第 16 讲 大数定律、中心极限定理

16.1 依概率收敛

若 $P\{A\}=1$, 则称事件 A 几乎必然发生.

若 $P\{Y=a+bX\}=1$, 则称 $Y=a+bX$ 几乎处处成立.

定义 16.1

设 $Y_1, Y_2, \cdots, Y_n, \cdots$ 是一个随机变量序列, a 是一个常数. 若对于任意正数 ε, 有

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

则称序列 $Y_1, Y_2, \cdots, Y_n, \cdots$ 依概率收敛于 a, 记为

$$Y_n \xrightarrow{P} a.$$

依概率收敛是收敛的推广, 若序列收敛, 则必依概率收敛. 收敛序列的任意子序列必收敛, 但依概率收敛序列可能存在子序列不一定依概率收敛. 即 $Y_n \xrightarrow{P} a.$, Y_{n_i} 为 Y_n 的子序列, $Y_{n_i} \xrightarrow{P} a$. 不一定成立.

命题 16.1

设 $X_n \xrightarrow{P} a, Y_n \xrightarrow{P} b$, 函数 $g(x, y)$ 在点 (a, b) 连续, 则

$$g\left(X_n, Y_n\right) \xrightarrow{P} g(a, b).$$

16.2 大数定律

令

$$\overline{X}=\frac{1}{n} \sum_{k=1}^{n} X_k.$$

数理统计中, $\overline{X}=\frac{1}{n} \sum_{k=1}^{n} X_k$ 称为样本均值.

16.2.1 切比雪夫不等式

引理 16.1

设 X 具有 $E(X)=\mu$, $D(X)=\sigma^2$, 则对任意正数 ϵ, 有

$$P\{|X-\mu| \geq \epsilon\} \leq \frac{\sigma^2}{\epsilon^2}$$

意义：分布未知, $E(X)$ 和 $D(X)$ 存在的条件下, 估计

$$P\{|X-E(X)|<\epsilon\} \geq 1-\frac{\sigma^2}{\epsilon^2}.$$

16.2.2 切比雪夫大数定律

定理 16.1 (切比雪夫大数定律)

设随机变量 $X_1, X_2, \cdots, X_n, \cdots$ 独立, 且存在常数 C, 使得 $D(X_i)<C, \forall i$. 则

$$\overline{X} \xrightarrow{P} E(\overline{X})=\frac{1}{n}\sum_{k=1}^{n} E(X_k).$$

证明 由切比雪夫不等式,

$$1 \geq P\left\{|\overline{X}-E\left(\overline{X}\right)|<\epsilon\right\} \geq 1-\frac{D\left(\overline{X}\right)}{\epsilon^2}.$$

其中

$$E\left(\overline{X}\right)=E\left(\frac{1}{n}\sum_{k=1}^{n} X_k\right)=\frac{1}{n}\sum_{k=1}^{n} E\left(X_k\right),$$

$$D\left(\overline{X}\right)=D\left(\frac{1}{n}\sum_{k=1}^{n} X_k\right)=\frac{1}{n^2}\sum_{k=1}^{n} D\left(X_k\right)<\frac{C}{n}.$$

所以

$$1 \geq P\left\{|\overline{X}-\frac{1}{n}\sum_{k=1}^{n} E\left(X_k\right)|<\epsilon\right\} \geq 1-\frac{D\left(\overline{X}\right)}{\epsilon^2} \geq 1-\frac{C}{n}.$$

令 $n \to \infty$, 则

$$\lim_{n\to\infty} P\left\{|\overline{X}-\frac{1}{n}\sum_{k=1}^{n} E\left(X_k\right)|<\epsilon\right\}=1.$$

16.2.3 辛钦大数定律

定理 16.2 (弱大数定律/辛钦大数定律)

设随机变量 $X_1, X_2, \cdots, X_n, \cdots$ 独立同分布且 $E\left(X_k\right)=\mu\ (k=1,2,\cdots)$, 则

$$\overline{X} \xrightarrow{P} E(\overline{X})=\mu$$

16.2.4 伯努利大数定律

定理 16.3 (伯努利大数定律)

$$\text{频率} \xrightarrow{P} \text{概率}$$

证明: 切比雪夫不等式 $\Longrightarrow$ 切比雪夫大数定律 $\Longrightarrow$ 弱大数定律 $\Longrightarrow$ 伯努利大数定律.

16.3 中心极限定理

记

$$Y_n = \sum_{k=1}^{n} X_k \qquad\qquad Y^* = \frac{Y - E(Y)}{\sqrt{D(Y)}}.$$

16.3.1 独立同分布的中心极限定理

定理 16.4 (独立同分布的中心极限定理)

设随机变量 $X_1, X_2, \cdots, X_n, \cdots$ 独立同分布, 则当 n 充分大时,

$$Y_n^* = \frac{\sum_{k=1}^{n} X_k - n\mu}{\sqrt{n}\sigma} \xrightarrow{\text{近似}} N(0,1).$$

其中 $\mu = E(X_k), \sigma = \sqrt{D(X_k)}$.

16.3.2 独立 (不一定同分布) 的中心极限定理

定理 16.5 (李雅普诺夫定理)

设随机变量 $X_1, X_2, \cdots, X_n, \cdots$ 独立, 则当 n 充分大时,

$$Y_n^* = \frac{\sum_{k=1}^{n} X_k - \sum_{k=1}^{n} \mu_k}{\sqrt{\sum_{k=1}^{n} \sigma_k^2}} \xrightarrow{\text{近似}} N(0,1).$$

其中 $\mu_k = E(X_k), \sigma_k^2 = D(X_k)$.

16.3.3 二项分布的中心极限定理

定理 16.6 (棣莫弗—拉普拉斯定理)

设随机变量 $\eta_n \sim b(n,p)$, 则当 n 充分大时,

$$\eta_n^* = \frac{\eta_n - np}{\sqrt{np(1-p)}} \xrightarrow{\text{近似}} N(0,1).$$

16.4 例题

例题 16.1 一加法器同时收到 20 个噪声电压 V_k $(k = 1, ..., 20)$, 设它们是独立同分布的，且 $V_k \sim U(0, 10)$. 记 $V = \sum_{k=1}^{20} V_k$, 求 $P\{V > 105\}$ 的近似值.

例题 16.2 一艘船舶在某海区航行，已知每遭受一次波浪的冲击，纵摇角大于 3° 的概率为 $p = 1/3$，若船舶遭受了 90000 次波浪冲击，问其中有 29500 ~ 30500 次纵摇角度大于 3° 的概率是多少？

例题 16.3 对于一个学生而言，来参加家长会的家长人数是一个随机变量，设一个学生无家长、有 1 名家长、有 2 名家长来参加会议的概率分别为 0.05、0.8、0.15. 若学校共有 400 名学生，设各学生参加会议的家长人数相互独立，且服从同一分布.

(1) 求参加会议的家长人数 X 超过 450 的概率.

(2) 求有 1 名家长来参加会议的学生人数不多于 340 的概率.

16.5 第 2、3、4 章自测题

例题 16.4 设二维随机变量 (X,Y) 的概率密度为

$$f(x,y)=\begin{cases} ce^{-y}, & 0<x<y<+\infty, \\ 0, & \text{else} \end{cases}$$

求

(1) 常数 c;

(2) $f_X(x)$ 和 $f_Y(y)$;

(3) 问 X,Y 是否相互独立? 为什么?

(4) $f_{X|Y}(x|y)$;

(5) $P\{X>1 \mid Y=2\}$;

(6) (X,Y) 的联合分布函数;

(7) $P\{X<1 \mid Y<2\}$;

(8) $Z=X+Y$ 的概率密度函数;

(9) 分别用 $f_Z(z)$ 和 $f(x,y)$ 计算 $P\{X+Y<2\}$;

(10) $P\{\min(X,Y)<2\}$;

(11) 设 X,Y 独立, 概率密度分别为 $f_X(x)$ 和 $f_Y(y)$, 求 $Z=X+Y$ 的概率密度函数;

(12) 设 X,Y 独立, 概率密度分别为 $f_X(x)$ 和 $f_Y(y)$, 求 $P\{\min(X,Y)<2\}$ 和 $P\{\max(X,Y)<2\}$;

(13) 分别利用 $f_X(x)$ 和 $f(x,y)$ 计算 $E(X)$;

(14) $D(X), D(Y)$;

(15) $Cov(X,Y)$;

(16) ρ_{XY};

(17) $Z_1=X+Y, Z_2=X-2Y$, 计算 $E(Z_1), E(Z_2), D(Z_1), D(Z_2), Cov(Z_1,Z_2), \rho_{Z_1Z_2}$.

解 翻书写下所用公式.

例题 16.5 将上题概率密度函数换为

$$f(x,y)=\begin{cases} cxe^{-y}, & 0<x<y<+\infty, \\ 0 & \text{其他}. \end{cases}.$$

解 闭卷独立作答自检.