概率基本概念的比较与分析

薄立军

数学系概率与统计教研室

2012 年 9 月 12 日

内容

古典概型

(a). $\Omega = \{\omega_1, \ldots, \omega_n\}$

(b). 事件域 $\mathcal{F}$:

(b1). 最大的 $\mathcal{F} = \{\Omega$ 的所有子集$\}$ (包含信息最多)
e.g. 当 $n = 2$, $\mathcal{F} = \{\Omega, \emptyset, \{\omega_1\}, \{\omega_2\}\}$

(b2). 最小的 $\mathcal{F} = \{\Omega, \emptyset\}$, 平凡事件域 (包含信息最少)

(c). $\mathbb{P}(A) = \frac{|A|}{n}$, $\forall\ A \in \mathcal{F}$ (等可能性)

Bernoulli 概型

- n 重 Bernoulli 概型：一个随机试验只有两种可能发生的结果 S (发生概率为 p) 和 F. 将此试验独立重复做 n 次.

(a). $\Omega = \{\omega = (\omega_1, \ldots, \omega_n);\ \omega_i = S \text{ or } F,\ i = 1, \ldots, n\}$

(b). $\mathcal{F} = \{\Omega$ 的所有子集$\}$

(c). 若 $\omega \in \Omega$ 中恰有 $m \le n$ 个 S, 则 $\mathbb{P}(\{\omega\}) = p^m(1-p)^{n-m}$

(d). 若 $A = \{\omega_i;\ i \le k\} \in \mathcal{F}$ 有 $k \le 2^n$ 个基本事件, 则

$$\mathbb{P}(A) = \sum_{i=1}^{k} \mathbb{P}(\{\omega_i\})$$

Bernoulli 概型

- 例: 1-重 Bernoulli 概型

(a). $\Omega = \{S, F\}$

(b). $\mathcal{F} = \{\Omega, \emptyset, \{S\}, \{F\}\}$

(c). $\mathbb{P}(\{S\}) = p = 1 - \mathbb{P}(\{F\})$

Bernoulli 概型

- 例: 3-重 Bernoulli 概型

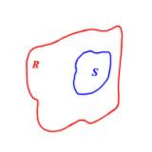
  (a). $\Omega = \{SSS, SSF, SFF, FFF, SFS, FSS, FFS, FSF\}$

  (b). $\mathcal{F} = \{\Omega$ 的子集$\}$

  (c). 例如事件 $A = \{SSF, SFS, FSS\} = \{$恰好成功 2 次$\}$, 则

  $$\mathbb{P}(A) = 3 \cdot \mathbb{P}(\{SSF\}) = C_3^2 p^2 (1-p)$$

几何概型



(a). $\Omega = R$ (区域 R 具有面积)

(b). $\mathcal{F} = \{$目标落入 R 中具有面积的子集 $S\}$

(c). $S \in \mathcal{F}$, 则 $\mathbb{P}(S) = \frac{|S|}{|R|}$ (等可能性)

r.v. 定义

(a). 两个空间 $\Omega \to \mathbb{R}^d$ (我们所熟悉的)

(b). 测度论角度：两个可测空间 $(\Omega, \mathcal{F}) \to (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$

(c). r.v.： $X(\omega) : \Omega \to \mathbb{R}^d$ 的 可测 函数

$$\forall\ A \in \mathcal{B}(\mathbb{R}^d), \quad X^{-1}(A) := \{\omega \in \Omega;\ X(\omega) \in A\} \in \mathcal{F}$$

随机变量与分布

- 举例

(1). 在 n 重 Bernoulli 试验中, $\forall\ \boldsymbol{\omega} = (\omega_1, \ldots, \omega_n) \in \Omega$, 定义

$$X_i(\boldsymbol{\omega}) = \mathbf{1}_A(\omega_i) = \begin{cases} 1, & if\ \omega_i = S \\ 0, & if\ \omega_i = F \end{cases}$$

(2). 对于 $i = 1, 2, \ldots, n$, 定义

$$X(\boldsymbol{\omega}) = \sum_{i=1}^{n} X_i(\boldsymbol{\omega})$$

随机变量与分布

- 累积分布函数 (cdf)

$$F(x)=\mathbb{P}(X\le x):=\mathbb{P}\left(\{\omega\in\Omega;\ X(\omega)\le x\}\right),\quad x\in\mathbb{R}$$

- 离散型 r.v. 的 cdf

$$F(x)=\sum_{\{i;\ a_i\le x\}} p_i$$

其中 $p_i=\mathbb{P}(X=a_i),\ i=1,2,\ldots$ 为分布律 (pmf)

随机变量与分布

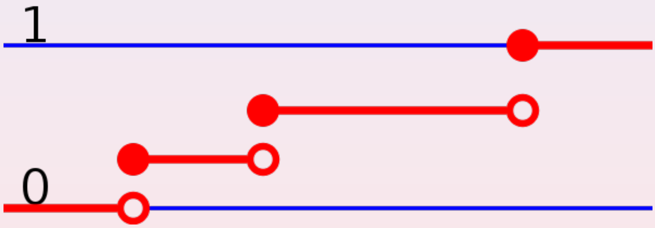


Figure: 离散型 r.v. 的 cdf

随机变量与分布

- 离散型 r.v. cdf:

  (a). 阶梯型函数 (简单函数)

  (b). r.v. 的取值 a_i 是 cdf 的跳点

  (c). cdf 在跳点跳的高度为 p_i

r.v. 分布

- 绝对连续型 r.v. cdf:

$$F(x) = \mathbb{P}(X \le x) = \int_{-\infty}^{x} p(y)\mathrm{d}y, \quad x \in \mathbb{R},$$

其中 $p(\cdot)$ 是概率密度函数 (pdf)

r.v. 分布

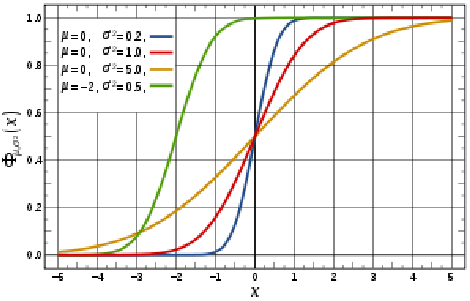


Figure: 绝对连续型 r.v. cdf

随机变量与分布

- 绝对连续型 r.v. cdf:

  (a). 绝对连续的 (因此是连续的)

  (b). 有限变差的 (两个单增函数之差)

  (c). 几乎处处可导的 (任意单调函数都是几乎处处可导的)

随机变量与分布

- 分布函数的 Lebesgue 分解:

Theorem

任意分布函数 $F(x)$ 是均可分解成一个离散型分布函数 $F_d(x)$、一个绝对连续型分布函数 $F_{ac}(x)$ 和一个连续奇异型分布函数 F_{cs} 的凸组合, 即

$$F(x) = a_1 \cdot F_d(x) + a_2 \cdot F_{ac}(x) + a_3 \cdot F_{cs}(x), \qquad \forall\ x \in \mathbb{R},$$

其中 $a_i \geq 0$ 且 $\sum_{i=1}^{3} a_i = 1$.

随机变量与分布

- 连续奇异型分布函数:

  (a). 定义: 一个分布函数 F_{cs} 被称为连续奇异型分布函数, 如果它是连续的且它的导数 $F'_{cs}=0$, a.e..

  (b). 例子: Cantor 分布

随机变量与分布

- Cantor 分布的构成:

(a). Cantor 集 $C = \cap_{n=0}^{\infty} C_n$ (G. Cantor in 1883):

$$
\begin{aligned}
C_0 &= [0,1],\\
C_1 &= [0,1/3] \cup [2/3,1],\\
C_2 &= [0,1/9] \cup [2/9,1/3] \cup [2/3,7/9] \cup [8/9,1]\\
C_3 &= [0,1/27] \cup [2/27,1/9] \cup [2/9,7/27] \cup [8/27,1/3]\\
&\quad \cup [2/3,19/27] \cup [20/27,7/9] \cup [8/9,25/27] \cup [26/27,1],\\
C_4 &= \cdots
\end{aligned}
$$

随机变量与分布

- Cantor 集 $C = \cap_{n=0}^{\infty} C_n$ 的性质:

(a1). C 是闭的 (因此是紧的)

(a2). C 是不可数的但其 Lebesgue 测度 $m(C) = 0$

(a3). $\forall\ x \in [0,1]$, $x = 0.a_1 a_2 a_3 \cdots$ 为 3 基底展开,
则 $x \in C \Longleftrightarrow a_n \in \{0,2\}$, $\forall\ n \in \mathbb{N}$

(a4). C 不包含任何开区间, 即 $(a,b) \not\subset C$, $\forall\ 0 < a < b < 1$

随机变量与分布

(b). Cantor r.v. X:

$$\mathbb{P}(X \in A) = 2^{-n}, \quad n = 0, 1, 2, \ldots$$

其中 A 是 C_n 中并的 2^n 个闭区间的任意一个闭区间

(c). Cantor 分布函数 $F(x) = \mathbb{P}(X \leq x)$ 称为 Cantor 函数

(d). Cantor r.v. 数字特征:

$$\mathbb{E}[X] = 1/2, \ D(X) = 1/8$$

r.v. 的数字特征

- 离散型 r.v. 数学期望 (均值) 的定义:

  X 取值 $\{a_i;\ i=1,2,\dots\}$, pmf 为 $\{p_i;\ i=1,2,\dots\}$. 若成立

$$\sum_{i=1}^{\infty}|a_i|p_i<+\infty$$

  则定义 $\mathbb{E}[X]=\sum_{i=1}^{\infty}a_ip_i$ 为 X 的数学期望.

r.v. 的数字特征

Theorem

设 σ 为 $\mathbb{N}$ 的任意排列. 如果无穷级数 $\{a_i p_i;\ i = 1, 2, \ldots\}$ 绝对收敛, 则

$$\sum_{i=1}^{\infty} a_{\sigma(i)} p_{\sigma(i)} = \sum_{i=1}^{\infty} a_i p_i$$

r.v. 的数字特征

- 绝对连续型 r.v. 数学期望:

如果

$$\int_{-\infty}^{\infty} |x| p(x) \mathrm{d}x < +\infty$$

则定义 $\mathbb{E}[X] = \int_{-\infty}^{\infty} x p(x) \mathrm{d}x$ 为 X 的数学期望.

r.v. 的数字特征

- r.v. 函数的数学期望:

  如果

$$\sum_{i=1}^{\infty}|g(a_i)|p_i<+\infty \quad \left(resp.\ \int_{-\infty}^{\infty}|g(x)|p(x)\mathrm{d}x<+\infty\right)$$

  则

$$\mathbb{E}[g(X)]=\sum_{i=1}^{\infty}g(a_i)p_i \quad \left(resp.\ \int_{-\infty}^{\infty}g(x)p(x)\mathrm{d}x\right)$$

r.v. 的数字特征

- r.v. 数字特征不一定存在

(a). 离散 r.v. X 的 pmf:

$$p_i = \begin{cases} 1/i, & i \in R = \{2^n;\ n = 1, 2, \ldots\}, \\ 0, & otherwise \end{cases}$$

考察

$$\sum_{i \in R} |i|\, p_i = \sum_{i \in R} i \cdot \frac{1}{i} = +\infty$$

r.v. 的数字特征

(b). 服从标准 *Cauchy* 分布 r.v. X 的 pdf:

$$p(x) = \frac{1}{\pi(1+x^2)}, \qquad x \in \mathbb{R}$$

考察: $\forall\ n > 0$,

$$\begin{aligned} \int_{-\infty}^{\infty} |x|^n p(x)\mathrm{d}x &= \frac{1}{\pi}\int_0^1 x^{\frac{1-n}{2}-1}(1-x)^{\frac{1+n}{2}-1}\mathrm{d}x \\ &< +\infty, \quad \textit{if}\ \frac{1\pm n}{2} > 0 \Longleftrightarrow n < 1 \end{aligned}$$

r.v. 的数字特征

- 如果 $\mathbb{E}[|X|^2]$ 存在, 则 $\mathbb{E}[X]$ 存在:

$$(\mathbb{E}[X])^2 \le \mathbb{E}[|X|^2].$$

- 对 2d r.v. (X, Y), 若 $D(X), D(Y)$ 存在, 则 $Cov(X, Y)$ 存在:

$$|Cov(X, Y)|^2 \le D(X) \cdot D(Y).$$

r.v. 的数字特征

- 随机向量的数字特征: 设 $\mathbf{X} = (X_1, X_2, \cdots, X_n)^*$ 是一 n 维 r.v.

(a). $\mathbf{X}$ 的数学期望: (if exists)

$$\mathbb{E}[\mathbf{X}] = (\mathbb{E}[X_1], \mathbb{E}[X_2], \cdots, \mathbb{E}[X_n])^* \in \mathbb{R}^n$$

(b). $\mathbf{X}$ 的方差: (if exists)

$$D(\mathbf{X}) = \mathbb{E}\left[(\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})^*\right],$$

其中 $\boldsymbol{\mu} = \mathbb{E}[\mathbf{X}]$.

r.v. 的数字特征

- $\mathbf{X}$ 数学期望的性质:

(a). $\mathbb{E}[\mathbf{X}^*] = \mathbb{E}[\mathbf{X}]^*$

(b). 设 $a \in \mathbb{R}$ 是一实数, 则 $\mathbb{E}[a \cdot \mathbf{X}] = a \cdot \mathbb{E}[\mathbf{X}]$

(c). $\mathbb{E}[\mathbf{X}+\mathbf{Y}] = \mathbb{E}[\mathbf{X}] + \mathbb{E}[\mathbf{Y}]$

(d). 若 $\mathbf{X}$ 与 $\mathbf{Y}$ 独立, 则 $\mathbb{E}[\mathbf{X}\mathbf{Y}^*] = \mathbb{E}[\mathbf{X}]\mathbb{E}[\mathbf{Y}]^*$

r.v. 的数字特征

- **X** 方差的性质:

  (a). $D(\mathbf{X})$ 是 $n \times n$ 对称方阵:

$$D(\mathbf{X}) = \begin{pmatrix} D(X_1) & Cov(X_1, X_2) & \cdots & Cov(X_1, X_n) \\ Cov(X_2, X_1) & D(X_2) & \cdots & Cov(X_2, X_n) \\ \cdots & \cdots & \cdots & \cdots \\ Cov(X_n, X_1) & Cov(X_n, X_2) & \cdots & D(X_n) \end{pmatrix}$$

r.v. 的数字特征

(b). 若 $X_1, \cdots, X_n$ 两两独立, 则 $D(\mathbf{X})$ 是对角阵:

$$D(\mathbf{X}) = \begin{pmatrix} D(X_1) & 0 & \cdots & 0 \\ 0 & D(X_2) & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & D(X_n) \end{pmatrix}$$

r.v. 的数字特征

(c). 若 $X_1, \cdots, X_n$ 两两独立且同分布, 则 $D(\mathbf{X}) = \sigma^2 \mathbf{I}_{n\times n}$

(d). $\mathbb{E}[\mathbf{X}\mathbf{X}^*] = D(\mathbf{X}) + \mathbb{E}[\mathbf{X}]\mathbb{E}[\mathbf{X}]^*$

(e). 设 $a \in \mathbb{R}$ 是一实数, 则 $D(a \cdot \mathbf{X}) = a^2 \cdot D(\mathbf{X})$

(f). 设 $\mathbf{A}$ 是 $m \times n$-常数矩阵, 则 $D(\mathbf{A}\mathbf{X} + a) = \mathbf{A}D(\mathbf{X})\mathbf{A}^*$

(g). $D(\mathbf{X})$ 总是非负定的, 即对任意常向量 $\mathbf{b}$, 由 (f) 得:

$$\mathbf{b}^* D(\mathbf{X})\mathbf{b} = D(\mathbf{b}^*\mathbf{X}) \geq 0$$

r.v. 的数字特征

- 多维正态分布

(a). 随机向量 $\mathbf{X} = (X_1, \cdots, X_n)^* \sim N(\boldsymbol{\mu}, \mathbf{C})$:

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{n}{2}} |\mathbf{C}|^{\frac{1}{2}}} \exp\left[-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^* \mathbf{C}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right],$$

其中 $\mathbf{x} = (x_1, \cdots, x_n)^* \in \mathbb{R}^n$, $\boldsymbol{\mu} = \mathbb{E}[\mathbf{X}]$ 和 $\mathbf{C} = D(\mathbf{X})$.

r.v. 的数字特征

(b). 多维正态的仿射变换:

设 $\mathbf{X} \sim N(\boldsymbol{\mu}, \mathbf{C})$ 且 $\mathbf{b}$ 为 $m \times 1$ 常向量和 $\mathbf{M}$ 为 $m \times n$ 常矩阵, 则

$$\mathbf{b} + \mathbf{M}\mathbf{X} \sim N(\mathbf{b} + \mathbf{M}\boldsymbol{\mu}, \mathbf{M}\mathbf{C}\mathbf{M}^*)$$

r.v. 的数字特征

- 2-维正态分布

(a). 随机向量 $\mathbf{X} = (X_1, X_2)^* \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2; \rho)$:

$$p(\mathbf{x}) = \frac{1}{2\pi|\mathbf{C}|^{\frac{1}{2}}} \exp\left[-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^* \mathbf{C}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right],$$

其中 $\boldsymbol{\mu} = (\mu_1, \mu_2)$ 以及

$$\mathbf{C} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$$

r.v. 的数字特征

(b). 2-维正态 r.v. 的条件数学期望函数:

$$C_{ef}(x) := \mathbb{E}[X_2|X_1 = x] = \mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x - \mu_1), \quad x \in \mathbb{R}$$

(c). 2-维正态 r.v. 的条件数学期望:

$$\mathbb{E}[X_2|X_1] := C_{ef}(X_1) = \mu_2 + \rho\frac{\sigma_2}{\sigma_1}(X_1 - \mu_1)$$

(d). 条件数学期望性质:

$$\mathbb{E}\left\{\mathbb{E}[X_2|X_1]\right\} = \mathbb{E}\left[\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(X_1 - \mu_1)\right] = \mu_2 = \mathbb{E}[X_2]$$

r.v. 的数字特征

- 1-维正态 r.v. 的矩:

n	$\mathbb{E}[X^n]$	$\mathbb{E}[(X-\mu)^n]$
1	μ	0
2	$\mu^2+\sigma^2$	σ^2
3	$\mu^3+3\mu\sigma^2$	0
4	$\mu^4+6\mu^2\sigma^2+3\sigma^4$	$3\sigma^4$
5	$\mu^5+10\mu^3\sigma^2+15\mu\sigma^4$	0
6	$\mu^6+15\mu^4\sigma^2+45\mu^2\sigma^4+15\sigma^6$	$15\sigma^6$
7	$\mu^7+21\mu^5\sigma^2+105\mu^3\sigma^4+105\mu\sigma^6$	0
8	$\mu^8+28\mu^6\sigma^2+210\mu^4\sigma^4+420\mu^2\sigma^6+105\sigma^8$	$105\sigma^8$

r.v. 的数字特征

- 1-维正态分布的偏态 (Skewness)

$$Skew[X] := \mathbb{E}\left[\left(\frac{X-\mu}{\sigma}\right)^3\right] = 0$$

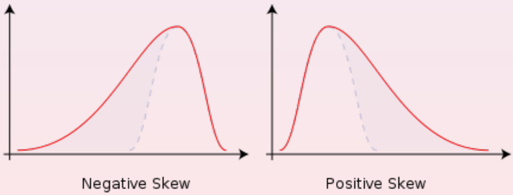


Figure: 偏态反应分布的非对称性

r.v. 的数字特征

- 1-维正态分布的峰度 (Kurtosis)

$$Kurt[X] := \mathbb{E}\left[\left(\frac{X-\mu}{\sigma}\right)^4\right] = 3$$

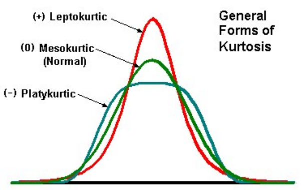


Figure: 峰度反应分布的峰值情况

r.v. 的数字特征

- 正态分布的历史
  1. A. de Moivre (1667-1754, 法国数学家) 在 1783 年发现了正态分布
  2. K. Gauss (1777-1855, 德国数学家) 在 1809 年研究最小二乘法时提出正态分布
  3. P. Laplace (1749-1827, 法国数学家) 在 1810 年证明了 CLT:

$$Z_n := \frac{1}{n}\sum_{i=1}^{n} X_i \text{ 的标准化 r.v. 依分布收敛于 } N(0,1), \quad n \to \infty$$

  4. Gauss 命名正态分布为 "Normal Distribution"

r.v. 的数字特征

- 正态分布的历史
  1. A. de Moivre (1667-1754, 法国数学家) 在 1783 年发现了正态分布
  2. K. Gauss (1777-1855, 德国数学家) 在 1809 年研究最小二乘法时提出正态分布
  3. P. Laplace (1749-1827, 法国数学家) 在 1810 年证明了 CLT:
  $$Z_n := \frac{1}{n}\sum_{i=1}^{n} X_i \text{ 的标准化 r.v. 依分布收敛于 } N(0,1), \quad n \to \infty$$
  4. Gauss 命名正态分布为 "Normal Distribution"

r.v. 的数字特征

- 正态分布的历史
  1. A. de Moivre (1667-1754, 法国数学家) 在 1783 年发现了正态分布
  2. K. Gauss (1777-1855, 德国数学家) 在 1809 年研究最小二乘法时提出正态分布
  3. P. Laplace (1749-1827, 法国数学家) 在 1810 年证明了 CLT：

$$Z_n := \frac{1}{n}\sum_{i=1}^{n} X_i \text{ 的标准化 r.v. 依分布收敛于 } N(0,1), \quad n \to \infty$$

  4. Gauss 命名正态分布为 "Normal Distribution"

r.v. 的数字特征

- 正态分布的历史
  1. A. de Moivre (1667-1754, 法国数学家) 在 1783 年发现了正态分布
  2. K. Gauss (1777-1855, 德国数学家) 在 1809 年研究最小二乘法时提出正态分布
  3. P. Laplace (1749-1827, 法国数学家) 在 1810 年证明了 CLT：

$$Z_n := \frac{1}{n}\sum_{i=1}^{n} X_i \text{ 的标准化 r.v. 依分布收敛于 } N(0,1), \quad n \to \infty$$

  4. Gauss 命名正态分布为 "Normal Distribution"

2-维 r.v. 函数的分布

- 2-维 r.v. (X, Y) 函数 r.v.:

$$Z = \Phi(X, Y),$$

其中函数 $\Phi(x, y)$ 主要形式有:

$$\begin{aligned}\Phi(x,y) &= x \pm y,\ x \cdot y,\ x/y, \\ &= \min\{x,y\},\ \max\{x,y\}, \\ &= \sqrt{x^2+y^2},\ \tan^{-1}(x/y)\end{aligned}$$

2-维 r.v. 函数的分布

- r.v. $Z = \Phi(X, Y)$ 的分布:

$$
\begin{aligned}
F_Z(z) &= \mathbb{P}(\Phi(X, Y) \leq z) \\
&= \iint_{\{(x,y)\in\mathbb{R}^2;\ \Phi(x,y)\leq z\}} p_{X,Y}(u, v)\mathrm{d}u\mathrm{d}v
\end{aligned}
$$

2-维 r.v. 函数的分布

- 函数 $\Phi(x,y)=x+y$:

  (a). 分布函数:

$$F_Z(z)=\mathbb{P}(X+Y\le z)=\int_{v=-\infty}^{\infty}\int_{u=-\infty}^{z-v}p_{X,Y}(u,v)\mathrm{d}u\mathrm{d}v$$

  (b). pdf:

$$p_Z(z)=\int_{-\infty}^{\infty}p_{X,Y}(z-v,v)\mathrm{d}v=\int_{-\infty}^{\infty}p_{X,Y}(u,z-u)\mathrm{d}u$$

2-维 r.v. 函数的分布

(c). 若 X 与 Y 独立, 则

$$p_Z(z) = \int_{-\infty}^{\infty} p_X(z-v)p_Y(v)\mathrm{d}v = \int_{-\infty}^{\infty} p_X(u)p_Y(z-u)\mathrm{d}u$$

2-维 r.v. 函数的分布

- 例: 设 r.v.s X 与 Y 独立同分布于参数为 $\lambda > 0$ 的指数分布, 求 $Z = X + Y$ 的 pdf.

  解:

  (a). 将 X, Y 的 pdf 写成示性函数形式:

$$p_X(x) = \lambda e^{-\lambda x}\mathbf{1}_{\{x>0\}}, \quad p_Y(y) = \lambda e^{-\lambda y}\mathbf{1}_{\{y>0\}}$$

2-维 r.v. 函数的分布

(b). 用卷积公式计算:

$$\begin{aligned}
p_Z(z) &= \int_{-\infty}^{\infty} p_X(z-v)p_Y(v)\mathrm{d}v = \int_0^{\infty} p_X(z-v)\lambda e^{-\lambda v}\mathrm{d}v \\
&= \int_0^{\infty} \lambda e^{-\lambda(z-v)}\mathbf{1}_{\{z-v>0\}}\lambda e^{-\lambda v}\mathrm{d}v \\
&= \lambda^2 e^{-\lambda z}\int_0^{\infty}\mathbf{1}_{\{z>v\}}\mathrm{d}v = \lambda^2 e^{-\lambda z}\int_0^{z}\mathrm{d}v\mathbf{1}_{\{z>0\}} \\
&= \lambda^2 z e^{-\lambda z}\mathbf{1}_{\{z>0\}}.
\end{aligned}$$

2-维 r.v. 函数的分布

- 函数 $\Phi(x,y)=x/y$:

  (a). 分布函数:

$$\begin{aligned}
F_Z(z) &= \mathbb{P}(X/Y\le z, Y>0)+\mathbb{P}(X/Y\le z, Y<0)\\
&= \mathbb{P}(X\le Yz, Y>0)+\mathbb{P}(X\ge Yz, Y<0)\\
&= \int_0^{\infty}\int_{-\infty}^{yz} p_{X,Y}(u,v)\mathrm{d}u\mathrm{d}v+\int_{-\infty}^{0}\int_{yz}^{\infty} p_{X,Y}(u,v)\mathrm{d}u\mathrm{d}v
\end{aligned}$$

2-维 r.v. 函数的分布

(b). pdf:

$$\begin{aligned} p_Z(z) &= \int_0^{\infty} v p_{X,Y}(vz, v)\mathrm{d}v + \int_{-\infty}^{0} (-v) p_{X,Y}(vz, v)\mathrm{d}v \\ &= \int_{-\infty}^{\infty} |v| p_{X,Y}(vz, v)\mathrm{d}v \end{aligned}$$

(c). 若 X 与 Y 独立, 则

$$p_Z(z) = \int_{-\infty}^{\infty} |v| \cdot p_X(vz) \cdot p_Y(v)\mathrm{d}v$$

2-维 r.v. 函数的分布

- 例: 设 $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2;\rho)$, 求 $Z=X/Y$ 的 pdf.

解:

(a). (X,Y) 的联合 pdf 为:

$$\begin{aligned} p_{X,Y}(x,y) &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \\ &\quad \times \exp\left[-\frac{1}{2(1-\rho^2)}\left(\frac{x^2}{\sigma_1^2}-\frac{2\rho xy}{\sigma_1\sigma_2}+\frac{y^2}{\sigma_2^2}\right)\right] \end{aligned}$$

2-维 r.v. 函数的分布

(b). 用公式计算:

$$
\begin{aligned}
p_Z(z) &= \int_{-\infty}^{\infty} |v| p_{X,Y}(vz, v)\mathrm{d}v \\
&= \frac{\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}{\sigma_2^2(z-\rho\sigma_1/\sigma_2)^2+\sigma_1^2(1-\rho^2)}
\end{aligned}
$$

这是 $Cauchy(\rho\sigma_1/\sigma_2)$ 分布的 pdf.

2-维 r.v. 函数的分布

- 函数 $\Phi(x, y) = x^2 + y^2$:

(a). 分布函数: 对于 $z > 0$,

$$F_Z(z) = \mathbb{P}(X^2 + Y^2 \le z) = \int_{-\sqrt{z}}^{\sqrt{z}} \int_{-\sqrt{z-v^2}}^{\sqrt{z-v^2}} p_{X,Y}(u, v)\mathrm{d}u\mathrm{d}v$$

(b). pdf 为: 对于 $z > 0$,

$$\begin{aligned} p_Z(z) \quad = \quad & \int_{-\sqrt{z}}^{\sqrt{z}} \frac{1}{2\sqrt{z - v^2}} \Big[p_{X,Y}(\sqrt{z - v^2}, v) \\ & + p_{X,Y}(-\sqrt{z - v^2}, v) \Big] \mathrm{d}v \end{aligned}$$

2-维 r.v. 函数的分布

- 例: 设 $(X, Y) \sim N(0, 0, 1, 1; 0)$, 求 $Z = X^2 + Y^2$ 的 pdf.

  解:

  (a). 用公式计算:

$$
\begin{aligned}
p_Z(z) &= \int_{-\sqrt{z}}^{\sqrt{z}} \frac{1}{2\sqrt{z - v^2}} \frac{1}{\pi} \exp\left(-\frac{z}{2}\right) \mathrm{d}v \mathbf{1}_{\{z>0\}} \\
&= \frac{1}{\pi} \exp\left(-\frac{z}{2}\right) \int_0^{\pi/2} \frac{\sqrt{z}\sin(\theta)}{\sqrt{z}\cos(\theta)} \mathrm{d}\theta \mathbf{1}_{\{z>0\}} \\
&= \frac{1}{2} \exp\left(-\frac{z}{2}\right) \mathbf{1}_{\{z>0\}}.
\end{aligned}
$$

2-维 r.v. 函数的分布

- 函数 $\Phi(x,y)=\sqrt{x^2+y^2}$:

(a). 分布函数：对于 $z>0$,

$$F_Z(z)=\mathbb{P}(X^2+Y^2\le z^2)=\int_{-z}^{z}\int_{-\sqrt{z^2-v^2}}^{\sqrt{z^2-v^2}}p_{X,Y}(u,v)\mathrm{d}u\mathrm{d}v$$

(b). pdf 为：对于 $z>0$,

$$\begin{aligned}p_Z(z) &= \int_{-z}^{z}\frac{z}{\sqrt{z^2-v^2}}\Big[p_{X,Y}(\sqrt{z^2-v^2},v)\\ &\quad +p_{X,Y}(-\sqrt{z^2-v^2},v)\Big]\mathrm{d}v\end{aligned}$$

2-维 r.v. 函数的分布

- 例: 设 $(X, Y) \sim N(0,0,1,1;0)$, 求 $Z = \sqrt{X^2+Y^2}$ 的 pdf.

  解:

  (a). 用公式计算:

$$p_Z(z) = z\exp\left(-\frac{z^2}{2}\right)\mathbf{1}_{\{z>0\}},$$

  这是 Rayleigh 分布的 pdf.

2-维 r.v. 函数的分布

- 例: 设 $(X, Y) \sim N(0,0,1,1;0)$, 定义复值高斯 r.v.:

$$W = X + i \cdot Y, \qquad i = \sqrt{-1}.$$

则 $|W|$ 服从 Rayleigh 分布. 求 W 的相位:

$$\Theta = \tan^{-1}\left(\frac{X}{Y}\right) \in (-\pi/2, \pi/2)$$

的分布?

2-维 r.v. 函数的分布

- 解:

(a). 定义 r.v. $U = \tan\Theta$, 则 $U = X/Y$

(b). 由前面的例子得: U 服从中心 Cauchy 分布, 其 pdf 为:

$$p_U(u) = \frac{1}{\pi(1+u^2)}, \qquad -\infty < u < +\infty.$$

2-维 r.v. 函数的分布

(c). 于是:

$$\begin{aligned} p_{\Theta}(\theta) &= \frac{1}{|\mathrm{d}\theta/\mathrm{d}u|} p_U(\tan\theta) = \frac{\sec^2\theta}{\pi(\tan^2\theta+1)} \\ &= \begin{cases} \frac{1}{\pi}, & \theta \in (-\frac{\pi}{2}, \frac{\pi}{2}), \\ 0, & otherwise \end{cases} \end{aligned}$$

这是 $(-\frac{\pi}{2}, \frac{\pi}{2})$ 上的均匀分布的 pdf.

(d). 进一步结果: $|W|$ 与 Θ 独立.

2-维 r.v. 函数的分布

- 函数 $\Phi(x, y) = \max\{x, y\}$:

  (a). 分布函数:

  $$F_Z(z) = \mathbb{P}(\max\{X, Y\} \leq z) = F_{X,Y}(z, z)$$

  (b). 若 X, Y 独立, 则 pdf 为:

  $$p_Z(z) = p_X(z)F_Y(z) + F_X(z)p_Y(z)$$

2-维 r.v. 函数的分布

- 函数 $\Phi(x,y)=\min\{x,y\}$:

  (a). 分布函数:

$$\begin{aligned} F_Z(z) &= \mathbb{P}(\min\{X,Y\}\le z) = 1-\mathbb{P}(X>z, Y>z) \\ &= F_X(z)+F_Y(z)-F_{X,Y}(z,z) \end{aligned}$$

  (b). 若 X, Y 独立, 则 pdf 为:

$$p_Z(z)=p_X(z)+p_Y(z)-p_X(z)F_Y(z)-F_X(z)p_Y(z)$$

2-维 r.v. 函数的分布

- 例: 设 $X, Y \sim Exp(\lambda)$ 且独立, 求 $Z = \min\{X, Y\}$ 的 pdf.

  解:

  (a). pdf 为:

$$\begin{aligned} p_Z(z) &= p_X(z) + p_Y(z) - p_X(z)F_Y(z) - F_X(z)p_Y(z) \\ &= 2\lambda e^{-\lambda z}\mathbf{1}_{\{z>0\}} - 2\lambda e^{-\lambda z}\left(1 - e^{-\lambda z}\right)\mathbf{1}_{\{z>0\}} \\ &= 2\lambda e^{-2\lambda z}\mathbf{1}_{\{z>0\}} \end{aligned}$$

  这说明 $Z = \min\{X, Y\} \sim Exp(2\lambda)$.

2-维 r.v. 函数的分布

- 例: 设 $X, Y \sim Exp(\lambda)$ 且独立, 定义 r.v.:

$$Z = \frac{\min\{X, Y\}}{\max\{X, Y\}} \in (0, 1)$$

计算 Z 的 pdf.

2-维 r.v. 函数的分布

- 解:

(a). (X, Y) 联合 pdf:

$$p_{X,Y}(x,y) = \lambda^2 e^{-\lambda(x+y)} \mathbf{1}_{\{x>0,y>0\}}$$

(b). Z 的分布函数: 对于 $0 < z < 1$,

$$\begin{aligned} F_Z(z) &= \mathbb{P}(\min\{X,Y\}/\max\{X,Y\} \le z) \\ &= \mathbb{P}(X \le z\cdot Y, X \le Y) + \mathbb{P}(Y \le z\cdot X, X > Y) \\ &= \int_0^\infty \int_0^{vz} p_{X,Y}(u,v)\mathrm{d}u\mathrm{d}v + \int_0^\infty \int_0^{uz} p_{X,Y}(u,v)\mathrm{d}v\mathrm{d}u \end{aligned}$$

2-维 r.v. 函数的分布

(c). Z 的 pdf 为:

$$
\begin{aligned}
p_Z(z) &= \int_0^{\infty} y \cdot p_{X,Y}(vz, v)\mathrm{d}v + \int_0^{\infty} x \cdot p_{X,Y}(u, uz)\mathrm{d}u \\
&= \frac{2}{(1+z)^2}\mathbf{1}_{\{0<z<1\}}
\end{aligned}
$$

事件的互斥与独立

- 事件 A 与 B 互斥: $A \cap B = \emptyset$

(a). 若 A 与 B 互斥, 则

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$$

(b). 若 $A_1, \cdots, A_n, \cdots$ 两两互斥, 则

$$\mathbb{P}\left(\cup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(A_n)$$

事件的互斥与独立

- 事件 A 与 B 独立:

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$$

(a). 若 A 与 B 独立, 则

$$\mathbb{P}(A|B) = \mathbb{P}(A), \quad \mathbb{P}(B) \neq 0,$$
$$\mathbb{P}(B|A) = \mathbb{P}(B), \quad \mathbb{P}(A) \neq 0$$

事件的互斥与独立

(b). 若 A 与 B 独立, 则

$$\mathbb{E}\left[\mathbf{1}_A \cdot \mathbf{1}_B\right] = \mathbb{E}\left[\mathbf{1}_A\right] \cdot \mathbb{E}\left[\mathbf{1}_B\right]$$

(c). 若 A 与 B 既互斥又独立, 则

$$0 = \mathbb{P}(A) \cdot \mathbb{P}(B) \iff \min\left\{\mathbb{P}(A), \mathbb{P}(B)\right\} = 0$$

事件的互斥与独立

- n 个事件两两独立 $(n \geq 3)$:

  称 n 个事件 $A_1, \ldots, A_n$ 两两独立, 如果

  $$\mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i) \cdot \mathbb{P}(A_j), \quad i \neq j,$$

  其中 $i, j = 1, 2, \ldots, n$.

事件的互斥与独立

- n 个事件相互独立 ($n \geq 3$):

  称 n 个事件 $A_1, \ldots, A_n$ 相互独立, 如果

$$\mathbb{P}\left(\cap_{i \in I_n} A_i\right) = \prod_{i \in I_n} \mathbb{P}(A_i),$$

  对所有 $\{1, 2, \ldots, n\}$ 的子集 I_n.

事件的互斥与独立

- 3 个事件相互独立:

(a). 3 个事件 A, B, C 相互独立 $\Longleftrightarrow$

$$\text{两两独立} \quad \begin{cases} \mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B) \\ \mathbb{P}(A \cap C) = \mathbb{P}(A) \cdot \mathbb{P}(C) \\ \mathbb{P}(B \cap C) = \mathbb{P}(B) \cdot \mathbb{P}(C) \end{cases}$$

$$+$$

$$\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A) \cdot \mathbb{P}(B) \cdot \mathbb{P}(C)$$

事件的互斥与独立

(b). 相互独立 $\Longrightarrow$ 两两独立

(c). 两两独立 $\not\Longrightarrow$ 相互独立

事件的互斥与独立

(d). 反例: 设有四张外型一样的卡片,上分别写有数字 $2, 3, 5, 30$, 今从中任取一张观察其上数字.

$$A = \{\text{取到是 } 2 \text{ 的倍数}\}, \quad B = \{\text{取到是 } 3 \text{ 的倍数}\}$$
$$C = \{\text{取到是 } 5 \text{ 的倍数}\},$$

则 A, B, C 两两独立但不相互独立. 因为

$$A = \{2, 30\},\ B = \{3, 30\},\ C = \{5, 30\},$$
$$AB = AC = BC = ABC = \{30\}$$

故 $\mathbb{P}(A) = \mathbb{P}(B) = \mathbb{P}(C) = 0.5$, $\mathbb{P}(AB) = \mathbb{P}(A)\mathbb{P}(B) = 0.25$,
类似 $\mathbb{P}(BC) = \mathbb{P}(B)\mathbb{P}(C)$, $\mathbb{P}(AC) = \mathbb{P}(A)\mathbb{P}(C)$,
而 $\mathbb{P}(ABC) = 0.25$,
但 $\mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C) = 0.5 \cdot 0.5 \cdot 0.5 = 0.125$ 两者不等

r.v. 的独立性

- 称取实值 r.v.s $X_1, \dots, X_n$ 相互独立, 如果 $\forall\, a_1, \dots, a_n \in \mathbb{R}$, 事件
$$\{X_1 \le a_1\},\ \{X_2 \le a_2\}, \cdots, \{X_n \le a_n\}$$
相互独立.

(a). 实值 r.v.s $X_1, \dots, X_n$ 相互独立 $\Longleftrightarrow \forall\, a_1, \dots, a_n \in \mathbb{R}$,
$$F_{X_1,\dots,X_n}(a_1, \cdots, a_n) = \prod_{i=1}^{n} F_{X_i}(a_i)$$

r.v. 的独立性

- 取 U-值 r.v.s $X_1, \ldots, X_n$ 相互独立, 若 $\forall\ A_1, \ldots, A_n \in \mathcal{B}(U)$,

$$\{X_1 \in A_1\},\ \{X_2 \in A_2\}, \cdots, \{X_n \in A_n\}$$

  相互独立

r.v. 的独立性

- r.v.s 相互独立用数学期望判别:

Theorem

r.v.s $X_1, \ldots, X_n$ 相互独立 $\Longleftrightarrow \forall$ 可测函数 $g_i(x)$ 使

$$\mathbb{E}\left[\prod_{i=1}^{n} g_i(X_i)\right] = \prod_{i=1}^{n} \mathbb{E}\left[g_i(X_i)\right],$$

其中 $\mathbb{E}\left[g_i(X_i)\right]$ 存在.

r.v. 的独立性

- r.v.s 相互独立用特征函数判别:

Theorem

实值 *r.v.s* $X_1, \dots, X_n$ 相互独立 $\iff \forall\ u_1, \dots, u_n \in \mathbb{R}$,

$$\mathbb{E}\left[\exp\left(\mathrm{i}\sum_{j=1}^{n} u_j X_j\right)\right] = \prod_{j=1}^{n} \mathbb{E}\left[\exp\left(\mathrm{i} u_j X_j\right)\right],$$

其中 $\mathrm{i} = \sqrt{-1}$.

r.v. 的独立性

- r.v.s X, Y 不相关：设 $D(X), D(Y)$ 存在

(a). $Cox(X, Y) = 0$

(b). $\rho_{XY} = \frac{Cov(X,Y)}{\sqrt{D(X)D(Y)}} = 0$

(c). $D(X + Y) = D(X) + D(Y)$

(d). $X - \mathbb{E}[X]$ 与 $Y - \mathbb{E}[Y]$ 正交

(e). $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$

r.v. 的独立性

- 不相关 $\not\Longrightarrow$ 独立

  (a). r.v.s (X, Y) 服从二维正态, 则 X, Y 不相关 $\Longleftrightarrow$ 独立

  (b). r.v. $X \sim U(-1, 1)$ 和 r.v. $Y = X^2$, 则 X, Y 不相关但不独立

  (c). r.v. $\mathbb{P}(X = 0) = \mathbb{P}(X = 1) = \frac{1}{2}$ 与 $\mathbb{P}(Y = -1) = \mathbb{P}(Y = 1) = \frac{1}{2}$, 则 $Z = XY$ 与 X 不相关但不独立

随机变量列的收敛

- 依分布收敛 (弱收敛)：$X_n \Longrightarrow X$

Definition

设 $(X_n;\ n = 1, 2, \ldots)$ 为一列 r.v.s 与 r.v. X, 如果对任意 F 的连续点 x 满足：

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

随机变量列的收敛

- 依概率收敛：$X_n \xrightarrow{P} X$

Definition

设 $(X_n;\ n=1,2,\dots)$ 为一列 r.v.s 与 r.v. X, 如果对任意 $\varepsilon>0$ 使

$$\lim_{n\to\infty}\mathbb{P}\left(|X_n-X|>\varepsilon\right)=0$$

随机变量列的收敛

- 几乎处处收敛：$X_n \xrightarrow{a.s.} X$

Definition

设 $(X_n;\ n = 1, 2, \ldots)$ 为一列 r.v.s 与 r.v. X, 如果

$$\mathbb{P}\left(\lim_{n\to\infty} X_n = X\right) = 1$$

随机变量列的收敛

- L^p $(p \geq 1)$ 收敛：$X_n \xrightarrow{L^p} X$

Definition

设 $(X_n;\ n = 1, 2, \ldots)$ 为一列 r.v.s 与 r.v. X, 如果

$$\lim_{n\to\infty} \mathbb{E}\left[|X_n - X|^p\right] = 0$$

当 $p = 2$ 时, L^2 收敛称为均方收敛

随机变量列的收敛

收敛的关系:

$$X_n \xrightarrow{L^p} X \rightsquigarrow X_n \xrightarrow{P} X \rightsquigarrow X_n \Longrightarrow X$$

$$X_n \xrightarrow{a.s.} X \rightsquigarrow X_n \xrightarrow{P} X \rightsquigarrow X_n \Longrightarrow X$$

$$X_n \xrightarrow{P} X \rightsquigarrow X_{k_n} \xrightarrow{a.s.} X$$

$$X_n \xrightarrow{P} C \leftrightsquigarrow X_n \Longrightarrow C$$

随机变量列的收敛

- Borel-Cantelli 引理:

$$\sum_{n=1}^{\infty} \mathbb{P}(|X_n - X| > \varepsilon) < \infty \rightsquigarrow X_n \xrightarrow{a.s.} X$$

- 控制收敛定理 (DCT):

$$\left.\begin{array}{r} X_n \xrightarrow{a.s.} X \\ |X_n| \le Y \\ \mathbb{E}[Y^p] < +\infty \end{array}\right\} \rightsquigarrow X_n \xrightarrow{L^p} X$$

随机变量列的收敛

- 连续映射定理: 设 g 是一连续函数, 则

$$X_n \xrightarrow{a.s.} X \rightsquigarrow g(X_n) \xrightarrow{a.s.} g(X)$$

$$X_n \xrightarrow{P} X \rightsquigarrow g(X_n) \xrightarrow{P} g(X)$$

$$X_n \Longrightarrow X \rightsquigarrow g(X_n) \Longrightarrow g(X)$$

大数定律

- 大数定律 (LLN) (Bernoulli, Chebyshev, Markov, Borel, Cantelli, Kolmogorov, Khinchin)

(a). 弱大数定律：$(X_n;\ n=1,2,\dots)$ i.i.d., 且 $\mu=\mathbb{E}[X_1]$ 存在, 则

$$\overline{X}_n=\frac{1}{n}\sum_{i=1}^{n}X_i \xrightarrow{P} \mu$$

(b). 强大数定律：

$$\overline{X}_n \xrightarrow{a.s.} \mu$$

大数定律

- 大数定律仿真试验：骰子标有 $1-6$ 的数字，X 表示掷骰子出现的点数，则平均点数 $\mathbb{E}[X]=3.5$

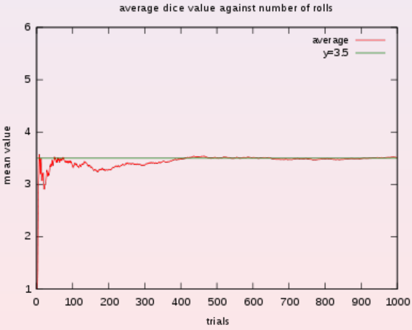


Figure: LLN

中心极限定理

- Lindeberg – Lévy CLT:

  设 $(X_i;\ i = 1, 2, \dots)$ i.i.d. r.v.s., μ, σ^2 为其均值和方差，则 $\overline{X}_n := \frac{1}{n}\sum_{i=1}^n X_i$ 的标准化:

$$\frac{\sqrt{n}\left[\overline{X}_n - \mu\right]}{\sigma} \Longrightarrow N(0,1)$$

中心极限定理

- Lyapunov CLT:

  设 $(X_i;\ i=1,2,\dots)$ 为独立的 r.v.s. μ_i, σ_i^2 为 X_i 均值和方差. 设 $s_n = \sum_{i=1}^n \sigma_i^2$, 若存在 $\delta > 0$ 使 Lyapunov 条件成立:

$$\lim_{n\to\infty} \frac{1}{s_n^{2+\delta}} \sum_{i=1}^{n} \mathbb{E}\left[|X_i - \mu_i|^{2+\delta}\right] = 0,$$

则

$$\frac{1}{s_n} \sum_{i=1}^{n} [X_i - \mu_i] \Longrightarrow N(0,1)$$

中心极限定理

- Lindeberg CLT:

  $\forall\ \varepsilon > 0$, 若 Lindeberg 条件成立 (比 Lyapunov 条件弱):

$$\lim_{n\to\infty} \frac{1}{s_n^2} \sum_{i=1}^{n} \mathbb{E}\left[|X_i - \mu_i|^2 \mathbf{1}_{\{|X_i-\mu_i|>\varepsilon s_n\}}\right] = 0,$$

  则

$$\frac{1}{s_n} \sum_{i=1}^{n} [X_i - \mu_i] \Longrightarrow N(0,1)$$

中心极限定理

- CLT 仿真试验：骰子标有 $1-6$ 的数字

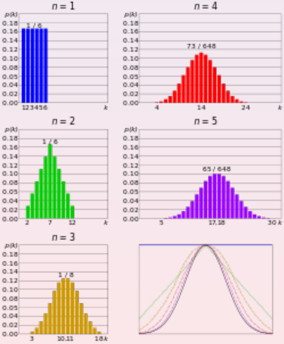


Figure: CLT

完

谢谢！