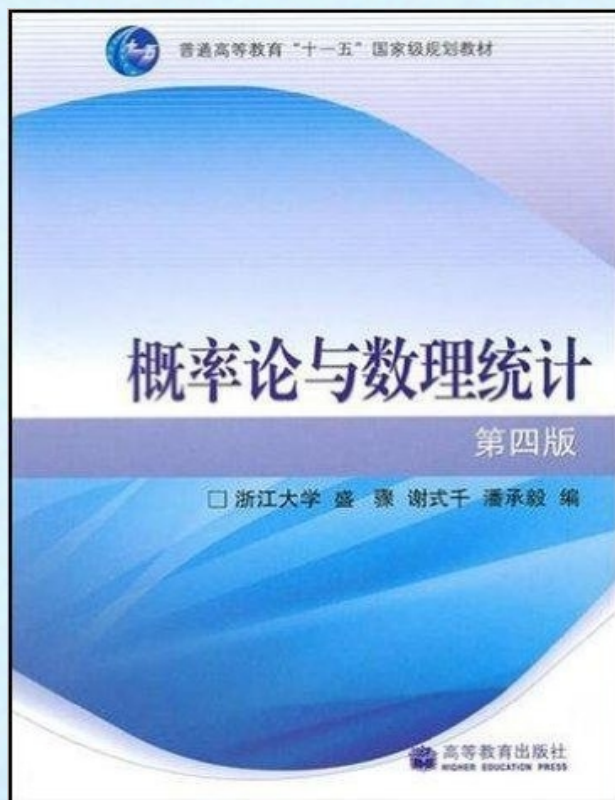


目 录



- 第一章 概率论的基本概念
- 第二章 随机变量及其分布
- 第三章 多维随机变量及其分布
- 第四章 随机变量的数字特征
- 第五章 大数定律及中心极限定理
- 第六章 样本及抽样分布
- 第七章 参数估计
- 第八章 假设检验
- 第九章 方差分析及回归分析

第五章 大数定律及中心极限定理

第一节 大数定律

第二节 中心极限定理

第二节 中心极限定理

(一) 独立同分布的中心极限定理

一、问题的引入

二、基本定理

三、典型例题

四、小结与思考

一、问题的引入

概率论与数理统计是研究随机现象统计规律性的学科。而随机现象的统计规律性是在相同条件下进行大量重复试验呈现出来的. 那么大量随机变量和的分布会呈现什么规律呢？

例如，为了考察某种产品的生产情况（关系的是重量），对产品的重量进行了大量的观测，从而推断的出整体的生产情况。同时从产量、存储、运输的角度，我们往往会关心一批产品重量总和的分布规律。

又例如：**考察射击命中点与靶心距离的偏差.**



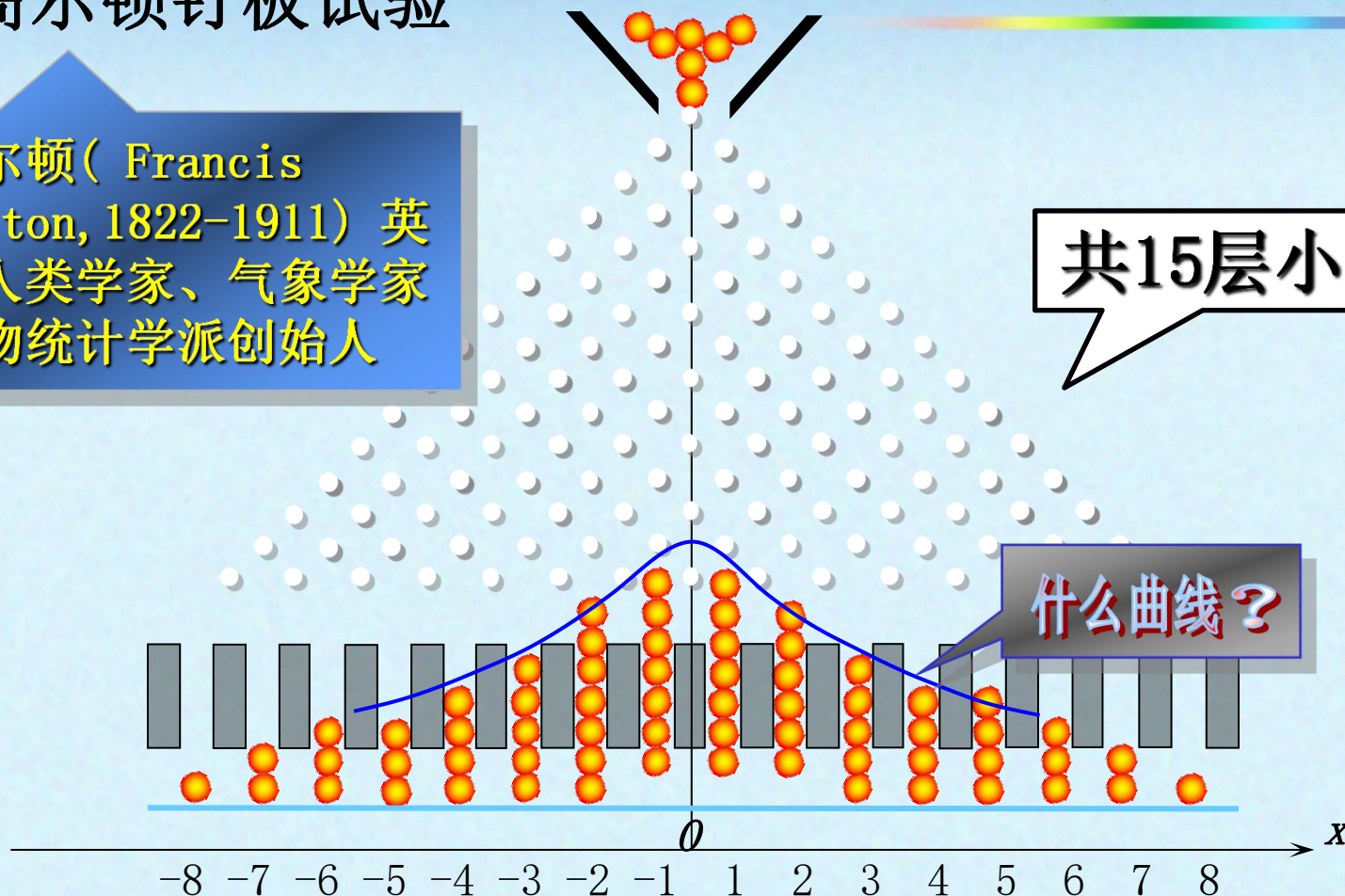
这种偏差是大量微小的偶然因素造成的微小误差的总和，这些因素包括：瞄准误差、测量误差、子弹制造过程方面(如外形、重量等)的误差以及射击时武器的振动、气象因素(如风速、风向、能见度、温度等)的作用，所有这些不同因素所引起的微小误差是相互独立的，并且它们中每一个对总和产生的影响不大。

另外，在第2章，我们提到在自然现象和社会现象中，很多随机变量都是服从或近似服从正态分布的。

高尔顿钉板试验

5.2 中心极限定理

高尔顿 (Francis Galton, 1822-1911) 英国人类学家、气象学家、生物统计学派创始人



共15层小钉

什么曲线?

记 $X_k = \begin{cases} 1, & \text{小球碰第 } k \text{ 层钉后向右落下} \\ -1, & \text{小球碰第 } k \text{ 层钉后向左落下} \end{cases} \quad (k=1, 2, \dots, n)$

$n=15$, $\mu=0$, $\sigma^2=1$, $n\sigma^2=15$, 15个两点分布和

$$\sum_{k=1}^{15} X_k \sim$$

二、基本定理

定理一（独立同分布的中心极限定理 Lindberg-levi）

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

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

的分布函数 $F_n(x)$ 对于任意 x 满足

$$\begin{aligned}\lim_{n \rightarrow \infty} F_n(x) &= \lim_{n \rightarrow \infty} P \left\{ \frac{\sum_{k=1}^n X_k - n\mu}{\sqrt{n} \sigma} \leq x \right\} \\ &= \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = \Phi(x).\end{aligned}$$

定理一表明:

当 $n \rightarrow \infty$, 随机变量序列 Y_n 的分布函数收敛于标准正态分布的分布函数.

□ 定理表明,当 n 充分大时, Y_n 近似服从标准正态分布.

$$Y_n = \frac{\sum_{i=1}^n X_i - nm}{\sqrt{ns}} \stackrel{\text{近似}}{\sim} N(0,1)$$

$$\sum_{i=1}^n X_i \stackrel{\text{近似}}{\sim} N(nm, ns^2)$$

$$\frac{1}{n} \sum_{i=1}^n X_i \stackrel{\text{近似}}{\sim} N\left(m, \frac{s^2}{n}\right)$$

三、典型例题

例1 一加法器同时收到 20 个噪声电压 V_k (单位:V), ($k = 1, 2, \dots, 20$), 设它们是相互独立的随机变量, 且都在区间 $(0, 10)$ 上服从均匀分布, 记 $V = \sum_{k=1}^{20} V_k$, 求 $P\{V > 105\}$ 的近似值.

解 易知 $E(V_k) = 5$, $D(V_k) = \frac{100}{12}$ ($k = 1, 2, \dots, 20$).

由定理一, 随机变量

$$Z = \frac{\sum_{k=1}^{20} V_k - 20 \times 5}{\sqrt{100/12} \sqrt{20}} = \frac{V - 20 \times 5}{\sqrt{100/12} \sqrt{20}}$$

近似服从正态分布 $N(0, 1)$, 于是

$$\begin{aligned} & P\{V > 105\} \\ &= P\left\{\frac{V - 20 \times 5}{(10/\sqrt{12})\sqrt{20}} > \frac{105 - 20 \times 5}{(10/\sqrt{12})\sqrt{20}}\right\} \\ &= P\left\{\frac{V - 100}{(10/\sqrt{12})\sqrt{20}} > 0.387\right\} \end{aligned}$$

$$= 1 - P\left\{\frac{V - 100}{(10 / \sqrt{12})\sqrt{20}} \leq 0.387\right\}$$

$$\approx 1 - \int_{-\infty}^{0.387} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

$$= 1 - \Phi(0.387)$$

$$= 0.348.$$

即有

$$P\{V > 105\} \approx 0.348 .$$

所以20个噪声电压和超过105V的概率约为0.348

例2 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、1名家长、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数相互独立, 且服从同一分布.

- (1) 求参加会议的家长数 X 超过450的概率;
- (2) 求有1名家长来参加会议的学生数不多于340的概率.

解 (1) 以 X_k ($k = 1, 2, \dots, 400$) 记第 k 个学生来参加会议的家长数,

则 X_k 的分布律为

X_k	0	1	2
p_k	0.05	0.8	0.15

易知 $E(X_k) = 1.1$, $D(X_k) = 0.19$, ($k = 1, 2, \dots, 400$)

而 $X = \sum_{k=1}^{400} X_k$, 由定理一, 随机变量

$$\frac{\sum_{k=1}^{400} X_k - 400 \times 1.1}{\sqrt{400} \sqrt{0.19}} = \frac{X - 400 \times 1.1}{\sqrt{400} \sqrt{0.19}}$$

近似服从正态分布 $N(0, 1)$, 于是

$$P\{X > 450\}$$

$$= 1 - P\{X \leq 450\}$$

$$= 1 - P\left\{\frac{X - 400' 1.1}{\sqrt{400} \sqrt{0.19}} \leq \frac{450 - 400' 1.1}{\sqrt{400} \sqrt{0.19}}\right\}$$

$$\gg 1 - \Phi(1.147)$$

$$= 0.1251.$$

(2) 令 $Y_j = \begin{cases} 1, & \text{第}j\text{个学生刚好有一名家长参加会议,} \\ 0, & \text{第}j\text{个学生的家长参加会议人数不为1.} \end{cases} j = 1, 2, \dots, 400$

$Y = \sum_{j=1}^{400} Y_j$ 表示400名学生中刚好有一名家长参加会议的人数

由定理一得, $P\{Y \leq 340\}$

$$= P\left\{ \frac{Y - 400 \times 0.8}{\sqrt{400 \times 0.8 \times 0.2}} \leq \frac{340 - 400 \times 0.8}{\sqrt{400 \times 0.8 \times 0.2}} \right\}$$

$$= P\left\{ \frac{Y - 400 \times 0.8}{\sqrt{400 \times 0.8 \times 0.2}} \leq 2.5 \right\}$$

$$\approx \Phi(2.5)$$

$$= 0.9938 .$$

四、小结与思考

- 1.回答了一定条件下独立随机变量和的分布问题;
- 2.解释了正态分布的形成机制, 突出了正态分布在各分布中的重要地位;
- 3.利用中心极限定理求解问题的步骤;

Q1:在独立同分布的中心极限定理中, 当服从0-1分布时, 会有什么样的结论?

Q2:当独立同分布中心极限定理的条件不满足时, 例如当各 X_i 独立但不同分布时, 有类似的结论吗?

谢 谢！