第五章：大数定律和中心极限定理

5.1 大数定律

问题引入 I

我们在生产线上一个接一个的检查产品的合格情况。记X_i为第i个产品的不合格数，$X_i \sim B(1,p)$。记$v_n = \sum_{i=1}^{n} X_i$，以及n次检查的不合格率$b_n = \frac{v_n}{n}$，考虑下面两个问题

1. 是否有$\lim_{n\to+\infty} b_n = p$?
2. 是否存在某个随机变量Y，使得$\lim_{n\to\infty} v_n = Y$?

回顾：概率的统计定义

频率法：定义1.2.1

为了考察某一随机试验的随机事件A发生的频率，我们重复的进行这一随机试验，并计算下面的数值

$$f_n(A) = \frac{n_A}{n}$$

其中n是试验的总次数，n_A为事件A发生的次数。随着试验次数的增加，准确的说当$n \to \infty$时，

$$\lim_{n\to\infty} f_n(A) = p, p \in [0, 1],$$

p称为事件A发生的概率。

依概率收敛与以概率 1 收敛

一般地，设有随机变量序列 $X_1, X_2, \cdots$ 和随机变量 Y

- 若 $\forall \epsilon > 0$ 有

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

则称随机变量序列 $X_1, X_2, \cdots$ 依概率收敛于随机变量 Y，记作 $X_n \xrightarrow{P} Y$.

- 若

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

则称随机变量序列 $X_1, X_2, \cdots$ 以概率 1 收敛（几乎处处收敛）于随机变量 Y，记作 $X_n \xrightarrow{a.s.} Y$. (a.s. almost surely)

弱大数定律

若对任意的$\epsilon > 0$，存在确定数列a_1，a_2，$\cdots$，使得

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

则$X_1, X_2, \cdots$服从弱大数定律.

具体地，在本章的学习中，我们将探寻形如

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

的大数定律.

强大数定律

存在确定数列 a_1，a_2，$\cdots$，使得

$$P\left(\omega : \lim_{n\to\infty}\left(\frac{1}{n}\sum_{i=1}^{n} X_i(\omega) - \frac{1}{n}\sum_{i=1}^{n} a_i\right) = 0\right) = 1,$$

则 $X_1, X_2, \cdots$ 服从强大数定律.

具体地，在强大数定律的学习中，将探寻形如

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

的大数定律.

大数定律区别

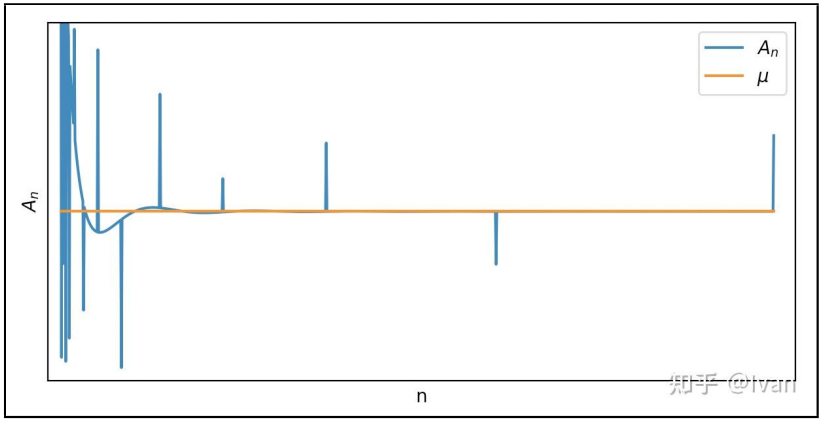


Figure: 弱大数定律

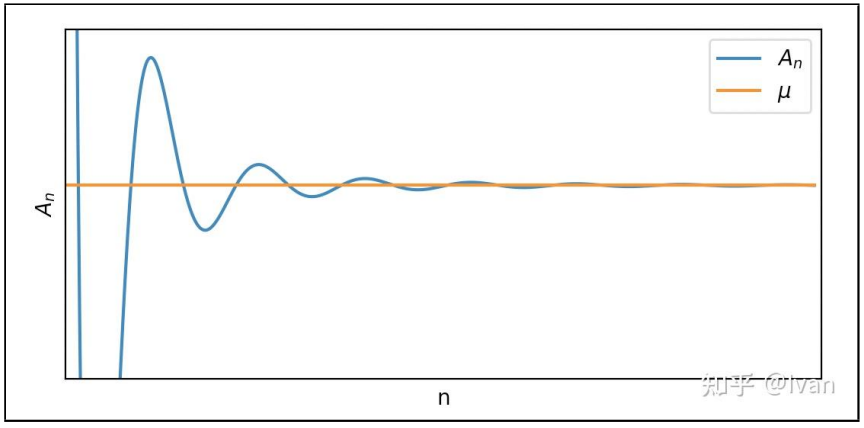


Figure: 强大数定律

几个常用的弱大数定律

1. 伯努利大数定律
2. 辛钦大数定律
3. 切比雪夫大数定律
4. 马尔可夫大数定律*

切比雪夫不等式 I

切比雪夫不等式

设随机变量 X 的方差 $Var[X]$ 存在，则对任意的 $\epsilon > 0$，有

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

切比雪夫不等式 II

证明：对于连续型随机变量

$$
\begin{aligned}
P(|X - E[X]| \geq \epsilon) &= \int_{|x-E[X]|\geq\epsilon} f_X(x)\,\mathrm{d}x \\
&\leq \int_{|x-E[X]|\geq\epsilon} \frac{(x-E[X])^2}{\epsilon^2} f_X(x)\,\mathrm{d}x \\
&\leq \int_{-\infty}^{\infty} \frac{(x-E[X])^2}{\epsilon^2} f_X(x)\,\mathrm{d}x \\
&= \frac{Var[X]}{\epsilon^2}.
\end{aligned}
$$

切比雪夫不等式 III

对于离散型随机变量

$$
\begin{aligned}
P(|X - E[X]| \geq \epsilon) &= \sum_{|x_i - E[X]| \geq \epsilon} p_i \\
&\leq \sum_{|x_i - E[X]| \geq \epsilon} \frac{(x_i - E[X])^2}{\epsilon^2} p_i \\
&\leq \sum_{i} \frac{(x_i - E[X])^2}{\epsilon^2} p_i \\
&= \frac{Var[X]}{\epsilon^2}.
\end{aligned}
$$

切比雪夫不等式的等价表述

1. $P(|X - E[X]| \geq k\sqrt{Var[X]}) \leq \frac{1}{k^2}$.
2. $P(|X - E[X]| < \epsilon) \geq 1 - \frac{Var[X]}{\epsilon^2}$.

练习

1. 随机变量 X 服从 $[a,5]$ 上的均匀分布，且由切比雪夫不等式

$$P(|X-3|<\varepsilon)\geq 0.99$$

求 a 和 ε 值。

2. 设 X 的数学期望 $E[X]=\mu$, 方差 $\mathrm{Var}[X]=\sigma^2$, 用切比雪夫不等式估计 $P(|X-\mu|\geq 2.5\sigma)$。若 $X\sim N(\mu,\sigma^2)$，对 $P(|X-\mu|\geq 2.5\sigma)$ 直接计算，并与估计值做比较。

伯努利大数定律

对于n重伯努利试验，弱大数定律阐述如下：

伯努利大数定律

设v_n为n重伯努利试验中事件A发生的次数（或指代n重伯努利试验中成功的次数），称$\frac{v_n}{n}$为事件A发生的频率，p为一次试验中事件A发生的概率.那么对任意的$\epsilon > 0$，有

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

当试验次数足够多时，频率将稳定于概率

伯努利大数定律

对于n重伯努利试验，弱大数定律阐述如下：

伯努利大数定律

设v_n为n重伯努利试验中事件A发生的次数（或指代n重伯努利试验中成功的次数），称$\frac{v_n}{n}$为事件A发生的频率，p为一次试验中事件A发生的概率.那么对任意的$\epsilon > 0$，有

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

当试验次数足够多时，频率将稳定于概率

伯努利大数定律使用条件是：X_1，X_2，$\cdots$为独立同分布的两点分布.

辛钦弱大数定律

辛钦弱大数定律

设 X_1，X_2，$\cdots$ 为独立同分布的随机变量序列，且具有相同的数学期望 μ，则对于任意的 $\epsilon > 0$ 有

$$\lim_{n\to\infty} P\left(\left|\frac{X_1 + X_2 + \cdots + X_n}{n} - \mu\right| \geq \epsilon\right) = 0.$$

辛钦弱大数定律的使用条件是

1. 随机变量序列 X_1，X_2，$\cdots$ 相互独立且同分布，
2. X_i 的期望（均值）存在且相同.

切比雪夫弱大数定律 I

定理5.1.1（切比雪夫弱大数定律）

设 X_1，X_2，$\cdots$ 为独立的随机变量序列，$E[X_i] = \mu$，$Var[X_i] \le C$，$i = 1, 2, \cdots,$，则对任意的 $\epsilon > 0$，有

$$\lim_{n\to\infty} P\left(\left|\frac{X_1 + X_2 + \cdots + X_n}{n} - \mu\right| \ge \epsilon\right) = 0$$

切比雪夫弱大数定律的使用条件是

1. 随机变量序列 X_1，X_2，$\cdots$ 相互独立，
2. X_i 有相同的期望（均值）（并不一定同分布），
3. X_i 的方差有公共上界.

切比雪夫弱大数定律 II

证明：记随机变量 $Y_n = \frac{X_1+X_2+\cdots+X_n}{n}$，那么由于期望是线性的，有

$$E[Y_n] = \mu.$$

由 X_1，X_2，$\cdots$，X_n 的独立性，有

$$Var[Y_n] = \frac{\sum_{i=1}^{n} Var[X_i]}{n^2} \leq \frac{C}{n}.$$

由切比雪夫不等式

$$0 \leq \lim_{n\to\infty} P(|Y_n - \mu| \geq \epsilon) \leq \lim_{n\to\infty} \frac{C}{n\epsilon} = 0.$$

马尔可夫大数定律*

马尔可夫大数定律

对于随机变量序列$\{X_n\}$，若有

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

成立，那么$\{X_n\}$服从马尔可夫大数定律，即对任意的$\epsilon>0$，有

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

马尔可夫弱大数定律的重要性在于，对随机变量序列已经没有任何独立性、同分布、不相关的设定，只是要求他们具有方差且满足定理要求.

大数定律总结

常见的弱大数定律	需要满足的条件
伯努利大数定律	v_n服从二项分布，单个样本X_n 服从独立同分布的伯努利分布
辛钦大数定律	X_n 独立同分布 X_n 的数学期望存在
切比雪夫大数定律	X_n独立，且有相同期望和方差上界
马尔可夫大数定律	没有其它的要求，只要求 $\frac{1}{n^2} Var\left(\sum_{i=1}^{n} X_i\right) \to 0$。

例子

1. 设 X_k 为独立的随机变量序列，且

$$P(X_k = \pm 2^k) = \frac{1}{2^{2k+1}}, \quad P(X_k = 0) = 1 - \frac{1}{2^{2k}}, \ k = 1, 2, \cdots$$

证明 $\{X_k\}$ 服从弱大数定律.

2. 设 $\{X_n\}$ 为独立随机变量序列，且

$$P(X_n = 1) = p_n, \quad P(X_n = 0) = 1 - p_n, \ n = 1, 2, \cdots$$

证明 $\{X_n\}$ 服从弱大数定律.

3. 设 $\{X_n\}$ 为独立同分布随机变量序列，且都服从参数为 λ 的泊松分布， 证明 $\{X_n\}$ 服从弱大数定律.

例子

1. 设 X_k 为独立的随机变量序列，且

$$P(X_k = \pm 2^k) = \frac{1}{2^{2k+1}}, \quad P(X_k = 0) = 1 - \frac{1}{2^{2k}}, \ k = 1, 2, \cdots$$

证明 $\{X_k\}$ 服从弱大数定律.

2. 设 $\{X_n\}$ 为独立随机变量序列，且

$$P(X_n = 1) = p_n, \quad P(X_n = 0) = 1 - p_n, \ n = 1, 2, \cdots$$

证明 $\{X_n\}$ 服从弱大数定律.

3. 设 $\{X_n\}$ 为独立同分布随机变量序列，且都服从参数为 λ 的泊松分布，证明 $\{X_n\}$ 服从弱大数定律.

4. (习题 5.5) 设 $\{X_n\}$ 是独立同分布的随机变量序列，且 $E[X_n] = 2$，$Var[X_n] = 6$，证明

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

并确定 a 的值.

几个常用的强大数定律*

1. 柯尔莫哥洛夫强大数定律
2. 博雷尔强大数定律

柯尔莫哥洛夫强大数定律*

柯尔莫哥洛夫强大数定律

1. 设 $X_1, X_2, \ldots$ 为独立随机变量序列，具有有限的数学期望，且

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

则

$$P\left(\lim_{n\to\infty} n^{-1} \sum_{k=1}^{n} (X_k - E[X_k]) = 0\right) = 1.$$

2. 设 $X_1, X_2, \ldots$ 为独立同分布随机变量序列，具有有限的数学期望 μ，则

$$P\left(\lim_{n\to\infty} n^{-1} \sum_{k=1}^{n} (X_k - \mu) = 0\right) = 1.$$

博雷尔（Borel）强大数定律*

博雷尔强大数定律

对于n重伯努利试验，记 v_n 为独立试验成功的次数，p 为一次试验成功的概率，则

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

5.2 中心极限定理

中心极限定理

$X_1, X_2, \cdots, X_n$独立同分布，问

$$Y_n = X_1 + X_2 + \cdots + X_n$$

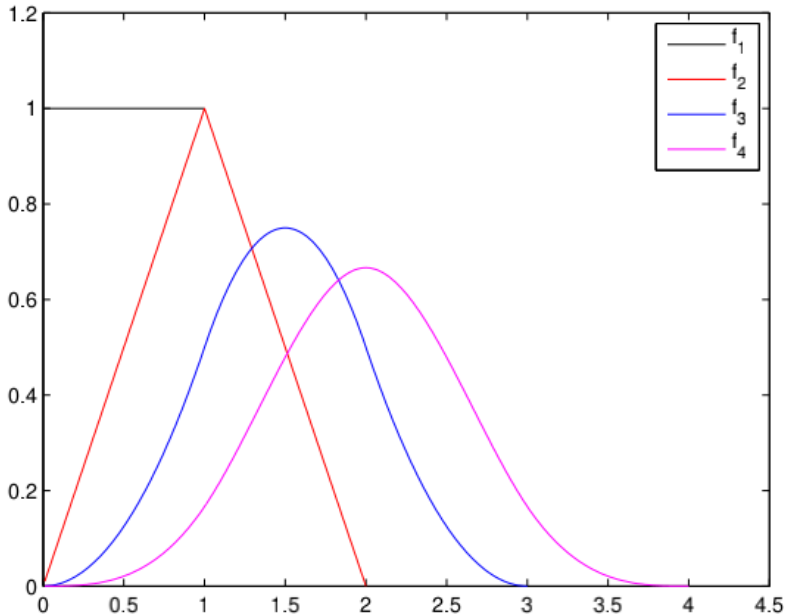
的分布函数是什么？当$n \to \infty$，这个极限分布是什么？

中心极限定理引入 I

设$\{X_n\}$为独立同分布的随机变量序列，且$X_i \sim B(1, 1/2)$，记$Y_n = \sum_{i=1}^{n} X_i \sim B(n, 1/2)$

中心极限定理引入 II

设$\{X_n\}$为独立同分布的随机变量序列，且$X_i \sim U(0,1)$，记$Y_n = \sum_{i=1}^{n} X_i$，设Y_1, Y_2, Y_3, Y_4的分布密度函数分别为f_1, f_2, f_3, f_4，有



随着n的增加，f_n越来越接近正态分布的密度函数。但$\sum_{i=1}^{n} X_i$的均值与方差都随着n变大而趋于无穷大。

中心极限定理引入 III

问题提法

当$n \to \infty$，$E[Y_n] \to \infty$，$Var[Y_n] \to \infty$，即Y_n不收敛到一个有限的分布。我们对Y_n作标准化

$$Y_n^* = \frac{Y_n - E[Y_n]}{\sqrt{Var[Y_n]}}$$

$E[Y_n^*] = 0$，$Var[Y_n^*] = 1$，此时Y_n^*就有可能用标准正态分布去代替，即

$$Y_n^* \to Z \sim N(0,1)$$

中心极限定理

定义 5.2.1

设 X_1，X_2，$\cdots$ 为随机变量序列，具有有限的数学期望和方差，记 $Y_n = \sum_{i=1}^{n} X_i$，若有

$$Z_n = \frac{Y_n - E[Y_n]}{\sqrt{Var[Y_n]}} \xrightarrow{d} Z, \quad Z \sim N(0,1),$$

则称 X_1，X_2，$\cdots$ 服从中心极限定理.

中心极限定理

1. 林德伯格-莱维 中心极限定理
2. 棣莫弗－拉普拉斯 中心极限定理

林德伯格-莱维 中心极限定理

定理 5.2.1

设 X_1，X_2，$\cdots$ 为独立同分布的随机变量序列，具有有限的的数学期望 μ 和方差 σ^2，那么 X_1，X_2，$\cdots$ 服从中心极限定理，即

$$\lim_{n\to\infty} P\left(\frac{1}{\sigma\sqrt{n}}\left[\sum_{k=1}^{n} X_k - n\mu\right] \le x\right) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}}\,\mathrm{d}t.$$

林德伯格-莱维 中心极限定理应用

对独立同分布随机变量序列的和 $\sum_{k=1}^{n} X_k$ 的分布可用标准正态分布近似计算。即当 n 较大时，有：

$$P\left(\sum_{k=1}^{n} X_k \leq b\right) \approx \Phi\left(\frac{b-n\mu}{\sigma\sqrt{n}}\right),$$

$$P\left(a \leq \sum_{k=1}^{n} X_k \leq b\right) \approx \Phi\left(\frac{b-n\mu}{\sigma\sqrt{n}}\right) - \Phi\left(\frac{a-n\mu}{\sigma\sqrt{n}}\right),$$

其中 μ 和 σ^2 分别为 X_k 的数学期望和方差，$\Phi(x)$ 表示标准正态分布的累积分布函数。

例题

设某银行服务窗口接待一位顾客的服务时间（单位：min）服从参数为 $1/10$ 的指数分布。

1. 求 $8h$ 以内该服务窗口能接待 48 位顾客的近似概率。
2. 若 $8h$ 以内该服务窗口能完成接待 n 位顾客任务的概率达 99%，顾客数 n 最多是多少？

棣莫弗－拉普拉斯 中心极限定理

定理 5.2.2

设 X_1，X_2，$\cdots$ 为独立同分布的随机变量序列，且都服从 $B(1,p)$，那么 X_1，X_2，$\cdots$ 服从中心极限定理，即

$$\lim_{n\to\infty} P\left(\frac{\sum_{k=1}^{n} X_k - np}{\sqrt{np(1-p)}} \le x\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}}\,\mathrm{d}t.$$

注：这个定理与泊松定理都是对二项分布的近似，当 $p \ll 1$ 的时候，用泊松分布近似较好，当 p 并不十分小的时候，用正态分布较好。

棣莫弗－拉普拉斯 中心极限定理的应用

若 Y_n 服从二项分布 $B(n,p)$，a,b 是两个非负整数且 $a<b$，当 n 很大时，有二项分布的近似计算公式：

$$P(a \le Y_n \le b) \approx \Phi\left(\frac{b-np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{a-np}{\sqrt{np(1-p)}}\right).$$

当 n 不太大时，这个近似公式有一个修正公式可提高计算精度：

$$P(a \le Y_n \le b) \approx \Phi\left(\frac{b+0.5-np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{a-0.5-np}{\sqrt{np(1-p)}}\right),$$

例题

1. 给定 n 和 x，求概率.
   - 100 个独立工作（工作的概率为 0.9）的部件组成一个系统，求系统中至少有 85 个部件工作的概率.
2. 给定 n 和概率，求 x.
   - 有 200 台独立工作（工作的概率为 0.7）的机床，每台机床工作时需 15kw 电力. 问共需多少电力, 才可有 95% 的可能性保证正常生产？
3. 给定 x 和概率，求 n.
   - 工厂生产的一批产品由于数量大，无法知道其次品率 p。现从这批产品中抽出 n 件产品进行检测。问 n 至少多大才能使所抽出的 n 件产品的次品率与全部产品的次品率 p 相差不超过 5% 的概率不小于 95%?