

《概率论与数理统计》课程教案

主讲教师_____ 所在单位_____

授课班级_____ 专业_____ 撰写时间_____

教案编号	14-0101	教案内容	第五章：大数定律及中心极限定理	学时	2
教学目标	基本要求	(1) 理解切比雪夫不等式的作用，能利用该不等式进行概率估算； (2) 理解大数定律的概念，理解伯努利大数定律； (3) 理解中心极限定理的概念，学会利用独立同分布的中心极限定理（莱维—林德伯格中心极限定理）及棣莫弗—拉普拉斯中心极限定理求解实际问题。			
	能力要求	1. 培养能力要求： a) 掌握概率论和数理统计中的基本概念和性质并能够运用到复杂工程问题的适当表述之中； b) 能够针对工程应用系统或过程的特点选择合适的概率分布来描述随机现象的统计规律性； 2. 这两种能力的训练体现在：中心极限定理在工程实践中的应用，主要是关于大量微小随机因素综合作用情况下的相关随机现象处理问题。			
教学重点	独立同分布的中心极限定理、棣莫佛—拉普拉斯中心极限定理的概念内涵及应用。				
教学难点	依概率收敛、依分布收敛的概念；切比雪夫大数定律、贝努利大数定律、辛钦大数定律的数学表述与内涵。				
教学方法	提问、讨论、讲授。				
工具仪器	多媒体教具、授课计划、教材、教案、教学课件、考勤表、平时成绩登记表				
教学安排	考勤、复习相关知识点、新课内容概述、组织教学、布置作业、课后小结				
教学过程	教学组织、具体教学内容及教学方法、手段、时间分配及其它说明				备注
旧知识点复习和新课内容概述(5-10 分钟)					
第一部分 知识点：绪论，所需时间：10 分钟					
概率论的起源、发展、应用					
(一) 概率论的起源					
早期的埃及人为了忘记饥饿，经常聚集在一起玩一种叫做“猎犬与胡狼”的游戏，实际上就是今天的 掷骰子游戏 ，相对面的数字之和是7 的骰子大约产生于公元					

前1400年的埃及，骰子就是游戏中常用的随机发生器，这类游戏也叫做机会性游戏。17世纪中叶，人们开始对机会性游戏的数学规律进行探讨。它的发展与数学史上一些伟大的名字相联系，如**帕斯卡、费马、惠更斯、詹姆斯、伯努利、棣莫弗、拉普拉斯**等。

1654年，**费马与帕斯卡的通信中关于分赌注问题的讨论被公认为是概率论诞生的标志**。问题是这样的：“两个赌徒相约赌若干局，谁先赢 s 局就算赢了，当赌徒A赢 a 局($a < s$)，而赌徒B赢 b 局($b < s$)时，赌博被迫中止，应该怎样分配赌注才合理？”在三年后，惠根斯亦用自己的方法解决了这一问题，并写成了《论赌博中的计算》一书，这就是概率论最早的论著，他们三人提出的解法中，都首先涉及了数学期望（mathematical expectation）这一概念，并由此奠定了古典概率的基础。

（二）概率论的发展

1713年，伯努利的遗嘱中发表了概率论中的第一个极限定理——**[伯努利大数定理]**，即“**在多次重复独立的试验中，事件发生的频率有越来越稳定的趋势。**”这正是频率稳定性的定理形式。到了1730年，**法国数学家棣莫弗**出版的著作《分析杂论》中包含了著名的「**棣莫弗—拉普拉斯定理**」。这就是概率论中第二个基本极限定理的雏形。接着拉普拉斯在1812年出版的《概率的分析理论》中，首先明确地对概率作了古典的定义。另外，他又和数个数学家建立了关于「正态分布」及「最小二乘法」的理论。后来，泊松将伯努利大数定律做了推广，研究得出了一种新的分布，就是**泊松分布**。概率论发展到1901年，**中心极限定理被严格的证明了**，数学家们利用这一定理第一次科学地**解释了为什么实际中遇到的许多随机变量近似服从以正态分布**。后期的中心研究课题则集中在推广和改进伯努利大数定律及中心极限定理，比如柯尔莫戈洛夫的概率公理化结构、以几乎处处收敛定义的强大数定律、林德伯格-费勒中心极限定理等等，另一方面，一些数学家将兴趣逐渐转向研究随机现象随时间演变过程的规律性，衍生出另一门重要的学科——**随机过程**。

（三）概率论的应用

目前，以概率理论作为基础的学科有很多，而最典型的莫过于统计学。各大高校非数学系本科生使用的概率统计教材都是建立在随机变量基础上的理论，很少有非数学类的学习以测度论为基础的概率理论。通过引入“随机变量”的定义，可以将抽象的样本空间映射到实空间中，方便我们能较好的用数学方法处理任何数据格式（比如实数数据和名义数据等）。

概率论中另一个重要的定义则是“**条件数学期望**”，让人们在做推断的时候想到了利用经验信息（先验信息），由此发展出来的**贝叶斯思想**（贝叶斯统计）现在可以用到任何领域。独立同分布场合的大数定律（**辛钦大数定律**）为一类参数估计奠定了理论基础，因为在简单随机抽样下得到的样本正好是独立同分布的，按照“**样本矩依概率收敛到总体矩**”的思想，**矩估计方法**诞生了。这正是为什么我们用样本

均值去估计总体期望的原因，它也启发人们用概率论的想法构造模型从而实现数值计算，比如**蒙特卡洛方法**。此外，参数估计中最著名的极大似然估计方法（MLE）则是来源于对已今发生的随机事件的概率的假定，人们承认一次观察中出现的那些样本就是最有可能出现的样本，极大它的概率得到了参数的估计，MLE 是目前十分流行的参数估计方法。反过来，利用“**小概率事件在一次试验中实际不发生**”的原理，人们实现了**假设检验、方差分析、相关分析、卡方检验、秩和检验等都是基本的假设检验方法**。

中心极限定理则解释了为什么正态分布在统计中占有不可替代的地位，也告诉我们现实当中什么样的数据可以认为是正态的。自从高斯认为误差服从正态分布以后，到今天，在正态总体下建立的许多估计方法和检验方法非常成熟，例如回归分析、判别分析、因子分析等等。同时，在非正态总体下，许多参数估计和检验也是稳健的，基于样本均值渐近无分布的参数方法的理论基础正是中心极限定理。但是，没有参数方法适用于处理名义变量或次序数据，因此而发展起来的**就是非参数统计**，典型的方法如：列联表、秩检验、核密度估计、局部多项式等等。介于二者之间则是半参数统计了。（更多精彩统计学相关文章，请访问“统计之都” Capital of Statistics—<http://www.cos.name>）

第二部分 知识点：大数定律，所需时间：30 分钟

§1 大数定律

本章研究内容主要考察大量试验后，事件发生的频率和算术平均值的稳定性问题，在第一章和第四章我们分别给出了一些结果

即随着实验次数的增加：

事件 A 发生的**频率** n_A/n 具有稳定性——事件 A 的**概率**

大量测量值的**算术平均值**也具有稳定性——测量值的**数学期望**

本章将对以上问题从理论上进行探讨，并给出证明

定理一（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值

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

则对于任意正数 ε ，有

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

定理的解释: 当 $n \rightarrow \infty$ 时, $\{|\frac{1}{n} \sum_{k=1}^n X_k - \mu| < \varepsilon\}$ 这一随机事件的概率趋近于 1

即对任意的正数 ε , 当 n 充分大时, 不等式 $|\bar{X} - \mu| < \varepsilon$ 成立的概率很大

证明: 由随机变量 $X_1, X_2, \dots, X_n, \dots$ 相互独立, 且具有相同的数学期望和方差

$$\text{有 } E[\bar{X}] = E[\frac{1}{n} \sum_{k=1}^n X_k] = \frac{1}{n} \sum_{k=1}^n E(X_k) = \frac{1}{n} \times n\mu = \mu$$

$$\text{及 } D[\bar{X}] = D[\frac{1}{n} \sum_{k=1}^n X_k] = \frac{1}{n^2} \sum_{k=1}^n D(X_k) = \frac{1}{n^2} \times n\sigma^2 = \frac{\sigma^2}{n},$$

由契比雪夫不等式

$$P\{|\frac{1}{n} \sum_{k=1}^n X_k - \mu| < \varepsilon\} \geq 1 - \frac{\sigma^2/n}{\varepsilon^2}$$

上式中令 $n \rightarrow \infty$ 时, 并注意到概率不能大于 1, 即得

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

定理的实际含义: 令 ε 充分小, 当 n 充分大时, 随机变量 $X_1, X_2, \dots, X_n$ 的算术平均值

$\bar{X} = \frac{1}{n} \sum_{k=1}^n X_k$ 无限接近数学期望 $E(X_1) = E(X_2) = \dots = E(X_n)$, 即当 n 无限增加时 n 个随机

变量的算术平均值将几乎变成一个常数。这种接近是概率意义上的接近

依概率收敛:

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

$$\lim_{n \rightarrow \infty} P\{|Y_n - a| < \varepsilon\} = 1,$$

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

$$Y_n \xrightarrow{P} a$$

依概率收敛序列的性质:

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

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

这样定理一可如下描述

定理一 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 ,

$$\text{即 } \bar{X} \xrightarrow{P} \mu$$

定理二 (伯努利大数定理) 关于频率与概率

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, 则对于任意的 $\varepsilon > 0$, 有

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

证: 参见 125 页例 6 的变量分解

由定理的条件易知 $n_A \sim b(n, p)$, 引入随机变量 $X_i, i=1, \dots, n$

$$\text{其中, } X_i = \begin{cases} 1, & \text{在第 } i \text{ 次实验中事件 } A \text{ 发生} \\ 0, & \text{在第 } i \text{ 次实验中事件 } A \text{ 不发生} \end{cases},$$

于是有 $n_A = X_1 + X_2 + \dots + X_n$

其中 $X_1, X_2, \dots, X_n$ 相互独立, 且都服从以 p 为参数的 $(0-1)$ 分布

X_i	0	1
p_k	$1-p$	p

易知 $E(X_k) = p, D(X_k) = p(1-p), k=1, 2, \dots, n$

$$\text{又 } \frac{n_A}{n} = \frac{1}{n} \sum_{k=1}^n X_k$$

由定理一知: $\lim_{n \rightarrow \infty} P\left\{\left|\frac{1}{n} \sum_{k=1}^n X_k - p\right| < \varepsilon\right\} = 1$

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

定理的含义: 伯努利大数定理表明, 事件发生的频率 $\frac{n_A}{n}$ 依概率收敛于事件的概率 p , 以严格的数学形式表达了频率的稳定性, 和概率的合理性

近似: 当 n 很大时, 事件发生的频率 $\frac{n_A}{n}$ 与概率有较大偏差的可能性很小, 因此由实际推断原理, 由于小概率事件几乎不发生, 当试验次数很大时, 可以用事件的频率来代替事件的概率, 因为产生较大偏差是小概率事件。

在定理一中要求随机变量 $X_1, X_2, \dots$ 的方差存在, 但当随机变量同分布时, 并不需要这一要求:

定理三 (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

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

显然伯努利大数定理是辛钦定理的特殊情况

三个定理之间的比较

	处理问题	定理	内容	定理条件	区别
大数定律	频率和算术平均值的稳定性	一, 契比雪夫定理的特殊情况, 算术平均值的稳定性, 期望	当 $n \rightarrow \infty$ 时, 算术均值依概率收敛于数学期望	$X_1, X_2, \dots, X_n, \dots$ 相互独立, 有相同的数学期望和方差	不必同分布
		二, 贝努利大数定理, 频率稳定性, 概率	当 $n \rightarrow \infty$ 时, 频率 n_A/n 依概率收敛于概率 p	n 次独立重复试验, $X_1, X_2, \dots, X_n$ 相互独立, 服从同一(0-1)分布, p 为事件 A 在每次试验中的概率	是定理三的特例
		三, 辛钦定理, 算术均值稳定性, 数学期望	同定理一	$X_1, X_2, \dots, X_n, \dots$ 相互独立, 服从同一分布, 有相同期望	方差不一定要存在, 但要同分布

第三部分 知识点: 中心极限定理, 所需时间: 45 分钟

§2 中心极限定理

在客观实际中有许多随机变量是由大量的相互独立的随机因素的综合影响所形成的, 而其中每一个因素在总的影响中所起的作用都是微小的。这种随机变量往往近似地服从正态分布。一>中心极限定理

例如: 一个城市一天的总用水量, 每一个用户的用水量是一个随机因素, 且每个用户的用水量的分布是千差万别的, 尽管如此, 在这些大量的微小的随机因素的综合作用下, 总用水量近似服从正态分布

定理四 (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 $\sum_{k=1}^n X_k$ 的标准化变量:

$$Y_n = \frac{\sum_{k=1}^n X_k - E(\sum_{k=1}^n X_k)}{\sqrt{D(\sum_{k=1}^n X_k)}} = \frac{\sum_{k=1}^n X_k - n\mu}{\sqrt{n\sigma^2}}$$

的分布函数 $F_n(x)$ 对于任意 x 满足

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

定理含义：这就是说，均值为 μ ，方差为 $\sigma^2 > 0$ 的独立同分布的随机变量 $X_1, X_2, \dots, X_n$ 之和 $\sum_{k=1}^n X_k$ 的标准化变量，当 n 充分大时，有

$$\frac{\sum_{k=1}^n X_k - n\mu}{\sqrt{n\sigma^2}} \sim \text{近似 } N(0,1)$$

这样可以用(标准)正态分布来对 $\sum_{k=1}^n X_k$ 作理论分析或实际计算，不必求分布函数

将上式改写为 $\frac{\frac{1}{n} \sum_{k=1}^n X_k - \mu}{\sigma/\sqrt{n}} = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$ ，这样有 $\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim \text{近似 } N(0,1)$ ，或 $\bar{X} \sim$

近似 $N(\mu, \sigma^2/n)$

这是独立同分布中心极限定理的结果的另一个形式，这就是说，均值为 μ ，方差为 $\sigma^2 > 0$ 的独立同分布的随机变量 $X_1, X_2, \dots, X_n$ 的算术平均 $\bar{X} = \frac{1}{n} \sum_{k=1}^n X_k$ ，当 n 充分大时，

近似的服从均值为 μ ，方差为 σ^2/n 的正态分布，这是数理统计中大样本统计的推断基础

※定理的证明参看高等概率论的相关内容

定理五 (李雅普诺夫 Liapunov 定理)

设随机变量 $X_1, X_2, \dots, X_n, \dots$ 相互独立，且具有数学期望和方差： $E(X_k) = \mu_k$ ， $D(X_k) = \sigma_k^2 > 0 (k=1, 2, \dots)$ ，

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 $\sum_{k=1}^n X_k$ 的标准化变量：

$$Z_n = \frac{\sum_{k=1}^n X_k - E(\sum_{k=1}^n X_k)}{\sqrt{D(\sum_{k=1}^n X_k)}} = \frac{\sum_{k=1}^n X_k - \sum_{k=1}^n \mu_k}{B_n} \text{ 的分布函数 } F_n(x) \text{ 对于任意 } x \text{ 满足}$$

$$\lim_{n \rightarrow \infty} F_n(x) = \lim_{n \rightarrow \infty} P \left\{ \frac{\sum_{k=1}^n X_k - \sum_{k=1}^n \mu_k}{B_n} \leq x \right\} = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt = \Phi(x)$$

定理五含义：在定理条件下，随机变量 $Z_n = \frac{\sum_{k=1}^n X_k - \sum_{k=1}^n \mu_k}{B_n}$ 当 n 很大时近似的服从正

态分布 $N(0,1)$ ，也即当 n 很大时，随机变量之和 $\sum_{k=1}^n X_k = B_n Z_n + \sum_{k=1}^n \mu_k$ 近似的服从正

态分布 $N(\sum_{k=1}^n \mu_k, B_n^2)$

即无论 $X_k(k=1,2,\dots)$ 服从什么分布，只要满足定理条件，当 n 很大时都有 $\sum_{k=1}^n X_k$ 近似

服从正态分布。这一结论均有普遍意义和广泛的实用价值

如：城市的总用电量，测量误差等等

定理六 (棣莫弗—拉普拉斯定理 De Moivre—Laplace)

设随机变量 $\eta_n \sim b(n,p)$ ，这里 $n=1,2,\dots$ ， $0 < p < 1$ ，则对于任意的 x 有

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

证：定理 6 是定理 4 的推论

由于二项分布随机变量 η_n 可以分解为 n 个相互独立的服从同一 $(0-1)$ 分布的随机变量 $X_1, X_2, \dots, X_n$ 之和，即有 $\eta_n = \sum_{k=1}^n X_k$

其中 $X_k(k=1,2,\dots)$ 的分布律为

X_i	0	1
p_k	$1-p$	p

易知 $E(X_k) = p$ ， $D(X_k) = p(1-p)$ ， $k=1,2,\dots,n$

由定理四

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

该定理表明：正态分布是二项分布的极限分布

中心极限的三个定理的比较

处理问题	定理	内容	定理条件	区别
中心极限定理	一，独立同分布	$X = \sum_{k=1}^n X_k$ $X \text{ 的标准化变量, 当 } n \rightarrow \infty \text{ 时, 近似的有}$ $\frac{\sum_{k=1}^n X_k - n\mu}{\sqrt{n}\sigma} \sim N(0,1)$	$X_1, X_2, \dots, X_n, \dots$ 相互独立, 同分布 数学期望和方差存在	
	二，李雅普诺夫定理	$X = \sum_{k=1}^n X_k$ $X \text{ 的标准化变量, 当 } n \rightarrow \infty \text{ 时, 近似的有}$ $\frac{\sum_{k=1}^n X_k - \sum_{k=1}^n \mu_k}{B_n} \sim N(0,1)$	$X_1, X_2, \dots, X_n, \dots$ 相互独立, 数学期望和方差存在; 存在正数 δ , 使得当 $n \rightarrow \infty$ 时, $\frac{1}{B_n^{2+\delta}} \sum_{k=1}^n E\{ X_k - \mu_k ^{2+\delta}\} \rightarrow 0$	不必同分布, 有特殊条件具有普遍性
	三，棣莫弗—拉普拉斯定理	$\eta_n \sim b(n, p)$ 当 $n \rightarrow \infty$ 时, 近似的有 $\eta_n \sim N(np, np(1-p))$ $\frac{\eta_n - np}{\sqrt{np(1-p)}} \sim N(0,1)$	$\eta_n = \sum_{k=1}^n X_k$ $X_1, X_2, \dots, X_n$ 相互独立且服从同一(0-1)分布	二项分布的极限分布定理四的推论

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

解: 直接求 V 的分布函数是一个 20 维的联合分布, 概率不易计算

由于所求为 20 个独立同分布的变量之和, 易知

$$E(V_k) = (0+10)/2 = 5, \quad D(V_k) = (10-0)^2/12 = 25/3$$

所以由定理四, 独立同分布的中心极限定理有

$$Z = (V - 20 \times 5) / \sqrt{20 \times 25/3} \sim N(0,1)$$

所以 $P\{V > 105\} = P\{(V - 100) / \sqrt{20 \times 25/3} > (105 - 100) / \sqrt{20 \times 25/3}\}$

$$=1-P\{(V-100)/10\sqrt{5/3}\leq 5/10\sqrt{5/3}\}$$

$$=1-\Phi(1/(2\sqrt{5/3}))=0.348$$

例 2: 一船舶在某海区航行, 每受一次波浪的冲击纵摇角大于 3° 的概率为 $p=1/3$, 若受了 90000 次波浪冲击, 问其中有 29500~30500 次纵摇角大于 3° 的概率是多少?

解: 设每一次冲击为一次试验, 且假定各次冲击独立, 在 90000 次冲击中纵摇角大于 3° 的次数记做 X . 则 $X \sim b(90000, 1/3)$ 的二项分布

而所求概率 $P\{29500 \leq X \leq 30500\} = \sum_{k=29500}^{30500} C_{90000}^k \left(\frac{1}{3}\right)^k \left(1 - \frac{1}{3}\right)^{90000-k}$, 此式有多项又涉及指数很

难计算, 由于 n 很大, 可以考虑用定理六, 棣莫弗拉普拉斯定理, 用二项分布的极限分布来计算

$$P\{29500 \leq X \leq 30500\} = P\left\{\frac{29500 - np}{\sqrt{np(1-p)}} \leq \frac{X - np}{\sqrt{np(1-p)}} \leq \frac{30500 - np}{\sqrt{np(1-p)}}\right\}$$

$$\approx \Phi\left(\frac{30500 - np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{29500 - np}{\sqrt{np(1-p)}}\right)$$

$$= \Phi(5\sqrt{2}/2) - \Phi(-5\sqrt{2}/2) = 0.9995$$

例 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、1 名家长和 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, 设个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)求有一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, 则其分布律为

X_i	0	1	2
p_k	0.05	0.8	0.15

易知 $E(X_k) = 1 \times 0.8 + 2 \times 0.15 = 1.1$, $D(X_k) = 0.19$, $k=1, 2, \dots, 400$,

而 $X = \sum_{k=1}^{400} X_k$ 由独立同分布的中心极限定理, 随机变量

$$\frac{\sum_{k=1}^{400} X_k - 400 \times 1.1}{\sqrt{400 \times 0.19}} = \frac{X - 400 \times 1.1}{\sqrt{400 \times 0.19}} \text{ 近似服从正态分布 } N(0, 1), \text{ 于是}$$

$$P\{X > 450\} = P\left\{\frac{X - 400 \times 1.1}{\sqrt{400 \times 0.19}} > \frac{450 - 400 \times 1.1}{\sqrt{400 \times 0.19}}\right\}$$

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

$$=1-\Phi\left(\frac{450-400\times 1.1}{\sqrt{400}\sqrt{0.19}}\right)=0.1257$$

(2) 以 Y 记有一名家长来参加会议的学生数, 则 $Y\sim b(400,0.8)$

由棣莫弗—拉普拉斯定理

$$P\{Y\leq 340\}=P\left\{\frac{Y-400\times 0.8}{\sqrt{400\times 0.8\times (1-0.8)}}\leq\frac{340-400\times 0.8}{\sqrt{400\times 0.8\times (1-0.8)}}\right\}$$

$$=P\left\{\frac{Y-400\times 0.8}{\sqrt{400\times 0.8\times (1-0.8)}}\leq 2.5\right\}$$

$$=\Phi(2.5)=0.9938$$

课后作业

作业: 126 页 1, 3, 5, 10, 11, 13 :数值实验作业、其它要求

课后小结

以下问题要重点讲:

大数定理的意义

中心极限定理所能解决的问题