概率论与数理统计

何志坚

华南理工大学数学学院

hezhijian@scut.edu.cn

第五章：大数定律和中心极限定理

目录

1 大数定律

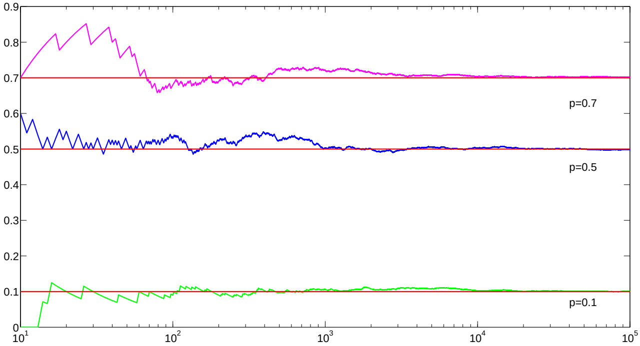
2 中心极限定理

试验1

考虑n重伯努利试验，事件A发生的概率为p. 记

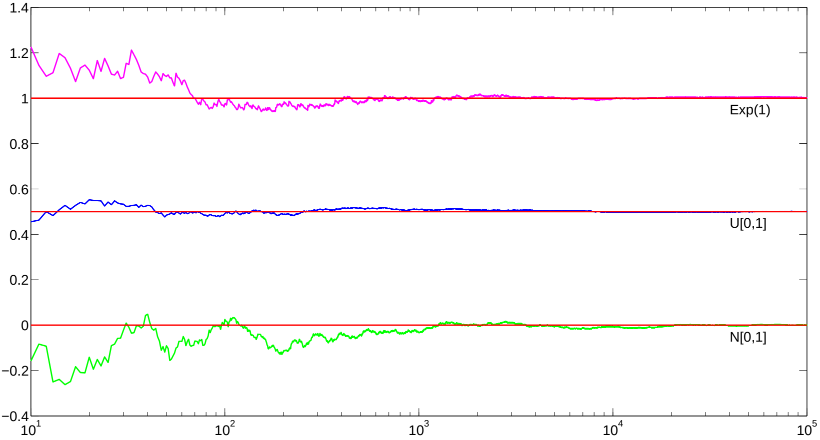
$$X_i = \begin{cases} 1, & \text{第}i\text{次试验事件}A\text{发生} \\ 0, & \text{第}i\text{次试验事件}A\text{不发生} \end{cases}$$

令$S_n = X_1 + X_2 + \cdots + X_n$. 试问$S_n/n$是否会收敛到一个常数?



试验2

对于一般的随机序列X_i呢？试问S_n/n是否会收敛到一个常数？



随机变量序列的收敛

定义： 如果对任意$\epsilon > 0$有

$$\lim_{n\to\infty} P(\omega : |Y_n(\omega) - Y(\omega)| \geq \epsilon) = 0,$$

则称Y_n依概率收敛于Y，记作$Y_n \xrightarrow{\mathrm{P}} Y$.

定义： 如果对任意$\epsilon > 0$有

$$P\left(\omega : \lim_{n\to\infty} Y_n(\omega) = Y(\omega)\right) = 1,$$

则称Y_n以概率1收敛(又称几乎处处收敛)于Y，记作$Y_n \xrightarrow{\mathrm{a.s}} Y$.

注： 概率为0不等于不可能，概率为1不等于一定。

形象解释： 开始上课了，慢慢地大家都安静下来，这是几乎处处收敛。绝大多数同学都安静下来，但每一个人都在不同的时间捣乱，这是依概率收敛。

依概率收敛与以概率1收敛的联系

定理： 如果$Y_n \xrightarrow{\text{a.s}} Y$，则$Y_n \xrightarrow{\text{P}} Y$.

证明： 令

$$C = \{\omega : \lim_{n\to\infty} Y_n(\omega) = Y(\omega)\}.$$

则$P(C) = 1$. 令$A_n = \{\omega : |Y_n(\omega) - Y(\omega)| \geq \epsilon\}$,

$$B_n = \bigcup_{i=n}^{\infty} A_n,$$

$B_\infty = \cap_{n=1}^{\infty} B_n$. 容易发现$C \cap B_\infty = \varnothing$. 所以，$B_\infty \subset \bar{C}$, $P(B_\infty) \leq P(\bar{C}) = 0$, 即$P(B_\infty) = 0$.

$$\lim_{n\to\infty} P(A_n) \leq \lim_{n\to\infty} P(B_n) = P(B_\infty) = 0.$$

注： 弱大数定律涉及依概率收敛，强大数定律涉及以概率1收敛。

伯努利弱大数定律

定理：设S_n为n重伯努利试验中事件A发生的次数，p为事件A出现的概率。则对任意$\epsilon > 0$有

$$\lim_{n\to\infty} P\left(\left|\frac{S_n}{n} - p\right| \geq \epsilon\right) = 0.$$

即$S_n/n \xrightarrow{\mathrm{P}} p$.

注：历史上，伯努利是第一个研究弱大数定理的，他在1713年发表的论文中，提出了上述定理，那是概率论的第一篇论文。

切比雪夫弱大数定律

定理： 设$\{X_n\}$为两两不相关随机变量序列。若每个X_i的方差存在且有共同的上界，即$\mathrm{Var}[X_i] \leq C$, 则对任意$\epsilon > 0$有

$$\lim_{n\to\infty} P\left(\left|\frac{\sum_{i=1}^n X_i}{n} - \frac{\sum_{i=1}^n \mathbb{E}[X_i]}{n}\right| \geq \epsilon\right) = 0.$$

特别地，如果所有的期望相等，即$\mathbb{E}[X_i] = \mu$，则

$$\frac{\sum_{i=1}^n X_i}{n} \xrightarrow{\mathrm{P}} \mu.$$

注： 伯努利弱大数定律是切比雪夫弱大数定律的特例，其中

$$X_i = \begin{cases} 1, & \text{第}i\text{次试验事件}A\text{发生} \\ 0, & \text{第}i\text{次试验事件}A\text{不发生} \end{cases}$$

$\mathrm{Var}[X_i] = p(1-p)$.

切比雪夫不等式

引理： 设k为正整数，随机变量X的k阶原点矩存在，则对任意$\epsilon > 0$有

$$P(|X| \geq \epsilon) \leq \frac{\mathbb{E}[|X|^k]}{\epsilon^k}.$$

切比雪夫不等式： 如果随机变量X的方差存在，则对任意$\epsilon > 0$有

$$P(|X - \mathbb{E}[X]| \geq \epsilon) \leq \frac{\mathrm{Var}[X]}{\epsilon^2}.$$

证明： 在引理中，替换X为$X - \mathbb{E}[X]$并取$k = 2$可得到切比雪夫不等式。只需证引理即可。

$$\begin{aligned} P(|X| \geq \epsilon) &= \mathbb{E}[1_{\{|X| \geq \epsilon\}}] \\ &\leq \mathbb{E}\left[\frac{|X|^k}{\epsilon^k} 1_{\{|X| \geq \epsilon\}}\right] \\ &\leq \mathbb{E}\left[\frac{|X|^k}{\epsilon^k}\right] = \frac{\mathbb{E}[|X|^k]}{\epsilon^k}. \end{aligned}$$

切比雪夫弱大数定律的证明

由切比雪夫不等式知，只需证$\lim_{n\to\infty}\mathrm{Var}[(\sum_{i=1}^n X_n)/n]=0$.

$$\begin{aligned}\mathrm{Var}\left[\frac{\sum_{i=1}^n X_n}{n}\right] &= \frac{1}{n^2}\mathrm{Var}\left[\sum_{i=1}^n X_n\right] \\ &= \frac{1}{n^2}\sum_{i=1}^n \mathrm{Var}\left[X_n\right] (\text{由于两两不相关}) \\ &\le \frac{nC}{n^2} = \frac{C}{n} \to 0.\end{aligned}$$

马尔科夫大数定律

定理：如果

$$\lim_{n\to\infty} \frac{1}{n^2} \mathrm{Var}\left[\sum_{i=1}^{n} X_i\right] = 0,$$

则

$$\lim_{n\to\infty} P\left(\left|\frac{\sum_{i=1}^{n} X_i}{n} - \frac{\sum_{i=1}^{n} \mathbb{E}[X_i]}{n}\right| \geq \epsilon\right) = 0.$$

辛钦大数定律

定理： 设$\{X_n\}$为独立同分布的随机变量序列，具有有限的期望μ，则对任意$\epsilon > 0$有

$$\lim_{n\to\infty} P\left(\left|\frac{\sum_{i=1}^{n} X_i}{n} - \mu\right| \geq \epsilon\right) = 0.$$

注： 辛钦大数定律不需要方差存在，但要求独立同分布。

博雷尔(Borel)强大数定律

定理： 设S_n为n重伯努利试验中事件A发生的次数，p为事件A出现的概率。则

$$P\left(\lim_{n\to\infty}\frac{S_n}{n}=p\right)=1.$$

即$S_n/n \overset{\text{a.s}}{\to} p$.

注： 博雷尔大数定律为伯努利大数定律的“加强”版本。

柯尔莫哥洛夫强大数定律

定理1：设$\{X_n\}$为独立的随机变量序列，具有有限的期望，且

$$\sum_{n=1}^{\infty} \frac{\mathrm{Var}[X_n]}{n^2} < \infty,$$

则

$$P\left(\lim_{n\to\infty} \frac{\sum_{i=1}^{n}(X_i - \mathbb{E}[X_i])}{n} = 0\right) = 1.$$

定理2：设$\{X_n\}$为独立同分布的随机变量序列，具有有限的期望μ, 则

$$P\left(\lim_{n\to\infty} \frac{\sum_{i=1}^{n} X_i}{n} = \mu\right) = 1.$$

注：由定理2可以得到辛钦大数定律。

例

例： 设$\{X_n\}$为独立同分布的随机变量序列，且$\mathbb{E}[X_n]=2, \mathrm{Var}[X_n]=6$，证明当$n\to\infty$，

$$\frac{X_1^2+X_2X_3+X_4^2+X_5X_6+\cdots+X_{3n-2}^2+X_{3n-1}X_{3n}}{n}\xrightarrow{\mathrm{P}} a,$$

并确定常数a的值。

证: 令$Y_k=X_{3k-2}^2+X_{3k-1}X_{3k}$，则$\{Y_k\}$为独立同分布的随机变量序列。又

$$\mathbb{E}[Y_k]=\mathbb{E}[X_{3k-2}^2]+\mathbb{E}[X_{3k-1}]\mathbb{E}[X_{3k}]=\mathrm{Var}[X_1^2]+2(\mathbb{E}[X_1])^2=14.$$

由辛钦大数定律得，$(Y_1+\cdots+Y_n)/n\xrightarrow{\mathrm{P}}14$. 所以$a=14$.

目录

1 大数定律

2 中心极限定理

中心极限定理

考虑随机变量序列$\{X_n\}$, 大数定律研究

$$\bar{X}_n := \frac{X_1 + X_2 + \cdots + X_n}{n}$$

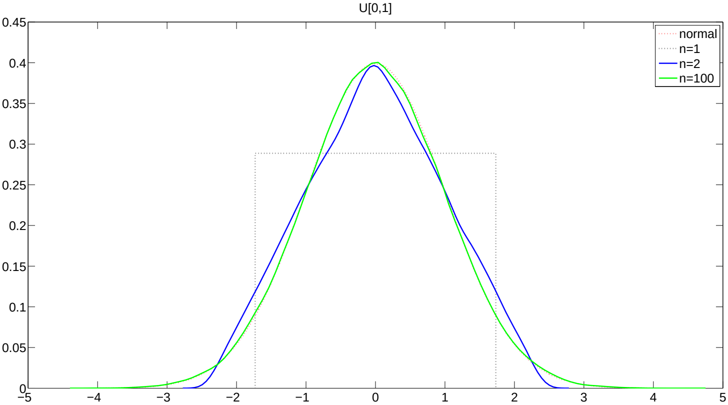
的收敛。中心极限定理研究$\bar{X}_n$的标准化后极限分布，即

$$\xi_n := \frac{\bar{X}_n - \mathbb{E}[\bar{X}_n]}{\sigma(\bar{X}_n)} = \frac{\sum_{i=1}^n (X_i - \mathbb{E}[X_i])}{\sqrt{\mathrm{Var}[\sum_{i=1}^n X_i]}}$$

中心极限定理将告诉我们只要X_n独立同分布且方差有限，ξ_n的极限分布均为标准正态分布，与X_n的分布无关。

注：“统计学之父”卡尔·皮尔逊(Karl Pearson,1857–1936)教授认为：正态分布是上帝赐给人们唯一正确的分布。

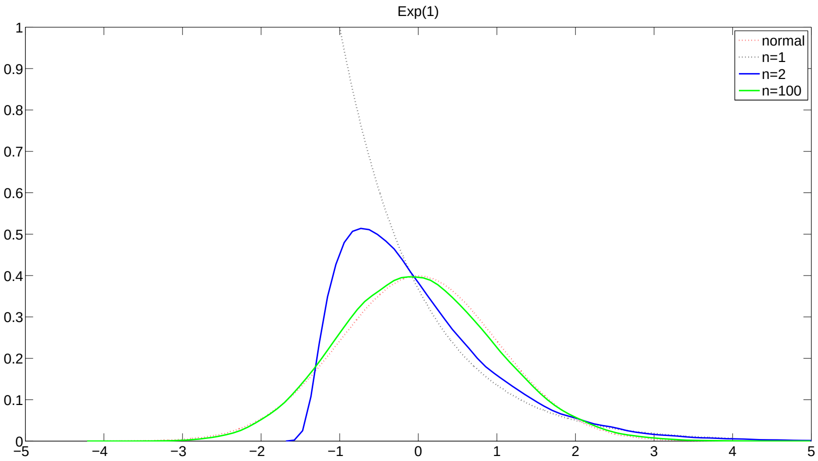
均匀分布



注： 通过12个独立同分布的标准均匀分布可以近似得到标准正态分布，即

$$Y=\sum_{i=1}^{12}X_i-6\approx N(0,1)$$

指数分布



依分布收敛

定义：设Y是连续型随机变量。如果对任意$x \in \mathbb{R}$[1]都有

$$\lim_{n\to\infty} F_{Y_n}(x) = F_Y(x),$$

则称Y_n依分布收敛于Y，记作$Y_n \xrightarrow{\mathrm{d}} Y$.

[1]对于一般的随机变量Y只要求对$F_Y(x)$的任意连续点x.

林德伯格–莱维(Lindberg–Lévy)中心极限定理

定理: 设$\{X_n\}$为独立同分布随机变量序列，且方差期望存在，即$\mathbb{E}[X_n] = \mu$, $\mathrm{Var}[X_n] = \sigma^2$. 令

$$\xi_n := \frac{\sum_{i=1}^n (X_i - \mathbb{E}[X_i])}{\sqrt{\mathrm{Var}[\sum_{i=1}^n X_i]}} = \frac{\sum_{i=1}^n X_i - n\mu}{\sigma\sqrt{n}}.$$

则$\xi_n \xrightarrow{\mathrm{d}} N(0,1)$, 即对任意$x \in \mathbb{R}$,

$$\lim_{n\to\infty} P(\xi_n \le x) = \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\,\mathrm{d}t.$$

棣莫弗–拉普拉斯(De Moivre–Laplace)中心极限定理

定理： 设$\{X_n\}$为独立同分布随机变量序列，同服从$B(1,p)$. 令

$$\xi_n := \frac{\sum_{i=1}^n (X_i - \mathbb{E}[X_i])}{\sqrt{\mathrm{Var}[\sum_{i=1}^n X_i]}} = \frac{\sum_{i=1}^n X_i - np}{\sqrt{np(1-p)}}.$$

则$\xi_n \xrightarrow{\mathrm{d}} N(0,1)$, 即对任意$x \in \mathbb{R}$,

$$\lim_{n\to\infty} P(\xi_n \le x) = \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\,\mathrm{d}t.$$

注1： 棣莫弗–拉普拉斯中心极限定理是林德伯格–莱维中心极限定理的特例。

注2： 更一般的中心极限定理参见林德伯格中心极限定理和李雅普诺夫中心极限定理，这两个定理不需要独立同分布的条件，只需满足相应的条件。

He, Z. & Zhu, L. Stat. Comput. (2017). Asymptotic normality of extensible grid sampling. 链接：`https://doi.org/10.1007/s11222-017-9794-y`

中心极限定理的应用

当n充分大时，可以认为

$$\xi_n = \frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}} \approx N(0,1)$$

或者

$$S_n = \sum_{i=1}^{n} X_i \approx N(n\mu, n\sigma^2)$$

或者

$$\frac{S_n}{n} \approx N(\mu, \sigma^2/n)$$

由此可以计算与S_n相关的事件的概率。比如

$$P(S_n \leq x) \approx \Phi\left(\frac{x - n\mu}{\sqrt{n}\sigma}\right)$$

例1

假设报名学习某门选修课的学生人数是服从均值100的泊松分布的随机变量。学校领导决定，如果报名人数不少于120人，就分成两个班上。如果少于120人，就集中在一个班上。试问该选修班分成两个班上的概率是多少?

解： 设报名人数为$X \sim Pois(100)$. 设$X_i \overset{\text{iid}}{\sim} Pois(1)$。由泊松分布的可加性得

$$X = \sum_{i=1}^{100} X_i$$

$\mathbb{E}[X] = \mathrm{Var}[X] = 100$. 由中心极限定理知，

$$P(X \geq 120) = 1 - \Phi((120 - 100)/\sqrt{100}) = 1 - \Phi(2) = 0.0228.$$

例2

例： 设某地区原有一家小电影院，现拟筹建一所较大的电影院。根据分析，该地区每天平均看电影者约有$n=1600$人，预计新电影院开业后，平均约有$3/4$的观众将去新电影院。现计划其座位数，要求座位数尽可能多，但"空座达到200或更多"的概率不能超过0.1，问设多少座位为好?

解： 设座位数为m. 设

$$X_i=\begin{cases}1, & \text{第}i\text{个观众去新院}\\ 0, & \text{其他}.\end{cases}$$

则$X_i \overset{\text{iid}}{\sim} B(1,3/4)$. 令$S_n=X_1+\cdots+X_n$. 依题意，

$$P(m-S_n\geq 200)=P(S_n\leq m-200)\leq 0.1.$$

由中心极限定理得，

$$P(S_n\leq m-200)\approx \Phi((m-200-3/4*1600)/\sqrt{1600*(3/4)(1-3/4)}).$$

所以，$m\leq \Phi^{-1}(0.1)*10\sqrt{3}+1400\approx 1377.8$. 应设1377个座位。

例3

例：用调查对象中的收看比例k/n作为某电视节目的收视率p的估计。要有90%的把握，使k/n与p的差异不大于0.05，问至少要调查多少个对象？

解：设S_n表示n个调查对象中收看此节目的人数，则$S_n \sim B(n,p)$. 依题得，

$$P(|S_n/n-p| \le 0.05) \ge 0.9.$$

由中心极限定理得

$$P(|S_n/n-p| \le 0.05) = 2\Phi(0.05/\sqrt{p(1-p)/n}) - 1 \ge 0.9.$$

所以，

$$n \ge 400\Phi^{-1}(0.95)^2 p(1-p) \ge 100\Phi^{-1}(0.95)^2 = 100 * 1.645^2 = 270.6.$$

至少调查271个对象。

作业

第五章习题：1, 2, 4, 8, 10, 11

5月23日（周三）交！