

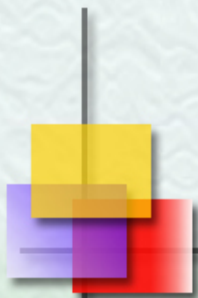
第五章 大数定律及中心极限定理

习 题 课

一、重点与难点

二、主要内容

三、典型例题



一、重点与难点

1.重点

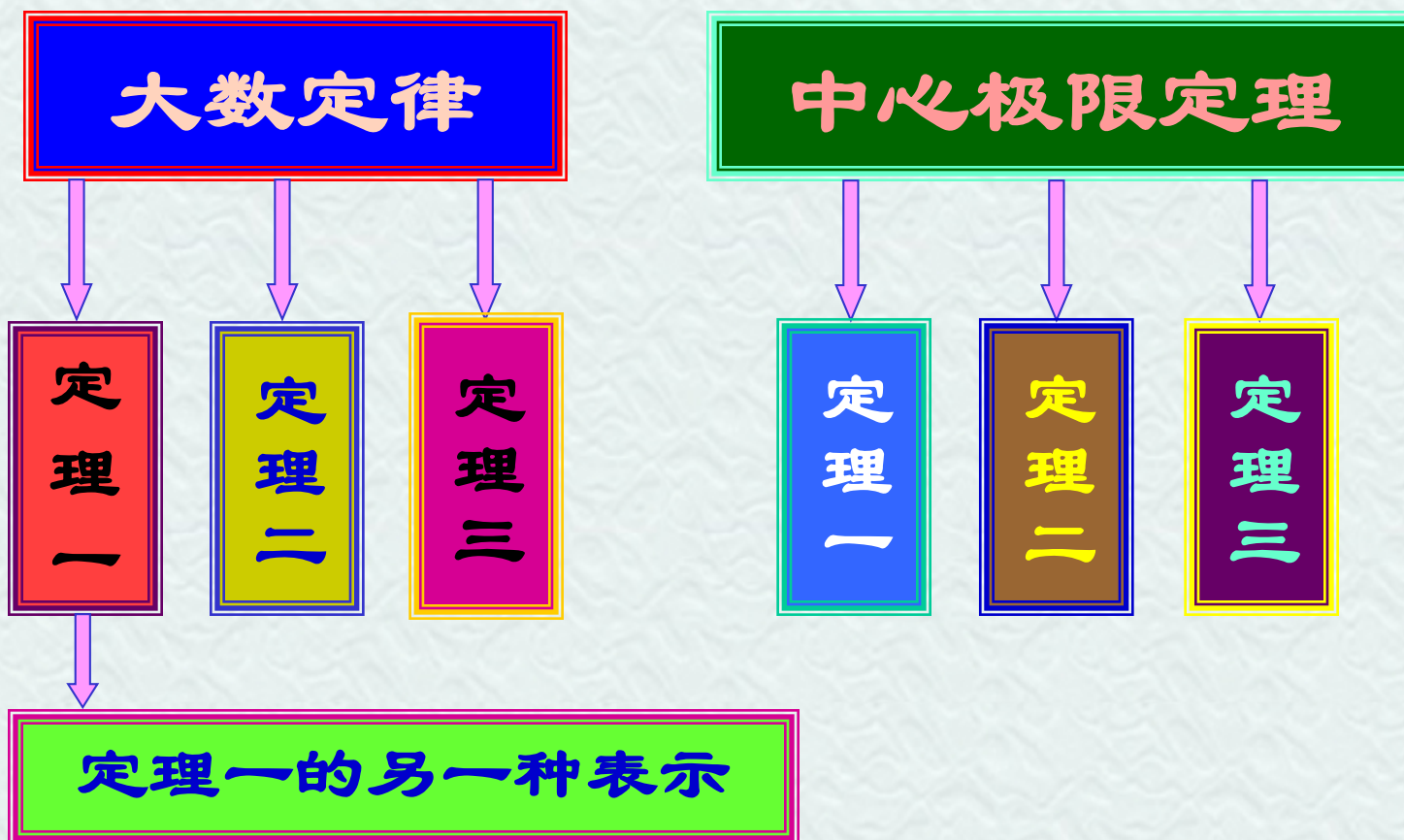
中心极限定理及其运用.

2.难点

证明随机变量服从大数定律.



二、主要内容



契比雪夫定理的特殊情况

设随机变量 $X_1, X_2, \dots, X_n, \dots$ 相互独立, 且具有相同的数学期望和方差: $E(X_k) = \mu$, $D(X_k) = \sigma^2$ ($k = 1, 2, \dots$), 作前 n 个随机变量

的算术平均 $\bar{X} = \frac{1}{n} \sum_{k=1}^n X_k$, 则对于任意正

数 ε 有

$$\lim_{n \rightarrow \infty} P\{|\bar{X} - \mu| < \varepsilon\} = \lim_{n \rightarrow \infty} P\left\{\left|\frac{1}{n} \sum_{k=1}^n X_k - \mu\right| < \varepsilon\right\} = 1.$$



定理一的另一种表示

设随机变量 $X_1, X_2, \dots, X_n, \dots$ 相互独立,
且具有相同的数学期望和方差: $E(X_k) = \mu$,

$D(X_k) = \sigma^2$ ($k = 1, 2, \dots$), 则序列 $\bar{X} = \frac{1}{n} \sum_{k=1}^n X_k$

依概率收敛于 μ , 即 $\bar{X} \xrightarrow{P} \mu$.



伯努利大数定理

设 n_A 是 n 次独立重复试验中事件 A 发生的次数, p 是事件 A 在每次试验中发生的概率, 则对于任意正数 $\varepsilon > 0$, 有

$$\lim_{n \rightarrow \infty} P\left\{\left|\frac{n_A}{n} - p\right| < \varepsilon\right\} = 1 \text{ 或 } \lim_{n \rightarrow \infty} P\left\{\left|\frac{n_A}{n} - p\right| \geq \varepsilon\right\} = 0.$$



辛钦定理

设随机变量 $X_1, X_2, \dots, X_n, \dots$ 相互独立, 服从同一分布, 且具有数学期望 $E(X_k) = \mu$ ($k = 1, 2, \dots$), 则对于任意正数 ε , 有

$$\lim_{n \rightarrow \infty} P \left\{ \left| \frac{1}{n} \sum_{k=1}^n X_k - \mu \right| < \varepsilon \right\} = 1.$$



独立同分布的中心极限定理

设随机变量 $X_1, X_2, \dots, X_n, \dots$ 相互独立, 服从同一分布, 且具有数学期望和方差: $E(X_k) = \mu$, $D(X_k) = \sigma^2 > 0$ ($k = 1, 2, \dots$), 则随机变量之和的

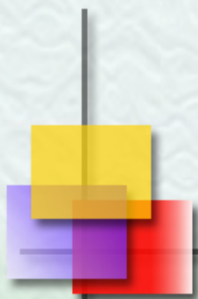
标准化变量 $Y_n = \frac{\sum_{k=1}^n X_k - E\left(\sum_{k=1}^n X_k\right)}{\sqrt{D\left(\sum_{k=1}^n X_k\right)}}.$



的分布函数 $F_n(x)$ 对于任意 x 满足

$$\lim_{n \rightarrow \infty} F_n(x) = \lim_{n \rightarrow \infty} P \left\{ \frac{\sum_{k=1}^n X_k - n\mu}{\sqrt{n} \sigma} \leq x \right\}$$

$$= \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = \Phi(x).$$



李雅普诺夫定理

设随机变量 $X_1, X_2, \dots, X_n, \dots$ 相互独立, 它们具有数学期望和方差:

$$E(X_k) = \mu_k, \quad D(X_k) = \sigma_k^2 \neq 0 \quad (k = 1, 2, \dots),$$

记
$$B_n^2 = \sum_{k=1}^n \sigma_k^2,$$

若存在正数 δ , 使得当 $n \rightarrow \infty$ 时,

$$\frac{1}{B_n^{2+\delta}} \sum_{k=1}^n E\{|X_k - \mu_k|^{2+\delta}\} \rightarrow 0,$$



则随机变量之和的标准化变量

$$Z_n = \frac{\sum_{k=1}^n X_k - E\left(\sum_{k=1}^n X_k\right)}{\sqrt{D\left(\sum_{k=1}^n X_k\right)}} = \frac{\sum_{k=1}^n X_k - \sum_{k=1}^n \mu_k}{B_n}$$

的分布函数 $F_n(x)$ 对于任意 x 满足

$$\lim_{n \rightarrow \infty} F_n(x) = \lim_{n \rightarrow \infty} P \left\{ \frac{\sum_{k=1}^n X_k - \sum_{k=1}^n \mu_k}{B_n} \leq x \right\}$$

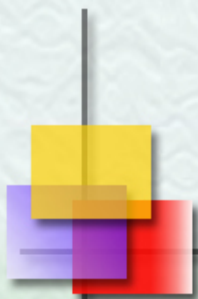
$$= \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = \Phi(x).$$



德莫佛—拉普拉斯定理

设随机变量 η_n ($n = 1, 2, \dots$) 服从参数为 n, p ($0 < p < 1$) 的二项分布, 则对于任意 x , 恒有

$$\lim_{n \rightarrow \infty} P\left\{\frac{\eta_n - np}{\sqrt{np(1-p)}} \leq x\right\} = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt.$$



三、典型例题

例1 假设 X_1, X_2, \dots, X_n 是来自总体 X 的简单随机样本, 已知 $E(X^k) = \alpha_k (k = 1, 2, 3, 4)$. 证明: 当 n 充分大时, 随机变量 $Z_n = \frac{1}{n} \sum_{i=1}^n X_i^2$ 近似服从正态分布, 并指出其分布参数.

解 因为 X_1, X_2, \dots, X_n 独立同分布,
所以 $X_1^2, X_2^2, \dots, X_n^2$ 也独立同分布,



且 $E(X_i^2) = \alpha_2$,

$$D(X_i^2) = E(X_i^4) - (EX_i^2)^2 = \alpha_4 - \alpha_2^2,$$

根据**独立同分布的中心极限定理**知

$$V_n = \frac{\sum_{i=1}^n X_i^2 - n\alpha_2}{\sqrt{n(\alpha_4 - \alpha_2^2)}} = \frac{\frac{1}{n} \sum_{i=1}^n X_i^2 - \alpha_2}{\sqrt{\frac{1}{n}(\alpha_4 - \alpha_2^2)}} = \frac{Z_n - \alpha_2}{\sqrt{\frac{1}{n}(\alpha_4 - \alpha_2^2)}}$$

的极限分布是标准正态分布.



故当 n 充分大时,

V_n 近似服从标准正态分布,

从而当 n 充分大时,

$$Z_n = \sqrt{\frac{1}{n}(\alpha_4 - \alpha_2^2)} V_n + \alpha_2 \text{ 近似服从}$$

参数为 $\mu = \alpha_2$, $\sigma^2 = \frac{\alpha_4 - \alpha_2^2}{n}$ 的正态分布.



例2 现有一批种子, 其中良种占 $\frac{1}{6}$, 今在其中任选 6000 粒, 试问在这些种子中良种所占的比例与 $\frac{1}{6}$ 之差的绝对值小于 $\frac{1}{1000}$ 的概率是多少?

解 令 $X_i = \begin{cases} 1, & \text{第 } i \text{ 粒是良种,} \\ 0, & \text{第 } i \text{ 粒不是良种,} \end{cases} \quad i = 1, 2, \dots, n.$

则 $P(X_i = 1) = \frac{1}{6}$, 记 $Y_n = \sum_{i=1}^n X_i$,

则 $Y_n \sim B\left(n, \frac{1}{6}\right), \quad n = 6000.$



根据题意, 所求概率为

$$P\left(\left|\frac{Y_n}{6000} - \frac{1}{6}\right| \leq \frac{1}{1000}\right) = P(|Y_n - 1000| \leq 6),$$

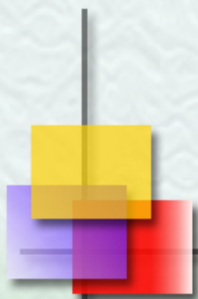
因为 $Y_n \sim B\left(6000, \frac{1}{6}\right)$,

由**中心极限定理**有:

Y_n 近似服从 $N\left(1000, 1000 \times \frac{5}{6}\right)$,



$$\begin{aligned}
 & \text{所以 } P\left(\left|\frac{Y_n}{6000} - \frac{1}{6}\right| \leq \frac{1}{1000}\right) \\
 &= P\left(\left|\frac{Y_n - 1000}{\sqrt{1000 \times 5/6}}\right| \leq \frac{6}{\sqrt{1000 \times 5/6}}\right) \\
 &= 2\Phi\left(\frac{6\sqrt{6}}{\sqrt{5000}}\right) - 1 = 2\Phi(0.208) - 1 \\
 &= 2 \times 0.5832 - 1 = 0.1664.
 \end{aligned}$$



例3 假设测量的随机误差 $X \sim N(0, 10^2)$, 试求在 100 次独立重复测量中, 至少有 3 次测量误差的绝对值大于 19.6 的概率 α , 并利用泊松分布求出 α 的近似值 (要求小数点后取两位有效数字, $\Phi(1.96) = 0.975$).

解 设 p 为每次测量误差的绝对值大于 19.6 的概率,

$$p = P\{|X| > 19.6\} = P\left\{\frac{|X|}{10} > \frac{19.6}{10}\right\}$$



$$\begin{aligned}
 &= P\left\{\frac{|X|}{10} > 1.96\right\} = 2 \cdot P\left\{\frac{X}{10} > 1.96\right\} \\
 &= 2 - 2\Phi(1.96) = 0.05,
 \end{aligned}$$

设 k 为 100 次独立测量中事件 $\{|X| > 19.6\}$ 出现的次数,

则 k 服从参数为 $n = 100, p = 0.05$ 的二项分布,

故 $\alpha = P\{k \geq 3\} = 1 - P\{k < 3\}$

$$\begin{aligned}
 &= 1 - 0.95^{100} - 100 \times 0.95^{99} \times 0.05 \\
 &\quad - \frac{100 \times 99}{2} \times 0.95^{98} \times 0.05^2
 \end{aligned}$$



由泊松定理知,

k 近似服从参数 $\lambda = np = 100 \times 0.05$ 的泊松分布,

$$\text{故 } \alpha = P\{k \geq 3\} = 1 - P\{k < 3\}$$

$$\approx 1 - \frac{5^0 e^{-5}}{0!} - \frac{5e^{-5}}{1!} - \frac{5^2 e^{-5}}{2!}$$

$$= 1 - e^{-5} \left(1 + 5 + \frac{25}{2} \right) \approx 0.87 .$$

