1 - np}{\sqrt{npq}} < x_{k_1+1} \equiv \frac{k_1 + 1 - np}{\sqrt{npq}} < \dots < x_{k_2} \equiv \frac{k_2 - np}{\sqrt{npq}} \leq b$$

$x_{k_1}, x_{k_1+1}, \dots, x_{k_2}$ 是区间 $[a, b]$ 的一个划分

$$\Rightarrow \lim_{n \rightarrow \infty} \sum_{k=k_1}^{k_2} \left[\varphi(x_k) \cdot \frac{1}{\sqrt{npq}} \right] = \int_a^b \varphi(x) dx = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

$$\Rightarrow P\left\{a \leq \frac{\mu_n - np}{\sqrt{npq}} < b\right\} = \sum_{k=k_1}^{k_2} P\{\mu_n = k\} = \sum_{k=k_1}^{k_2} \left[\varphi(x_k) \cdot \frac{1}{\sqrt{npq}} \right] + \sum_{k=k_1}^{k_2} \frac{\varepsilon_k}{\sqrt{npq}}$$

四、De Moivre - Laplace 极限定理的应用

● 推导 Bernoulli 大数定律

积分极限定理

$$\frac{\mu_n - np}{\sqrt{npq}} \sim N(0, 1) \quad (n \rightarrow \infty)$$

$$\Rightarrow P \left\{ \left| \frac{\mu_n}{n} - p \right| < \varepsilon \right\} = P \left\{ \left| \frac{\mu_n - np}{\sqrt{npq}} \right| < \frac{\sqrt{n}}{\sqrt{pq}} \cdot \varepsilon \right\} = P \left\{ -\frac{\varepsilon\sqrt{n}}{\sqrt{pq}} < \frac{\mu_n - np}{\sqrt{npq}} < \frac{\varepsilon\sqrt{n}}{\sqrt{pq}} \right\}$$

$$\Rightarrow \lim_{n \rightarrow \infty} P \left\{ \left| \frac{\mu_n}{n} - p \right| < \varepsilon \right\} = \lim_{n \rightarrow \infty} \int_{-\frac{\varepsilon\sqrt{n}}{\sqrt{pq}}}{\frac{\varepsilon\sqrt{n}}{\sqrt{pq}}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

$\forall \varepsilon > 0$, 只要 n 足够大, $\frac{\varepsilon\sqrt{n}}{\sqrt{pq}} \rightarrow \infty$

$$= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 1$$

定理: (Bernoulli 大数定律)

设 μ_n 是 n 重 Bernoulli 试验中事件 A 发生的次数, $p = P\{A\}$, 则 $\forall \varepsilon > 0$,

$$\lim_{n \rightarrow \infty} P \left\{ \left| \frac{\mu_n}{n} - p \right| < \varepsilon \right\} = 1 \quad \text{或} \quad \lim_{n \rightarrow \infty} P \left\{ \left| \frac{\mu_n}{n} - p \right| \geq \varepsilon \right\} = 0$$

Chebyshev 不等式: 设 ξ 为具有有限均值 $E(\xi)$ 与有限方差 $D(\xi)$ 的随机变量, 则 $\forall \varepsilon > 0$ 有

$$P \left\{ |\xi - E(\xi)| \geq \varepsilon \right\} \leq \frac{D(\xi)}{\varepsilon^2}$$

四、De Moivre - Laplace 极限定理的应用

- 用频率估计概率时的计算问题

积分极限定理 \rightarrow $\frac{\mu_n - np}{\sqrt{npq}} \sim N(0, 1) \quad (n \rightarrow \infty)$

$$\Rightarrow P \left\{ \left| \frac{\mu_n}{n} - p \right| < \varepsilon \right\} = P \left\{ \left| \frac{\mu_n - np}{\sqrt{npq}} \right| < \frac{\sqrt{n}}{\sqrt{pq}} \cdot \varepsilon \right\} = P \left\{ -\frac{\varepsilon\sqrt{n}}{\sqrt{pq}} < \frac{\mu_n - np}{\sqrt{npq}} < \frac{\varepsilon\sqrt{n}}{\sqrt{pq}} \right\}$$

当 n 很大时 \rightarrow $\approx \Phi \left(\frac{\varepsilon\sqrt{n}}{\sqrt{pq}} \right) - \Phi \left(-\frac{\varepsilon\sqrt{n}}{\sqrt{pq}} \right) = 2\Phi \left(\frac{\varepsilon\sqrt{n}}{\sqrt{pq}} \right) - 1$

- ▶ 该式可用来解决很多计算问题，现举例如下.

四、De Moivre - Laplace 极限定理的应用

● 用频率估计概率时的计算问题 $\implies P \left\{ \left| \frac{\mu_n}{n} - p \right| < \varepsilon \right\} \approx 2\Phi \left(\frac{\varepsilon\sqrt{n}}{\sqrt{pq}} \right) - 1$ (当 n 很大时)

▶ 例：Puffon 掷 (均匀) 硬币试验中，掷硬币进行了 4040 次，硬币出现正面向上 2048 次，求正面出现的频率与概率的偏差不超过试验中出现的偏差的概率。

均匀硬币 $\rightarrow p = \frac{1}{2}$ 试验中的偏差 $\rightarrow \varepsilon = \frac{2048}{4040} - \frac{1}{2} \approx 0.00693$

$n = 4040$ 足够大 $\rightarrow P \left\{ \left| \frac{\mu_n}{4040} - \frac{1}{2} \right| < 0.00693 \right\} \approx 2\Phi \left(\frac{0.00693 \times \sqrt{4040}}{\sqrt{\frac{1}{2} \times \frac{1}{2}}} \right) - 1 \approx 2\Phi(0.8809554) - 1 = 0.621658$

```
2 * pnorm(0.8809554, 0, 1) - 1
```

```
> 2 * pnorm(0.8809554, 0, 1) - 1
[1] 0.621658
```

四、De Moivre - Laplace 极限定理的应用

● 用频率估计概率时的计算问题 $\implies P \left\{ \left| \frac{\mu_n}{n} - p \right| < \varepsilon \right\} \approx 2\Phi \left(\frac{\varepsilon\sqrt{n}}{\sqrt{pq}} \right) - 1$ (当 n 很大时)

▶ 例：某品牌以往的市场占有率为 15%，现公司计划再作一次市场调查，要求市场占有率调查结果的误差小于 1% 的概率达到 95%，问至少要抽样调查多少个用户？

$p = 0.15$

$\varepsilon = 0.01$

结果误差小于 1% 的概率达到 95%

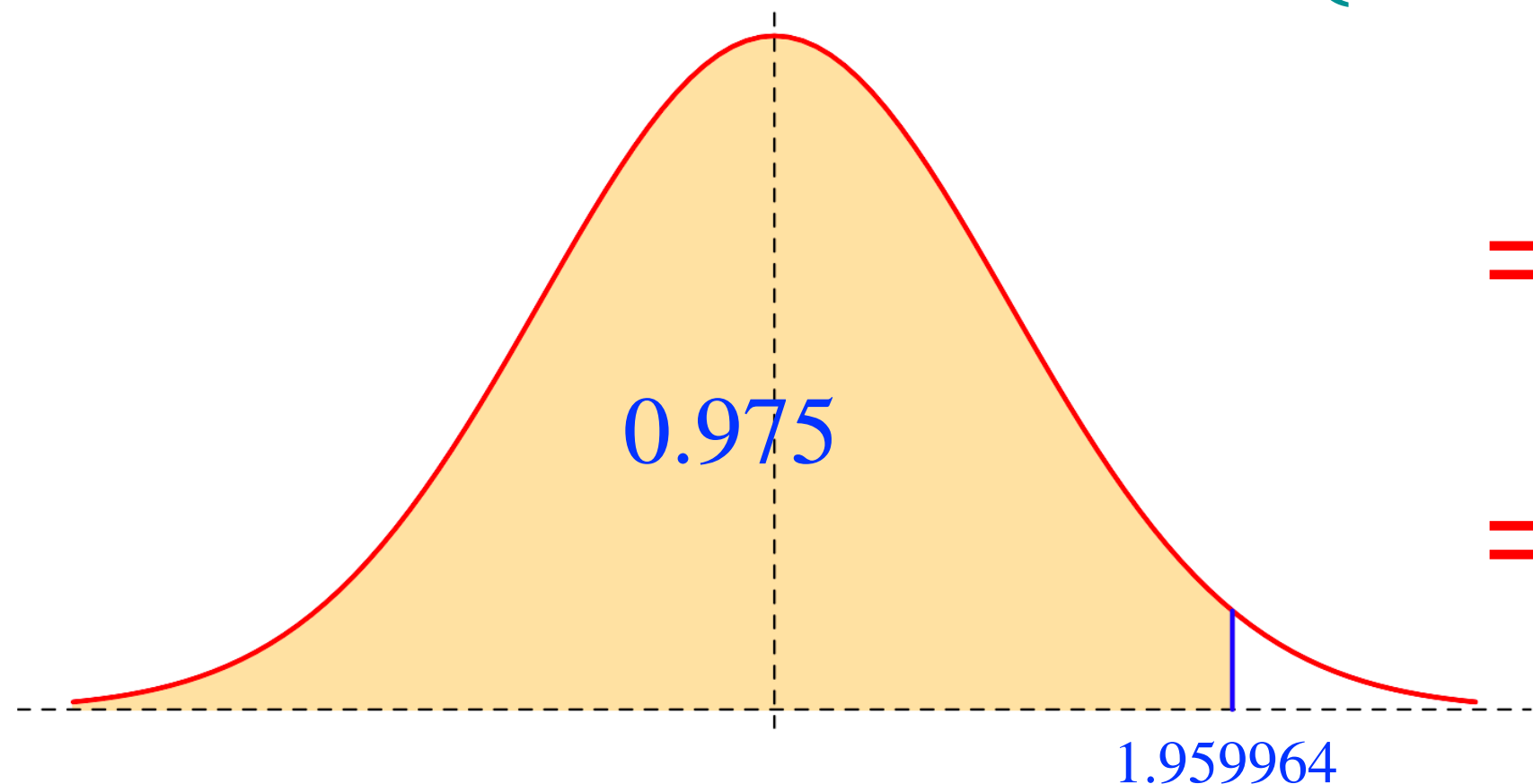
$\implies P \left\{ \left| \frac{\mu_n}{n} - p \right| < \varepsilon \right\} \approx 2\Phi \left(\frac{\varepsilon\sqrt{n}}{\sqrt{pq}} \right) - 1 \geq 0.95 \implies 2\Phi \left(\frac{0.01\sqrt{n}}{\sqrt{0.15 \times 0.85}} \right) - 1 \geq 0.95$

$\implies n \geq 4897.86$

$\implies \Phi \left(\frac{0.01\sqrt{n}}{\sqrt{0.15 \times 0.85}} \right) \geq 0.975$

$\implies n = 4898$

$\implies \frac{0.01\sqrt{n}}{\sqrt{0.15 \times 0.85}} \geq 1.959964$



`qnorm(0.975, 0, 1)`

`> qnorm(0.975, 0, 1)`
`[1] 1.959964`

四、De Moivre - Laplace 极限定理的应用

● 用频率估计概率时的计算问题 $\implies P \left\{ \left| \frac{\mu_n}{n} - p \right| < \varepsilon \right\} \approx 2\Phi \left(\frac{\varepsilon\sqrt{n}}{\sqrt{pq}} \right) - 1$ (当 n 很大时)

▶ 例：某品牌以往的市场占有率为 15%，现公司计划再作一次市场调查，要求市场占有率调查结果的可靠概率依然是 95%，但预算只允许调查 2000 户，问调查的误差是多少？

$p = 0.15$

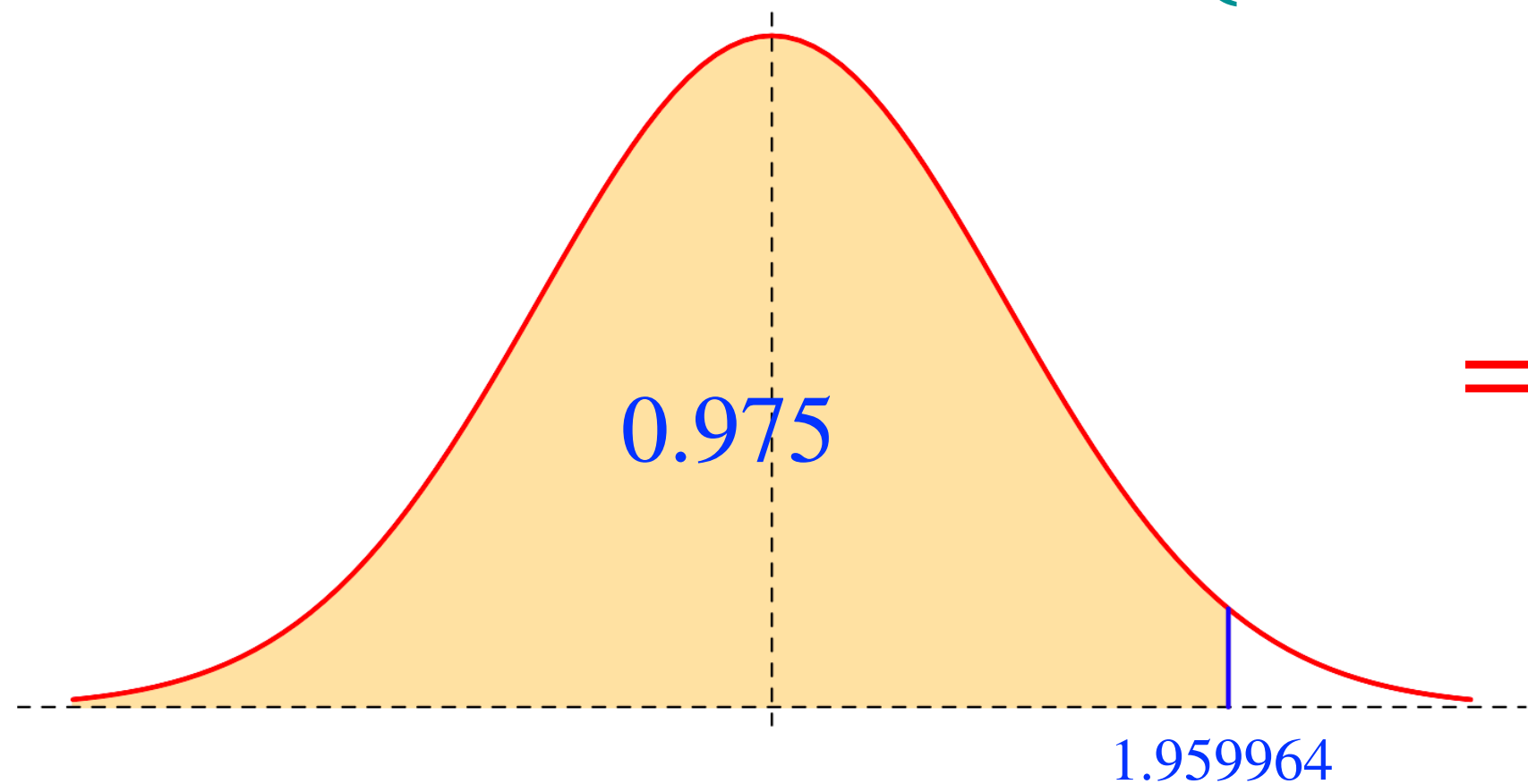
$n = 2000$

调查结果的可靠概率为 95%

$P \left\{ \left| \frac{\mu_n}{n} - p \right| < \varepsilon \right\} \approx 2\Phi \left(\frac{\varepsilon\sqrt{n}}{\sqrt{pq}} \right) - 1 \geq 0.95 \implies 2\Phi \left(\frac{\varepsilon\sqrt{2000}}{\sqrt{0.15 \times 0.85}} \right) - 1 \geq 0.95$

$\implies \Phi \left(\frac{\varepsilon\sqrt{2000}}{\sqrt{0.15 \times 0.85}} \right) \geq 0.975$

$\implies \frac{\varepsilon\sqrt{2000}}{\sqrt{0.15 \times 0.85}} \geq 1.959964$



$\implies \varepsilon \geq 0.01564906$

四、De Moivre - Laplace 极限定理的应用

- 局部极限定理在二项分布计算中的应用

局部极限定理 $\rightarrow P\{\mu_n = k\} \div \left(\frac{1}{\sqrt{np(1-p)}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_k^2} \right) \rightarrow 1 \quad (n \rightarrow \infty)$

$\Rightarrow \binom{n}{k} p^k q^{n-k} \div \left(\frac{1}{\sqrt{npq}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{k-np}{\sqrt{npq}}\right)^2} \right) \rightarrow 1 \quad (n \rightarrow \infty)$

$\Rightarrow \binom{n}{k} p^k q^{n-k} \approx \frac{1}{\sqrt{npq}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{k-np}{\sqrt{npq}}\right)^2} = \frac{1}{\sqrt{npq}} \cdot \varphi\left(\frac{k-np}{\sqrt{npq}}\right) \quad (\text{当 } n \text{ 很大时})$

