# 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 |
|---|---|---|
| $\Omega$ | 空间或全集 | 样本空间或必然事件 |
| $\Phi$ | 空集 | 不可能事件 |
| $\omega$ | 元素 | 样本点 |
| $A$ | 子集 | 随机事件 |
| $\omega \in A$ | $\omega$ 是 $A$ 的元素 | 事件 $A$ 包含样本点 $\omega$ |
| $A \subset B$ | $A$ 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：$\left(\Omega, \mathscr{F}, P\right)$
    - 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 $\sigma$ 域，保证事件对交、并、逆、差作可列次运算的封闭性.
    - 在这种 $\sigma$ 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：$P(A\,|\,B)=\dfrac{P(AB)}{P(B)}.$
    - 乘法公式：$P(A_1A_2\cdots A_n)=P(A_1)\cdot P\left(A_2\middle|A_1\right)\cdot P\left(A_3\middle|A_1A_2\right)\cdot\cdots\cdot P\left(A_n\middle|A_1A_2\cdots A_{n-1}\right)$
    - 全概率公式：$P(B)=\sum\limits_{i=1}^{\infty}P(A_i)\cdot P\left(B\middle|A_i\right)$
    - Bayes 公式：$P\left(A_k\middle|B\right)=\dfrac{P(A_k)\cdot P\left(B\middle|A_k\right)}{\sum\limits_{i=1}^{\infty}P(A_i)\cdot P\left(B\middle|A_i\right)}$

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

# 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 $\left(\Omega, \mathscr{F}, P\right)$，其中

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

  - 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 $\xi$ 是概率空间 $(\Omega, \mathscr{F}, P)$ 中定义域为 $\Omega$、取值于 $\mathbb{R}$ 的一个单值 Borel 函数.

    - 分布函数 $F(x) = P\left\{\xi(\omega) < x\right\}$，$-\infty < x < \infty$ 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：$\begin{cases} \text{离散型：} & P\{x \in D\} = \sum\limits_{x_i \in D} p_i, \quad P\left\{(x, y) \in D\right\} = \sum\limits_{(x_i, y_j) \in D} p_{ij} \\ \text{连续型：} & P\{x \in D\} = \int\limits_D p(x)\mathrm{d}x, \quad P\left\{(x, y) \in D\right\} = \iint\limits_D p(x, y)\mathrm{d}x\mathrm{d}y \end{cases}$

# 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

$$(\xi,\eta)\sim N\left(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho\right)\sim N\left(\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix},\begin{pmatrix}\sigma_1^2 & \sigma_1\sigma_2\rho\\ \sigma_1\sigma_2\rho & \sigma_2^2\end{pmatrix}\right)\Longrightarrow\begin{cases}\xi\sim N\left(\mu_1,\sigma_1^2\right),\qquad \eta\sim N\left(\mu_2,\sigma_2^2\right)\\ \left(\eta\,\middle|\,\xi=x\right)\sim N\left(\mu_2+\rho\dfrac{\sigma_2}{\sigma_1}\left(x-\mu_1\right),\sigma_2^2\left(1-\rho^2\right)\right)\\ \left(\xi\,\middle|\,\eta=y\right)\sim N\left(\mu_1+\rho\dfrac{\sigma_1}{\sigma_2}\left(y-\mu_2\right),\sigma_1^2\left(1-\rho^2\right)\right)\\ \xi\text{ 与 }\eta\text{ 相互独立}\Longleftrightarrow\rho=0\end{cases}$$

# 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：

$$\begin{cases} \eta = g(\xi): \text{直接法 } F_\eta(y) = P\{\eta < y\} = \displaystyle\int_{g(x)<y} p_\xi(x)\mathrm{d}x \\ \zeta = g(\xi, \eta): \text{直接法 } F_\zeta(z) = P\{\zeta < z\} = \displaystyle\iint_{g(x,\, y)<z} p(x,\, y)\mathrm{d}x\mathrm{d}y \\ \qquad\qquad \text{和 (卷积公式)、差、商、积、 max , min 的公式} \\ \begin{cases} \zeta_1 = g_1(\xi, \eta) \\ \zeta_2 = g_2(\xi, \eta) \end{cases}: \text{变换法 } \begin{cases} u = g_1(x, y) \\ v = g_2(x, y) \end{cases} \Longrightarrow \begin{cases} x = h_1^{-1}(u, v) \\ y = h_2^{-1}(u, v) \end{cases} \Longrightarrow J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} \end{cases}$$

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

# 已学知识点

- 第四章 数字特征与特征函数

  - 数学期望

    - 定义： $E\left(\xi\right)=\begin{cases}\sum\limits_i x_i\cdot p_i\,, & \text{离散型}\\ \int\limits_{-\infty}^{\infty} x\cdot p_\xi(x)\mathrm{d}x\,, & \text{连续型}\end{cases}$ ，绝对收敛.

    - 随机变量函数的数学期望： $E\left[g(\xi)\right]=\begin{cases}\sum\limits_i g\left(x_i\right)\cdot p_i\,, & \text{离散型}\\ \int\limits_{-\infty}^{\infty} g(x)\cdot p_\xi(x)\mathrm{d}x\,, & \text{连续型}\end{cases}$

$$E\left[g\left(\xi,\,\eta\right)\right]=\begin{cases}\sum\limits_i\sum\limits_j g\left(x_i,\,y_j\right)\cdot p_{ij}\,, & \text{离散型}\\ \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} g\left(x,\,y\right)\cdot p_{(\xi,\,\eta)}(x,\,y)\mathrm{d}x\mathrm{d}y\,, & \text{连续型}\end{cases}$$

数学期望的性质：

① $E(c)=c$

② $a\le\xi\le b \implies a\le E(\xi)\le b$

③ $\xi\le\eta \implies E(\xi)\le E(\eta)$

④ $E\left(a\xi+b\eta\right)=a\cdot E\left(\xi\right)+b\cdot E\left(\eta\right)$

⑤ $\xi,\,\eta$ 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：$D(\xi)=E(\xi^2)-\left[E(\xi)\right]^2$

  - Chebyshev 不等式：

$$P\left\{\left|\xi-E(\xi)\right|\geq\varepsilon\right\}\leq\frac{D(\xi)}{\varepsilon^2}\iff P\left\{\left|\xi-E(\xi)\right|<\varepsilon\right\}\geq1-\frac{D(\xi)}{\varepsilon^2}$$

  - 协方差：

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

  - 和的方差：$D(X\pm Y)=D(X)+D(Y)\pm2\operatorname{cov}(X,Y)$

方差的性质：

① 设 $C$ 是常数，则 $D(C)=0$.

② 设 $\xi$ 是随机变量，$c$ 为常数，则 $D(\xi+c)=D(\xi)$.

③ 设 $\xi$ 是随机变量，$c$ 为常数，则 $D(c\xi)=c^2D(\xi)$.

④ 若 $c\neq E(\xi)$，则 $D(\xi)<E\left[(\xi-c)^2\right]$.

⑤ $D(\xi)=0\iff P\{\xi=C\}=1$.

⑥ 若 $\xi$ 与 $\eta$ 独立，则 $D(\xi\pm\eta)=D(\xi)+D(\eta)$.

协方差的性质：

① $\operatorname{cov}(X,X)=D(X)$.

② $\operatorname{cov}(X,Y)=\operatorname{cov}(Y,X)$

③ $\operatorname{cov}(aX,bY)=ab\cdot\operatorname{cov}(X,Y)$，$a,b$ 为常数.

④ $\operatorname{cov}(X_1+X_2,Y)=\operatorname{cov}(X_1,Y)+\operatorname{cov}(X_2,Y)$.

⑤ 若 $X,Y$ 相互独立，则 $\operatorname{cov}(X,Y)=0$.

⑥ 若 $C$ 为常数，则 $\operatorname{cov}(X,C)=0$.

# 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 | $\xi \sim \mathrm{Exp}(\lambda)$ | $E(\xi) = \frac{1}{\lambda}$ | $D(\xi) = \frac{1}{\lambda^2}$ |
| $\chi^2$ 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：$\rho_{\xi\eta} \triangleq \dfrac{\operatorname{cov}(\xi,\eta)}{\sqrt{D(\xi)}\sqrt{D(\eta)}}$

  - Cauchy-Schwarz 不等式：$\left|E(\xi\eta)\right|^2 \le E(\xi^2)\cdot E(\eta^2)$

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

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

相关系数的性质：

① $\left|\rho_{\xi\eta}\right| \le 1.$

② 等价命题：$\operatorname{cov}(\xi,\eta)=0$；$\xi$ 与 $\eta$ 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 $\xi$ 与 $\eta$ 独立，则 $\xi$ 与 $\eta$ 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：

$$\begin{cases} E\left(\eta \,\middle|\, \xi = x_i\right) = \sum\limits_j y_j \cdot P\left\{\eta = y_j \,\middle|\, \xi = x_i\right\} = \sum\limits_j y_j \cdot \dfrac{p\left(x_i, y_j\right)}{p_1\left(x_i\right)} \\ E\left(\xi \,\middle|\, \eta = y_j\right) = \sum\limits_i x_i \cdot P\left\{\xi = x_i \,\middle|\, \eta = y_j\right\} = \sum\limits_i x_i \cdot \dfrac{p\left(x_i, y_j\right)}{p_2\left(y_j\right)} \end{cases}$$

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

  - 连续型随机变量的条件期望：$E\left(\eta \,\middle|\, \xi = x\right) = \displaystyle\int_{-\infty}^{+\infty} y \cdot p\left(y \,|\, x\right) \mathrm{d}y$

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

  - 重期望公式：$E(\xi) = E\left[E\left(\xi \,\middle|\, \eta\right)\right]$.

# 已学知识点

- 第四章 数字特征与特征函数

  - 复随机变量：若 $\xi$ 与 $\eta$ 是概率空间 $(\Omega, \mathscr{F}, P)$ 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：$\mathrm{e}^{ix}=\cos x+i\sin x$.
  - 特征函数：$f_\xi(t)=E\left(\mathrm{e}^{it\xi}\right)=\int_{-\infty}^{+\infty}\mathrm{e}^{itx}\mathrm{d}F_\xi(x)=\begin{cases}\sum_{j=1}^{\infty}p_j\cdot\mathrm{e}^{itx_j}, & \text{离散型}\\ \int_{-\infty}^{+\infty}\mathrm{e}^{itx}\cdot p_\xi(x)\,\mathrm{d}x, & \text{连续型}\end{cases}$

**特征函数的性质：**

① 特征函数 $f(t)$ 有如下性质：$|f(t)|\le f(0)=1$，$f(-t)=\overline{f(t)}$.

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 $t_1, t_2, \ldots, t_n$、任意复数 $\lambda_1, \lambda_2, \ldots, \lambda_n$，均有 $\sum_{k=1}^{n}\sum_{j=1}^{n}f\left(t_k-t_j\right)\lambda_k\overline{\lambda_j}\ge 0$

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 $\xi$ 的 $n$ 阶矩存在，则它的特征函数可微分 $n$ 次，且当 $k\le n$ 时：$f^{(k)}(0)=i^kE\left(\xi^k\right)$.

⑥ 设 $\eta=a\xi+b$，这里 $a, b$ 为常数，则 $f_\eta(t)=\mathrm{e}^{ibt}\cdot f_\xi(at)$.

**常见分布的特征函数：**

① 退化分布 $I_c(x)$：$f(t)=\mathrm{e}^{itc}$

② 二项分布 $b(n, p)$：$f(t)=\left(p\mathrm{e}^{it}+1-p\right)^n$

③ Poisson 分布：$f(t)=\mathrm{e}^{\lambda\left(\mathrm{e}^{it}-1\right)}$

④ $\Gamma$ 分布：$f(t)=\left(1-\dfrac{it}{\lambda}\right)^{-r}$

⑤ 正态分布：$f(t)=\mathrm{e}^{i\mu t-\frac{1}{2}\sigma^2t^2}$

# 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：$f_\xi(t) \triangleq E\left(\mathrm{e}^{it\xi}\right) = \int_{-\infty}^{+\infty} \mathrm{e}^{itx}\,\mathrm{d}F_\xi(x) = \begin{cases} \sum_{j=1}^{\infty} p_j \cdot \mathrm{e}^{itx_j}, & \text{离散型} \\ \int_{-\infty}^{+\infty} \mathrm{e}^{itx} \cdot p_\xi(x)\,\mathrm{d}x, & \text{连续型} \end{cases}$
    - (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 $F'(x) = p(x) = \dfrac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{e}^{-itx} f(t)\,\mathrm{d}t$ .

  - 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、$\Gamma$ 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.

# 已学知识点

- 第五章 极限定理

  - 大数定律：若随机变量序列 $\{\xi_k\}$ 满足 $\lim\limits_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{k=1}^{n}\xi_k - a\right| < \varepsilon\right\} = 1$，则称它服从大数定律.

  - 中心极限定理：若独立的随机变量序列 $\{\xi_k\}$ 满足

$$\lim_{n\to\infty} P\left\{\frac{\sum\limits_{i=1}^{n}\xi_i - \sum\limits_{i=1}^{n}E(\xi_i)}{\sqrt{\sum\limits_{i=1}^{n}D(\xi_i)}} < x\right\} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} \mathrm{e}^{-\frac{t^2}{2}}\mathrm{d}t,$$

  则称它服从中心极限定理.

  - 问题：寻找大数定律、中心极限定理成立的条件.

# 已学知识点

- 第五章 极限定理

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

- Bernoulli 大数定律：设 $\mu_n$ 是 $n$ 重 Bernoulli 试验中事件 $A$ 发生的次数，$p = P\{A\}$，则 $\forall \varepsilon > 0$,

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

- Chebyshev 大数定律：设 $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ 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，且它们有公共的上界 $D(\xi_1) \leq C$ , $D(\xi_2) \leq C$ , …, $D(\xi_n) \leq C$ , …, 则 $\forall \varepsilon > 0$,

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

- Markov 大数定律：对于随机变量序列 $\xi_1, \xi_2, \ldots, \xi_n, \ldots$，若 $\frac{1}{n^2}\cdot D\left(\sum_{k=1}^{n}\xi_k\right) \longrightarrow 0$，则 $\forall \varepsilon > 0$,

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

# 已学知识点

## 第五章 极限定理

- 独立同分布下的大数定律：设 $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ 是独立同分布的随机变量序列，且 $E(\xi_i) = \mu$，$D(\xi_i) = \sigma^2, i = 1, 2, \ldots$则 $\forall \varepsilon > 0$,

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

- 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 $\mu_n$ 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

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

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.

# 已学知识点

- 第五章 极限定理

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

▸ De Moivre - Laplace 定理：若 $\mu_n$ 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 $a \le x_k \equiv \dfrac{k-np}{\sqrt{np(1-p)}} \le b$ 及 $n \to \infty$ 时，一致地有 $P\{\mu_n = k\} \div \left(\dfrac{1}{\sqrt{np(1-p)}} \cdot \dfrac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_k^2}\right) \longrightarrow 1.$

局部极限定理

② 当 $n \to \infty$ 时，一致地有 $P\left\{a \le \dfrac{\mu_n - np}{\sqrt{np(1-p)}} < b\right\} \longrightarrow \displaystyle\int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \qquad (-\infty < x < +\infty).$

积分极限定理

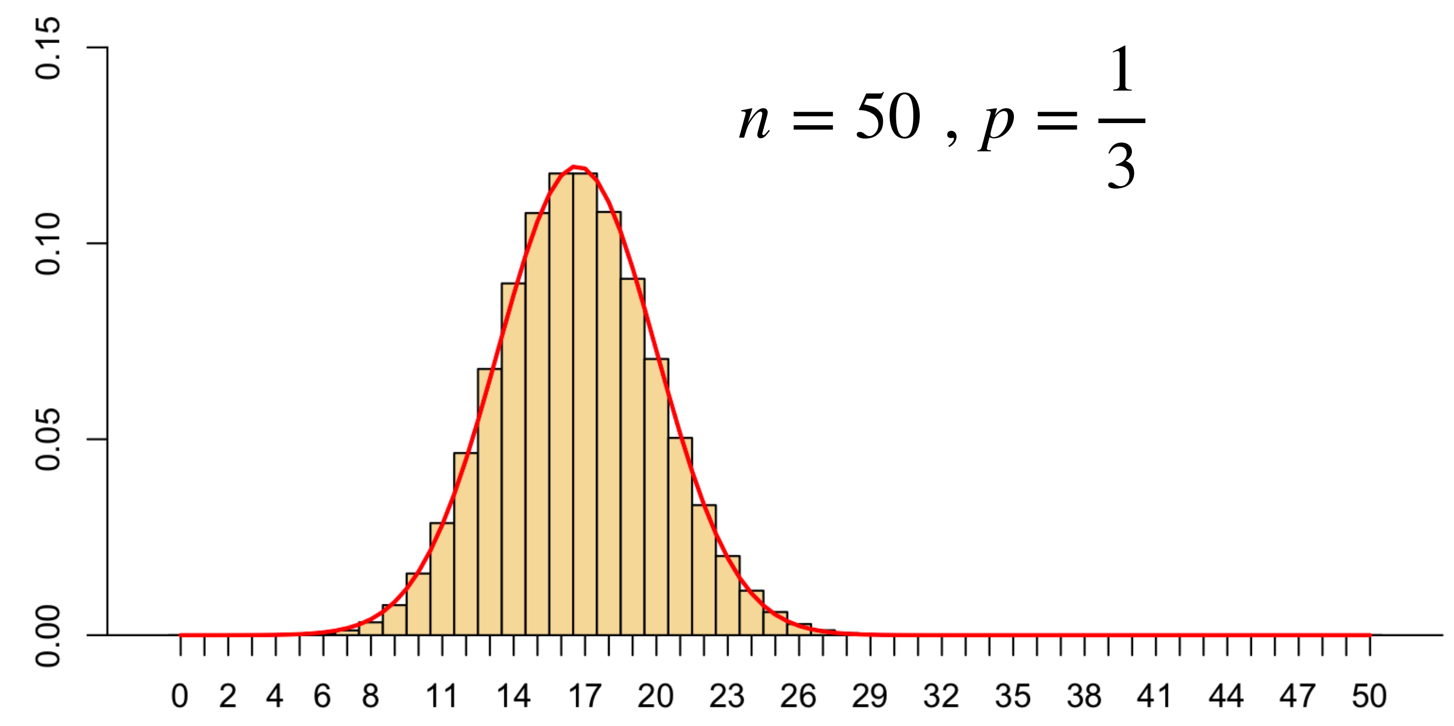
$$P\{\mu_n = k\} \approx \frac{1}{\sqrt{np(1-p)}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_k^2}, \quad x_k = \frac{k-np}{\sqrt{np(1-p)}}; \quad n \to \infty$$

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

▸ 应用：用频率估计概率时的计算问题

# 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



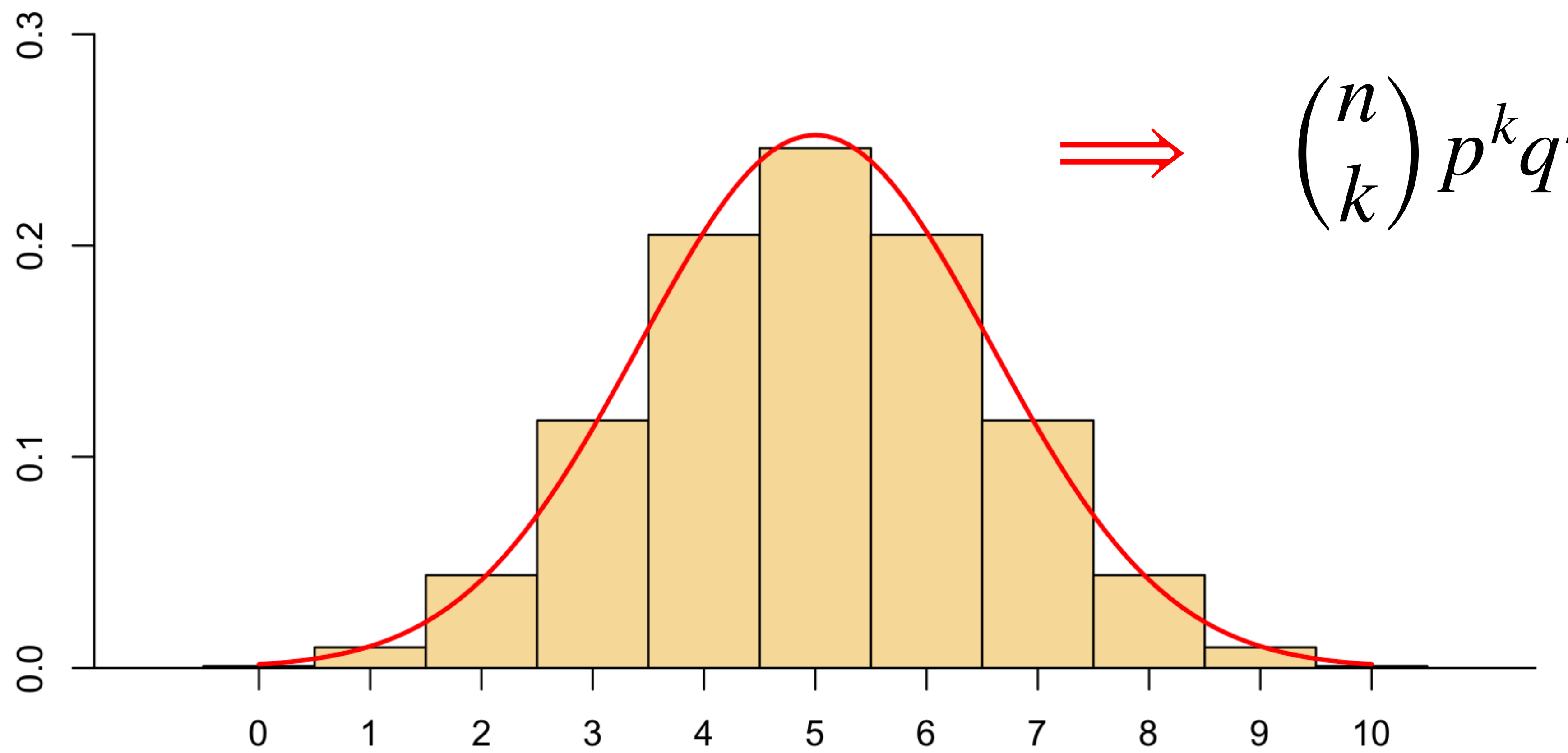

局部极限定理

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

$$\Longrightarrow\quad\binom{n}{k}p^kq^{n-k}\div\left(\frac{1}{\sqrt{npq}}\cdot\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}\left(\frac{k-np}{\sqrt{npq}}\right)^2}\right)\longrightarrow 1\qquad(n\to\infty)$$

$$\Longrightarrow\quad\binom{n}{k}p^kq^{n-k}\approx\frac{1}{\sqrt{npq}}\cdot\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}\left(\frac{k-np}{\sqrt{npq}}\right)^2}=\frac{1}{\sqrt{npq}}\cdot\varphi\left(\frac{k-np}{\sqrt{npq}}\right)$$

( 当 $n$ 很大时)




$n=10\ ,\ p=\dfrac{1}{2}$

# 5.2 收敛性

一、分布函数弱收敛

二、连续性定理

三、随机变量的收敛性

# 一、分布函数弱收敛

定理：若 $\mu_n$ 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 $a \le x_k \equiv \dfrac{k-np}{\sqrt{np(1-p)}} \le b$ 及 $n\to\infty$ 时，一致地有

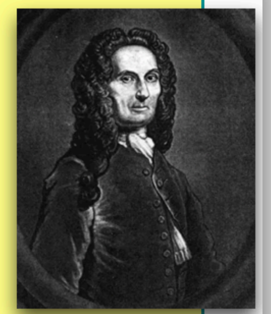
局部极限定理

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

② 当 $n\to\infty$ 时一致地有

积分极限定理

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




$$F_n(x) \triangleq P\left\{ \frac{\mu_n - np}{\sqrt{npq}} < x \right\}$$

$$\Longrightarrow \quad \forall x \in (-\infty\,,\,+\infty)\ :\quad F_n(x) \longrightarrow \Phi(x) \qquad n\to\infty$$

▶ 例：考虑分布

$$F_n(x) = \begin{cases} 0\,, & x < \dfrac{1}{n} \\ 1\,, & x \ge \dfrac{1}{n} \end{cases}, \qquad n = 1, 2, \ldots$$

分布函数

退化分布

| $\xi_n$ | $\dfrac{1}{n}$ |
|---|---|
| $P\{\xi_n = x\}$ | 1 |

$$F_n(0) \equiv 0$$

$$F(x) = \begin{cases} 0\,, & x < 0 \\ 1\,, & x \ge 0 \end{cases}$$

分布函数

$$F(0) = 1$$

$x = 0$ 是不连续点

$n \longrightarrow \infty$ 时收敛于

极限分布

| $\xi$ | 0 |
|---|---|
| $P\{\xi = x\}$ | 1 |

$$\Longrightarrow \quad \lim_{n\to\infty} F_n(0) \ne F(0)$$

$\Longrightarrow$ 要求分布函数列在所有点都收敛于极限分布函数，过于严格了.

## 一、分布函数弱收敛

**定义**：对于分布函数列 $\{F_n(x)\}$，如果存在一个非降函数 $F(x)$，使得在 $F(x)$ 的每一连续点上都有

$$\lim_{n\to\infty} F_n(x) = F(x)$$

成立，则称 $F_n(x)$ 弱收敛于 $F(x)$，记为 $F_n(x) \xrightarrow{w} F(x)$.

- 这样得到的极限函数是一个有界的非降函数.
- 可作适当选择使得函数 $F(x)$ 是右连续的.
- 例：对于分布函数列 $F_n(x) = \begin{cases} 0, & x < n \\ 1, & x \geq n \end{cases}$， $n = 1, 2, \ldots$ $\Longrightarrow$ $\lim_{n\to\infty} F_n(x) = 0 = F(x)$ 不是分布函数
- 分布函数列的弱收敛极限不一定是一个分布函数.
- 如果 $F(x)$ 是一个分布函数，其右连续的性质保证了 $F(x)$ 在不连续点上的值完全由它在连续点集 $C_F$ 上的值唯一确定，此时，分布函数列的弱收敛极限是唯一的.

## 一、分布函数弱收敛

- 我们来讨论一个分布函数序列弱收敛到一个分布函数的充分必要条件.

证明过程略!

**引理**：设 $\{F_n(x)\}$ 是实变量 $x$ 的非降函数序列，$D$ 是 $\mathbb{R}^1$ 上的稠密集，若对于 $D$ 中的所有点，序列 $\{F_n(x)\}$ 收敛于 $F(x)$，则对 $F(x)$ 的一切连续点 $x$ 都有

$$\lim_{n\to\infty} F_n(x) = F(x)$$

有界数列 $\{x_n\}, \forall n \in \mathbb{N}$, 令 $\alpha_n = \inf\limits_{k\geq n}\{x_n\}$ ， $\beta_n = \sup\limits_{k\geq n}\{x_n\}$ $\Longrightarrow$ $\{\alpha_n\}$ 递增，$\{\beta_n\}$ 递减

下极限： $\underline{\lim}_{n\to\infty} x_n \triangleq \lim_{n\to\infty} \alpha_n = \sup\limits_{n\in\mathbb{N}} \inf\limits_{k\geq n} \{x_n\}$ ， 上极限： $\overline{\lim}_{n\to\infty} x_n \triangleq \lim_{n\to\infty} \beta_n = \inf\limits_{n\in\mathbb{N}} \sup\limits_{k\geq n} \{x_n\}$

如果数列 $\{x_n\}$ 无下界，则 $\underline{\lim}_{n\to\infty} x_n = -\infty$ ， 如果数列 $\{x_n\}$ 无上界，则 $\overline{\lim}_{n\to\infty} x_n = +\infty$

$\Longrightarrow$ 对任何数列取上极限和下极限都是有意义的.

## 一、分布函数弱收敛

- 我们来讨论一个分布函数序列弱收敛到一个分布函数的充分必要条件.

**引理**：设 $\{F_n(x)\}$ 是实变量 $x$ 的非降函数序列，$D$ 是 $\mathbb{R}^1$ 上的稠密集，若对于 $D$ 中的所有点，序列 $\{F_n(x)\}$ 收敛于 $F(x)$，则对 $F(x)$ 的一切连续点 $x$ 都有

$$\lim_{n\to\infty} F_n(x) = F(x)$$

上、下极限的性质：

① $\underline{\lim}_{n\to\infty} x_n \le \overline{\lim}_{n\to\infty} x_n$

② $\lim_{n\to\infty} x_n = A \iff \underline{\lim}_{n\to\infty} x_n = \overline{\lim}_{n\to\infty} x_n = A$

③ $\overline{A}$ 为有界数列 $\{x_n\}$ 的上极限 $\iff \forall\varepsilon > 0$，(1) $\exists N > 0$：$n > N$ 时有 $x_n < \overline{A} + \varepsilon$；

(2) 存在子列 $\{x_{n_k}\}$ 使 $x_{n_k} > \overline{A} - \varepsilon$，

④ $\underline{A}$ 为有界数列 $\{x_n\}$ 的下极限 $\iff \forall\varepsilon > 0$，(1) $\exists N > 0$：$n > N$ 时有 $x_n > \underline{A} - \varepsilon$；

(2) 存在子列 $\{x_{n_k}\}$ 使 $x_{n_k} < \underline{A} + \varepsilon$，$k = 1, 2, \ldots$

## 一、分布函数弱收敛

- 我们来讨论一个分布函数序列弱收敛到一个分布函数的充分必要条件.

**引理**：设 $\{F_n(x)\}$ 是实变量 $x$ 的非降函数序列，$D$ 是 $\mathbb{R}^1$ 上的稠密集，若对于 $D$ 中的所有点，序列 $\{F_n(x)\}$ 收敛于 $F(x)$，则对 $F(x)$ 的一切连续点 $x$ 都有

$$\lim_{n\to\infty} F_n(x) = F(x)$$

$\forall\, x \in \mathbb{R}^1$， 选择 $x' \in D,\ x'' \in D$, 使得 $x' \le x \le x''$

$$\xrightarrow{F_n(x)\ \text{是非降函数}} F_n(x') \le F_n(x) \le F_n(x'') \implies F(x') \le \varliminf_{n\to\infty} F_n(x) \le \varlimsup_{n\to\infty} F_n(x) \le F(x'')$$

$$\xrightarrow{D\ \text{是}\ \mathbb{R}^1\ \text{上的稠密集}} F(x-0) \le \varliminf_{n\to\infty} F_n(x) \le \varlimsup_{n\to\infty} F_n(x) \le F(x+0)$$

$$\xrightarrow{\text{若}\ x\ \text{是}\ F(x)\ \text{的连续点}} F(x-0) = F(x+0) = F(x) \implies \lim_{n\to\infty} F_n(x) = F(x)$$

## 一、分布函数弱收敛

证明过程略!

**定理：(Helly 第一定理)**

任一一致有界的非降函数序列 $\{F_n(x)\}$ 中，必有一子序列 $\left\{F_{n_k}(x)\right\}$ 弱收敛于某一有界的非降函数 $F(x)$.

定理：(Helly 第一定理)

任一一致有界的非降函数序列 $\{F_n(x)\}$ 中，必有一子序列 $\{F_{n_k}(x)\}$ 弱收敛于某一有界的非降函数 $F(x)$.

## 一、分布函数弱收敛

任取 $\mathbb{R}^1$ 上处处稠密的一个可数点集 $D$，在此，我们取 $D$ 为全体有理数的点集，记为

$$D=\{r_1,\ r_2,\ \ldots,\ r_m,\ \ldots\}$$

对数列 $\{F_n(r_1)\}_{n=1}^{\infty}$：$F_1(r_1), F_2(r_1),\ \ldots,\ F_n(r_1),\ \ldots$，这是一个单调有界数列,

必有收敛子列 $\{F_{1,\,n}(r_1)\}_{n=1}^{\infty}$：$F_{1,\,1}(r_1), F_{1,\,2}(r_1),\ \ldots,\ F_{1,\,n}(r_1),\ \ldots \xrightarrow{\text{收敛于某个数}} G(r_1)$

$$\Longrightarrow \lim_{n\to\infty} F_{1,\,n}(r_1)=G(r_1)$$

对数列 $\{F_{1,\,n}(r_2)\}_{n=1}^{\infty}$：$F_{1,\,1}(r_2), F_{1,\,2}(r_2),\ \ldots,\ F_{1,\,n}(r_2),\ \ldots$，这也是一个单调有界数列,

必有收敛子列 $\{F_{2,\,n}(r_2)\}_{n=1}^{\infty}$：$F_{2,\,1}(r_2), F_{2,\,2}(r_2),\ \ldots,\ F_{2,\,n}(r_2),\ \ldots \xrightarrow{\text{收敛于某个数}} G(r_2)$

$$\Longrightarrow \lim_{n\to\infty} F_{2,\,n}(r_2)=G(r_2),\quad \lim_{n\to\infty} F_{2,\,n}(r_1)=G(r_1)$$

定理：(Helly 第一定理)

任一一致有界的非降函数序列 $\{F_n(x)\}$ 中，必有一子序列 $\{F_{n_k}(x)\}$ 弱收敛于某一有界的非降函数 $F(x)$.

## 一、分布函数弱收敛

再对数列 $\{F_{2,n}(r_3)\}_{n=1}^{\infty}$： $F_{2,1}(r_3), F_{2,2}(r_3), \ldots, F_{2,n}(r_3), \ldots$，这也是一个单调有界数列，

必有收敛子列 $\{F_{3,n}(r_3)\}_{n=1}^{\infty}$： $F_{3,1}(r_3), F_{3,2}(r_3), \ldots, F_{3,n}(r_3), \ldots \xrightarrow{\text{收敛于某个数}} G(r_3)$

$$\Longrightarrow \lim_{n\to\infty} F_{3,n}(r_3) = G(r_3)\ ,\ \lim_{n\to\infty} F_{3,n}(r_2) = G(r_2)\ ,\ \lim_{n\to\infty} F_{3,n}(r_1) = G(r_1)$$

持续上述过程，可得数列 $\{F_{m,n}(x)\}_{n=1}^{\infty}$： $F_{m,1}(x), F_{m,2}(x), \ldots, F_{m,n}(x), \ldots$，它满足

$$\lim_{n\to\infty} F_{m,n}(r_k) = G(r_k)\ ,\quad k = 1, 2, \ldots, m$$

定理：(Helly 第一定理)

任一一致有界的非降函数序列 $\{F_n(x)\}$ 中，必有一子序列 $\{F_{n_k}(x)\}$ 弱收敛于某一有界的非降函数 $F(x)$.

## 一、分布函数弱收敛

于是，我们得到了数列 $\{F_n(x)\}_{n=1}^{\infty}$ 的如下子数列：

$$
\begin{array}{cccccc}
F_{1,1}(x), & F_{1,2}(x), & F_{1,3}(x), & \ldots, & F_{1,n}(x), & \ldots \\
F_{2,1}(x), & F_{2,2}(x), & F_{2,3}(x), & \ldots, & F_{2,n}(x), & \ldots \\
\vdots & \vdots & \vdots & \cdots & \vdots & \cdots \\
F_{m,1}(x), & F_{m,2}(x), & F_{m,3}(x), & \ldots, & F_{m,n}(x), & \ldots \\
\vdots & \vdots & \vdots & \cdots & \vdots & \cdots
\end{array}
$$

每一行都是前一行的子数列

$\lim\limits_{n\to\infty} F_{m,n}(r_k) = G(r_k)$

取对角线上的元素

构造新数列 $\{F_{n,n}(x)\}_{n=1}^{\infty}$ ：$F_{1,1}(x),\ F_{2,2}(x),\ F_{3,3}(x),\ \ldots,\ F_{n,n}(x),\ \ldots$

是 $F_{1,n}(x)$ 的子数列 $\Longrightarrow$ $\lim\limits_{n\to\infty} F_{n,n}(r_1) = G(r_1)$

是 $F_{2,n}(x)$ 的子数列 $\Longrightarrow$ $\lim\limits_{n\to\infty} F_{n,n}(r_2) = G(r_2)$

是 $F_{3,n}(x)$ 的子数列 $\Longrightarrow$ $\lim\limits_{n\to\infty} F_{n,n}(r_3) = G(r_3)$

$\Longrightarrow$ $\lim\limits_{n\to\infty} F_{n,n}(r_k) = G(r_k)$

任意有理数

定理：(Helly 第一定理)

任一一致有界的非降函数序列 $\{F_n(x)\}$ 中，必有一子序列 $\{F_{n_k}(x)\}$ 弱收敛于某一有界的非降函数 $F(x)$.

## 一、分布函数弱收敛

于是，我们得到了数列 $\{F_n(x)\}_{n=1}^{\infty}$ 的如下子数列：

$$\begin{array}{cccccc} F_{1,1}(x), & F_{1,2}(x), & F_{1,3}(x), & \dots, & F_{1,n}(x), & \dots \\ F_{2,1}(x), & F_{2,2}(x), & F_{2,3}(x), & \dots, & F_{2,n}(x), & \dots \\ \vdots & \vdots & \vdots & \cdots & \vdots & \cdots \\ F_{m,1}(x), & F_{m,2}(x), & F_{m,3}(x), & \dots, & F_{m,n}(x), & \dots \\ \vdots & \vdots & \vdots & \cdots & \vdots & \cdots \end{array}$$

每一行都是前一行的子数列

$$\lim_{n\to\infty} F_{m,n}(r_k) = G(r_k)$$

取对角线上的元素

构造新数列 $\{F_{n,n}(x)\}_{n=1}^{\infty}$ : $F_{1,1}(x), F_{2,2}(x), F_{3,3}(x), \dots, F_{n,n}(x), \dots$

$$\xrightarrow{\text{对一切有理数 } r} \quad \lim_{n\to\infty} F_{n,n}(r) = G(r)$$

是定义在有理数上的一个函数

该函数有界、非降

$$\forall x \in \mathbb{R}^1, \quad F(x) \triangleq \sup_{r_k \le x} G(r_k) \quad \xrightarrow{\text{当 } x \text{ 是有理数时}} \quad F(x) = G(x)$$

有界、非降

引理：设 $\{F_n(x)\}$ 是实变量 $x$ 的非降函数序列，$D$ 是 $\mathbb{R}^1$ 上的稠密集，若对于 $D$ 中的所有点，序列 $\{F_n(x)\}$ 收敛于 $F(x)$，则对 $F(x)$ 的一切连续点 $x$ 都有

$$\lim_{n\to\infty} F_n(x) = F(x)$$

$$\Longrightarrow \quad \lim_{n\to\infty} F_{n,n}(x) = F(x) \quad \text{对 } F(x) \text{ 的一切连续点成立}$$

## 一、分布函数弱收敛

**定理：(Helly 第二定理)**

设 $f(x)$ 是 $[a, b]$ 上的连续函数，又 $\{F_n(x)\}$ 是在 $[a, b]$ 上弱收敛于函数 $F(x)$ 的一致有界非降函数序列，且 $a$ 和 $b$ 是 $F(x)$ 的连续点，则

$$\lim_{n\to\infty}\int_a^b f(x)\,\mathrm{d}F_n(x)=\int_a^b f(x)\,\mathrm{d}F(x)$$

**定理：(推广的 Helly 第二定理)**

设 $f(x)$ 在 $(-\infty, +\infty)$ 上的有界连续，又 $\{F_n(x)\}$ 是 $(-\infty, +\infty)$ 上弱收敛于函数 $F(x)$ 的一致有界非降函数序列，且

$$\lim_{n\to\infty}F_n(-\infty)=F(-\infty)\ ,\quad \lim_{n\to\infty}F_n(+\infty)=F(+\infty)$$

则

$$\lim_{n\to\infty}\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}F_n(x)=\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}F(x)$$

证明过程略！

## 二、连续性定理 (Levy - Cramer)

**定理：(正极限定理)**

设分布函数列 $\{F_n(x)\}$ 弱收敛于某一分布函数 $F(x)$，则相应的特征函数列 $\{f_n(t)\}$ 收敛于特征函数 $f(t)$，且在 $t$ 的任一有限区间内收敛是一致的.

$\mathrm{e}^{itx}$ 在 $-\infty < x < +\infty$ 上有界连续

$$f_n(t) = \int_{-\infty}^{+\infty} \mathrm{e}^{itx} \mathrm{d}F_n(x)$$

$$f(t) = \int_{-\infty}^{+\infty} \mathrm{e}^{itx} \mathrm{d}F(x)$$

**定理：(推广的 Helly 第二定理)**

设 $f(x)$ 在 $(-\infty, +\infty)$ 上的有界连续，又 $\{F_n(x)\}$ 是 $(-\infty, +\infty)$ 上弱收敛于函数 $F(x)$ 的一致有界非降函数序列，且

$$\lim_{n\to\infty} F_n(-\infty) = F(-\infty)\ , \quad \lim_{n\to\infty} F_n(+\infty) = F(+\infty)$$

则

$$\lim_{n\to\infty} \int_{-\infty}^{+\infty} f(x)\,\mathrm{d}F_n(x) = \int_{-\infty}^{+\infty} f(x)\,\mathrm{d}F(x)$$

$$\Longrightarrow \quad \lim_{n\to\infty} f_n(t) = \lim_{n\to\infty} \int_{-\infty}^{+\infty} \mathrm{e}^{itx} \mathrm{d}F_n(x) = \int_{-\infty}^{+\infty} \mathrm{e}^{itx} \mathrm{d}F(x) = f(t)$$

## 二、连续性定理 (Levy - Cramer)

**定理：(逆极限定理)**

设特征函数列 $\{f_n(t)\}$ 收敛于某一函数 $f(t)$，且 $f(t)$ 在 $t=0$ 连续，则相应的分布函数列 $\{F_n(x)\}$ 弱收敛于某一分布函数 $F(x)$，而 $f(t)$ 是 $F(x)$ 的特征函数.

- 通常把正、逆极限定理合称为连续性定理.

**定理：(Levy - Cramer 连续性定理)**

分布函数列 $\{F_n(x)\}$ 弱收敛于某一分布函数 $F(x)$，当且仅当 $\{F_n(x)\}$ 对应的特征函数列 $\{f_n(t)\}$ 在任意有限区间内一致收敛到某个函数 $f(t)$.

# 三、随机变量的收敛性

或 $\xi_n(\omega) \xrightarrow{d} \xi(\omega)$

**定义：(依概率收敛)** $\Longleftrightarrow \lim\limits_{n\to\infty} P\left\{\left|\xi_n(\omega)-\xi(\omega)\right| < \varepsilon\right\} = 1$

如果 $\forall\, \varepsilon > 0$，都有 $\lim\limits_{n\to\infty} P\left\{\left|\xi_n(\omega)-\xi(\omega)\right| \geq \varepsilon\right\} = 0$ 成立，则称 $\{\xi_n(\omega)\}$ 依概率收敛

(convergence in probability) 于 $\xi(\omega)$，记为 $\xi_n(\omega) \xrightarrow{P} \xi(\omega)$.

**定理：(Bernoulli 大数定律)**

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

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

$\Longrightarrow$

设 $\mu_n$ 是 $n$ 重 Bernoulli 试验中事件 $A$ 发生的次数，$p = P(A)$，则 $\dfrac{\mu_n}{n} \xrightarrow{P} p$

## 三、随机变量的收敛性

- 使用最多的是 $r=2$ 的情形，称为**均方收敛**.

## 三、随机变量的收敛性

## 三、随机变量的收敛性

**定理：** $\xi_n \xrightarrow{P} \xi \quad \Longrightarrow \quad \xi_n \xrightarrow{L} \xi\,.$

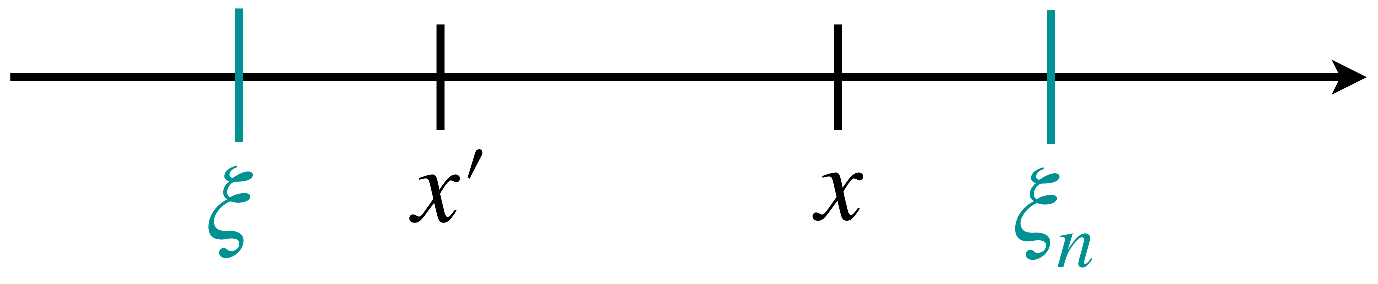
$\forall\, x' < x:$

$$\{\xi < x'\} = \{\xi_n < x,\ \xi < x'\} + \{\xi_n \geq x,\ \xi < x'\}$$

$$\subset \{\xi_n < x\} + \{\xi_n \geq x,\ \xi < x'\}$$

$$\Longrightarrow \quad P\{\xi < x'\} \leq P\{\xi_n < x\} + P\{\xi_n \geq x,\ \xi < x'\}$$

$$\Longrightarrow \quad F(x') \leq F_n(x) + P\{\xi_n \geq x,\ \xi < x'\}$$




$$\xi_n \xrightarrow{P} \xi \quad \Longrightarrow \quad \lim_{n\to\infty} P\left\{\left|\xi_n - \xi\right| \geq \varepsilon\right\} = 0$$

$$\Longrightarrow \quad P\{\xi_n \geq x,\ \xi < x'\} \leq P\left\{\left|\xi_n - \xi\right| \geq x - x'\right\} \longrightarrow 0\,, \quad n \to \infty$$

$$\Longrightarrow \quad F(x') \leq \varliminf_{n\to\infty} F_n(x)$$

## 三、随机变量的收敛性

**定理：** $\xi_n \xrightarrow{P} \xi \quad \Longrightarrow \quad \xi_n \xrightarrow{L} \xi$ .

$\forall\, x < x'':$

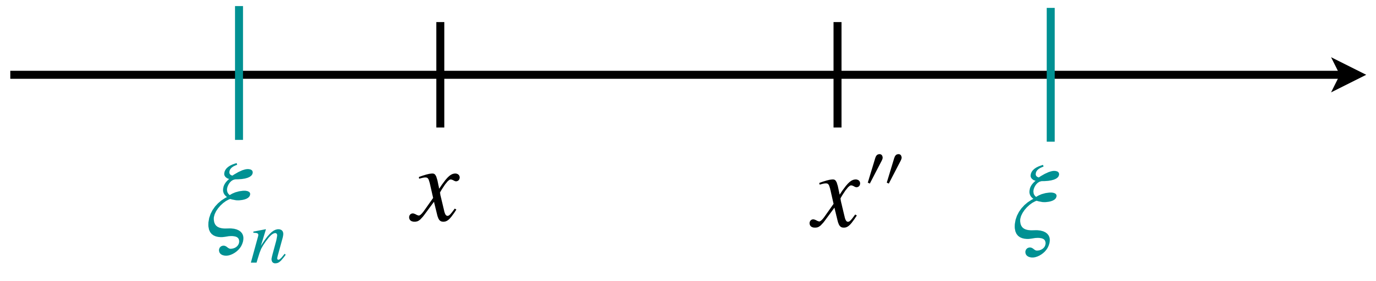
$$\{\xi \ge x''\} = \{\xi_n < x,\ \xi \ge x''\} + \{\xi_n \ge x,\ \xi \ge x''\}$$

$$\subset \{\xi_n < x,\ \xi \ge x''\} + \{\xi_n \ge x\}$$

$$\Longrightarrow \quad P\{\xi \ge x''\} \le P\{\xi_n \ge x\} + P\{\xi_n < x,\ \xi \ge x''\}$$

$$\Longrightarrow \quad 1 - F(x'') \le 1 - F_n(x) + P\{\xi_n < x,\ \xi \ge x''\}$$

$$\Longrightarrow \quad F_n(x) \le F(x'') + P\{\xi_n < x,\ \xi \ge x''\}$$




$$\xi_n \xrightarrow{P} \xi \quad \Longrightarrow \quad \lim_{n\to\infty} P\left\{\left|\xi_n - \xi\right| \ge \varepsilon\right\} = 0$$

$$\Longrightarrow \quad P\{\xi_n < x,\ \xi \ge x''\} \le P\left\{\left|\xi_n - \xi\right| \ge x'' - x\right\} \longrightarrow 0\,, \quad n \to \infty$$

$$\Longrightarrow \quad \overline{\lim_{n\to\infty}} F_n(x) \le F(x'')$$

## 三、随机变量的收敛性

**定理：** $\xi_n \xrightarrow{P} \xi \quad \Longrightarrow \quad \xi_n \xrightarrow{L} \xi$ .

$$\forall\ x' < x < x'' \quad \Longrightarrow \quad F(x') \le \varliminf_{n\to\infty} F_n(x) \le \varlimsup_{n\to\infty} F_n(x) \le F(x'')$$

如果 $x$ 是 $F(x)$ 的连续点，令 $x' \to x$ ， $x'' \to x$ $\quad \Longrightarrow \quad F(x) = \lim\limits_{n\to\infty} F_n(x)$

## 三、随机变量的收敛性

**定理**：$\xi_n \xrightarrow{P} \xi \quad \Longrightarrow \quad \xi_n \xrightarrow{L} \xi$ .

- 该定理的逆命题不成立.

样本空间 $\Omega = \{\omega_1 , \omega_2\}$ $\quad P(\omega_1) = P(\omega_2) = \dfrac{1}{2}$

定义随机变量 $\xi$ ：$\xi(\omega_1) = -1 , \xi(\omega_2) = 1$ $\Longrightarrow$ $\xi$ 的分布列

| $\xi$ | $-1$ | $1$ |
|---|---|---|
| $p_k$ | $\frac{1}{2}$ | $\frac{1}{2}$ |

定义随机变量 $\xi_n = -\xi$ ：$\xi_n(\omega_1) = 1 , \xi_n(\omega_2) = -1$

$\Longrightarrow$ $\xi_n$ 的分布列

| $\xi_n$ | $-1$ | $1$ |
|---|---|---|
| $p_k$ | $\frac{1}{2}$ | $\frac{1}{2}$ |

相同

$\Longrightarrow$ $\xi_n \xrightarrow{L} \xi$ 依分布收敛

$\forall\ 0 < \varepsilon < 2$ ：

$$P\left\{ \left|\xi_n(\omega) - \xi(\omega)\right| > \varepsilon \right\} = P\{\Omega\} = 1 \neq 0 \quad \Longrightarrow \quad \{\xi_n\}\ \text{不依概率收敛于}\ \xi$$

## 三、随机变量的收敛性

- 在特殊情形时有以下结果.

**定理：** 设 $C$ 为常数，则 $\xi_n \xrightarrow{P} C \iff \xi_n \xrightarrow{L} C$.

**定理：** $\xi_n \xrightarrow{P} \xi \implies \xi_n \xrightarrow{L} \xi$. $\Longrightarrow$ $\xi_n \xrightarrow{P} C \implies \xi_n \xrightarrow{L} C$

$$\forall\, \varepsilon > 0: \quad P\left\{\left|\xi_n - C\right| \ge \varepsilon\right\} = P\{\xi_n \ge C+\varepsilon\} + P\{\xi_n \le C-\varepsilon\}$$

$$= 1 - F_n(C+\varepsilon) + F_n(C-\varepsilon+0)$$

$$\xrightarrow{n\to\infty} \quad 1 - F(C+\varepsilon) + F(C-\varepsilon+0) = 1 - 1 + 0 = 0$$

$$F(x) = \begin{cases} 0, & x \le C \\ 1, & x > C \end{cases}$$

## 三、随机变量的收敛性

- $r$ 阶收敛与依概率收敛的关系.

**定理**：$\xi_n \xrightarrow{r} \xi \quad \Longrightarrow \quad \xi_n \xrightarrow{P} \xi.$

$r=2$ 即为 Chebyshev 不等式

先来证明 Markov 不等式： $\forall\, \varepsilon > 0$， $P\left\{\left|\xi_n - \xi\right| \geq \varepsilon\right\} \leq \dfrac{E\left(\left|\xi_n - \xi\right|^r\right)}{\varepsilon^r}$

如果 $F(x)$ 表示随机变量 $\eta = \xi_n - \xi$ 的分布函数

$$\Longrightarrow \quad P\left\{\left|\xi_n - \xi\right| \geq \varepsilon\right\} = \int_{|t|\geq\varepsilon} \mathrm{d}F(t) \leq \int_{|t|\geq\varepsilon} \left(\frac{|t|}{\varepsilon}\right)^r \mathrm{d}F(t) = \frac{1}{\varepsilon^r}\int_{|t|\geq\varepsilon} |t|^r \mathrm{d}F(t)$$

$$\leq \frac{1}{\varepsilon^r}\int_{-\infty}^{\infty} |t|^r \mathrm{d}F(t) = \frac{E\left(\left|\xi_n - \xi\right|^r\right)}{\varepsilon^r}$$

$$\xi_n \xrightarrow{r} \xi \quad \Longrightarrow \quad \lim_{n\to\infty} E\left(\left|\xi_n - \xi\right|^r\right) = 0 \qquad \Longrightarrow \quad \xrightarrow{n\to\infty} 0$$

## 三、随机变量的收敛性

- 上述定理的逆命题不成立.

定理：$\xi_n \xrightarrow{r} \xi \implies \xi_n \xrightarrow{P} \xi.$

样本空间：$\Omega = (0\,,\,1]$ ，定义：

$$\begin{cases} \xi_n(\omega) = \begin{cases} n^{\frac{1}{r}}, 0 < \omega \le \dfrac{1}{n} \\ 0,\ \dfrac{1}{n} < \omega \le 1 \end{cases}, \quad n = 1, 2, \ldots \\ \xi(\omega) \equiv 0 \end{cases}$$

$\Longrightarrow$ 对每一个 $\omega \in \Omega$， $\xi_n(\omega) \longrightarrow \xi(\omega) \equiv 0$

$$\forall\, \varepsilon > 0: \quad P\left\{ \left| \xi_n(\omega) - \xi(\omega) \right| \ge \varepsilon \right\} = P\left\{ \xi_n(\omega) \ge \varepsilon \right\} \longrightarrow 0\,, \quad n \to \infty \implies \xi_n(\omega) \xrightarrow{P} \xi$$

$$E\left( \left| \xi_n - \xi \right|^r \right) = E\left[ \left( n^{\frac{1}{r}} \right)^r \right] = \left( n^{\frac{1}{r}} \right)^r \cdot \frac{1}{n} = 1 \ne 0 \implies$$

$\{\xi_n\}$ 不 $r$ 阶 (矩) 收敛于 $\xi$

## 三、随机变量的收敛性

**定理：** $\xi_n(\omega) \xrightarrow{\text{a.s.}} \xi(\omega) \quad \Longrightarrow \quad \xi_n(\omega) \xrightarrow{P} \xi(\omega).$

证明略 (教材第 329 ~ 334 页).

- 上述定理的逆命题不成立，即：由依概率收敛或矩收敛不能推出几乎处处收敛.

反例见教材第 334 页，此处略.

## 三、随机变量的收敛性

四种收敛性的定义：

① 依分布收敛 $\xi_n \xrightarrow{L} \xi$ : $F_{\xi_n}(x) \xrightarrow{W} F_{\xi}(x)$

② 依概率收敛 $\xi_n \xrightarrow{p} \xi$ : $\lim\limits_{n\to\infty} P\left\{\left|\xi_n - \xi\right| \geq \varepsilon\right\} = 0$

③ $r$ 阶矩收敛 $\xi_n \xrightarrow{r} \xi$ : $\lim\limits_{n\to\infty} E\left(\left|\xi_n - \xi\right|^r\right) = 0$， $r > 0$

④ 几乎处处收敛 $\xi_n \xrightarrow{a.s.} \xi$ : $P\left\{\lim\limits_{n\to\infty} \xi_n(\omega) = \xi(\omega)\right\} = 1$