第五章　大数定理与中心极限定理

2009 考试内容（本大纲为数学 1，数学 3 需要根据大纲作部分增删）

切比雪夫（Chebyshev）不等式　切比雪夫大数定律　伯努利（Bernoulli）大数定律　辛钦（Khinchine）大数定律　棣莫弗—拉普拉斯（De Moivre—Laplace）定理　列维—林德伯格（Levy—Lindberg）定理

考试要求

1. 了解切比雪夫不等式。
2. 了解切比雪夫大数定律、伯努利大数定律和辛钦大数定律（独立同分布随机变量序列的大数定律）
3. 了解棣莫弗—拉普拉斯定理（二项分布以正态分布为极限分布）和列维—林德伯格定理（独立同分布随机变量序列的中心极限定理）

考点导读 3 大 2 中 1 不等（3 个大数定理、2 个中心极限定理和一个不等式）。

一、切贝雪夫不等式

1. 切贝雪夫不等式及其应用范围

如果不知道 X 属于何种分布，只要 $E\left(X\right)$ 和 $D(X)$ 存在，就可以估算出以 $E\left(X\right)$ 为中心的对称区间上取值的概率。即任给 $\varepsilon>0$, 有 $\boxed{P\{|X-E(X)|\geq\varepsilon\}\leq\frac{D(X)}{\varepsilon^2}}$ 或 $\boxed{P\{|X-E(X)|<\varepsilon\}\geq1-\frac{D(X)}{\varepsilon^2}}$

证 明：由积分比较定理可知：

$$DX=\int_{-\infty}^{\infty}\left[x-EX\right]^2f\left(x\right)dx\geq\int\limits_{|x-EX|\geq\varepsilon}\left[x-EX\right]^2f\left(x\right)dx\geq\int\limits_{|x-EX|\geq\varepsilon}\varepsilon^2f\left(x\right)dx$$

$$=\varepsilon^2\int\limits_{|x-EX|\geq\varepsilon}f\left(x\right)dx=\varepsilon^2P\left\{\left|X-EX\right|\geq\varepsilon\right\}$$

$$\Rightarrow P\left\{\left|X-EX\right|\geq\varepsilon\right\}\leq\frac{DX}{\varepsilon^2}$$

$$\Rightarrow1-P\left\{\left|X-EX\right|<\varepsilon\right\}\leq\frac{DX}{\varepsilon^2}\Rightarrow P\left\{\left|X-EX\right|<\varepsilon\right\}\geq1-\frac{DX}{\varepsilon^2}$$

智轩第 8 技 对于任意事件 $a<X<b$，应用切贝雪夫不等式的条件是 $EX=\frac{a+b}{2}$，并有

$$\boxed{P\left\{a<X<b\right\}=P\left\{\frac{a-b}{2}<X-\frac{a+b}{2}<\frac{b-a}{2}\right\}=P\left\{\left|X-\frac{a+b}{2}\right|<\frac{b-a}{2}\right\}\geq1-\frac{4DX}{\left(b-a\right)^2}}$$

2. 依概率收敛的定义

设 a 是一个常数，X_n 为一随机变量序列，$\forall\varepsilon>0,\ \exists P\{|X_n-a|<\varepsilon\}=1$ 或 $P\{|X_n-a|\geq\varepsilon\}=0$，则称 $\{X_n\}$ 依概率收敛于 a，记为 $X_n\xrightarrow{P}a$ 或 $\lim\limits_{n\to+\infty}X_n=a\left(P\right)$ 。

■切贝雪夫不等式题型题法

【例 1】设 X 为连续型随机变量，则是对任意常数 C，必有

（A）$P(|X-C|\geq\varepsilon)=\frac{E|X-C|}{\varepsilon}$　　（B）$P(|X-C|\geq\varepsilon)\geq\frac{E|X-C|}{\varepsilon}$

（C）$P(|X-C|\geq\varepsilon)\leq\frac{E|X-C|}{\varepsilon}$　　（D）$P(|X-C|\geq\varepsilon)\leq\frac{D(X)}{\varepsilon^2}$

解：$P(|X-C|\ge\varepsilon)=\int_{|x-C|\ge\varepsilon}f(x)dx\le\int_{-\infty}^{+\infty}\frac{|x-C|}{\varepsilon}f(x)dx=\frac{1}{\varepsilon}\int_{-\infty}^{+\infty}|x-C|f(x)dx=\frac{E|X-C|}{\varepsilon}$，应选（C)。

【例 2】设 X 的概率密度为 $f(x)=\begin{cases}\dfrac{x^ne^{-x}}{n!}, & x\ge 0\\ 0, & x<0\end{cases}$，证明：$P\{0<X<2(n+1)\}\ge\dfrac{n}{n+1}$。

解：根据 $\boxed{P\{a<X<b\}=P\left\{\frac{a-b}{2}<X-\frac{a+b}{2}<\frac{b-a}{2}\right\}=P\left\{\left|X-\frac{a+b}{2}\right|<\frac{b-a}{2}\right\}\ge1-\frac{4DX}{(b-a)^2}}$

取 $a=0,\ b=2(n+1)$，可以推知 $EX=\dfrac{a+b}{2}=\dfrac{0+2(n+1)}{2}=n+1$，否则，命题就存在错误，事实上

$$EX=\int_{-\infty}^{+\infty}xf(x)dx=\frac{1}{n!}\int_0^{+\infty}x^{n+1}e^{-x}dx=\frac{1}{n!}\Gamma(n+2)=n+1$$

$$EX^2=\int_{-\infty}^{+\infty}x^2f(x)dx=\frac{1}{n!}\int_0^{+\infty}x^{n+2}e^{-x}dx=\frac{1}{n!}\Gamma(n+3)=(n+2)(n+1)$$

$$DX=EX^2-(EX)^2=(n+2)(n+1)-(n+1)^2=n+1$$

$$P\{0<X<2(n+1)\}=P\left\{\left|X-\frac{a+b}{2}\right|<\frac{b-a}{2}\right\}=P\{|X-(n+1)|<n+1\}\ge1-\frac{DX}{(n+1)^2}=\frac{n}{n+1}$$

【例 3】已知随机变量 X,Y 的数学期望分别为 -2 和 2，方差分别为 1 和 4，相关系数为 -0.5，

试估计 $P\{|X+Y|\ge6\}$。

解：由于未知 X,Y 的具体分布，故使用切贝雪夫不等式 $P\{|X-E(X)|\ge\varepsilon\}\le\dfrac{D(X)}{\varepsilon^2}$

$$Z=X+Y\Rightarrow EZ=EX+EY=-2+2=0$$

$$DZ=DX+DY+2\sqrt{DX\cdot DY}\cdot\rho_{XY}=1+4+2\times\sqrt{1\times4}\times(-0.5)=3$$

$$P\{|X-E(X)|\ge\varepsilon\}\le\frac{D(X)}{\varepsilon^2}\Rightarrow P\{|Z-E(Z)|\ge6\}\le\frac{3}{6^2}=\frac{1}{12}$$

$$\Rightarrow P\{|X+Y|\ge6\}\le\frac{1}{12}$$

【例 4】随机掷 6 颗骰子，利用切比雪夫不等式估计 6 颗骰子点数之和大于 14 小于 28 的概率至少为多少？

解：设 $X_i=\{\text{第 } i \text{ 颗骰子出现的点数}\}$

$$X_i\sim\begin{pmatrix}1&2&3&4&5&6\\ \frac{1}{6}&\frac{1}{6}&\frac{1}{6}&\frac{1}{6}&\frac{1}{6}&\frac{1}{6}\end{pmatrix}\Rightarrow X=\sum_{i=1}^{6}X_i$$

$$EX_i=\frac{1}{6}(1+2+3+4+5+6)=\frac{7}{2};\ EX_i^2=\frac{1}{6}(1^2+2^2+3^2+4^2+5^2+6^2)=\frac{91}{6};\ DX_i=EX_i^2-(EX_i)^2=\frac{91}{6}-\left(\frac{7}{2}\right)^2=\frac{35}{12}$$

$$EX=E\left(\sum_{i=1}^{6}X_i\right)=\sum_{i=1}^{6}EX_i=6\times\frac{7}{2}=21;\ DX=D\left(\sum_{i=1}^{6}X_i\right)=\sum_{i=1}^{6}DX_i=6\times\frac{35}{12}=\frac{35}{2}$$

$$P\{14<X<28\}=P\{-7<X-21<7\}=P\{|X-21|<7\}\ge1-\frac{\frac{35}{2}}{7^2}=\frac{9}{14}$$

【例 5】假设某一年龄段女孩平均身高 130cm，标准差是 8 厘米，现在从该年龄段女孩中随机抽取 5 名女孩，测其身高，估计她们的平均身高 $\overline{X}$ 在120cm-140cm之间的概率。

解：不知分布估计概率使用切贝雪夫不等式

设 X_i 为第 i 名被测女孩的身高，显然 $X_1,\cdots X_5$ 相互独立同分布

$$E(X_i)=130;\ D(X_i)=\sigma^2=8^2=64;\ \ \overline{X}=\frac{1}{5}\sum_{i=1}^{5}X_i\Rightarrow E(\overline{X})=\frac{1}{5}\sum_{i=1}^{5}E(X_i)=\frac{1}{5}\times5\times130=130$$

$$D(\overline{X})=\frac{1}{25}D(\Sigma X_i)=\frac{1}{25}\Sigma D(X_i)=\frac{1}{25}\times5\times64=\frac{64}{5}=12.8$$

应用切贝雪夫不等式，有 $P\{120<\overline{X}<140\}=P\{\left|\overline{X}-130\right|<10\}\ge1-\frac{12.8}{10^2}=0.872$ 。

二、3 个大数定理

1．大数定理的定义

设随机变量序列 $\{X_n\}$ 的数学期望 $\{EX_n\}$ 存在（不要求方差一定存在），如果对任意 $\varepsilon>0$，有

$\lim\limits_{n\to\infty}P\left(\left|\overline{X}-\frac{1}{n}\sum_{i=1}^{n}EX_i\right|<\varepsilon\right)=1$，则称 $\{X_n\}$ 服从大数定理。其中 $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$ 。

2．大数定理的特征：体现一个“均”字。

大数定理中的随机变量、数学期望、方差（标准差）均是对“均”而言。如 $\overline{X}$ 和 $\frac{1}{n}\sum_{i=1}^{n}EX_i$

3．　切比雪夫大数定理

设随机变量 $X_1,X_2,\cdots,X_n$ 相互独立，不要求同分布，但**数学期望和方差都存在**（注意 $DX=cn$，当 $n\to+\infty$ 方差不存在），则 $\forall\varepsilon>0,$ 有

$\boxed{\lim\limits_{n\to\infty}P\left\{\left|\overline{X}-\frac{1}{n}\sum_{i=1}^{n}EX_i\right|<\varepsilon\right\}=1}$。

4．辛钦大数定理

设随机变量 $X_1,X_2,\cdots,X_n$ 相互独立同分布，且具有**相同的数学期望** $E(X_i)=\mu\ \left(i=1,2,\cdots,n\right)$，**不要求方差存在**，则 $\forall\varepsilon>0,$ 有 $\boxed{\lim\limits_{n\to\infty}P\{\left|\overline{X}-\mu\right|<\varepsilon\}=1}$

辛钦大数定理说明，在大量（一般要求 $n>45$ ）的测量值中，算术平均当 n 无限增加时将接近期望值。

5．伯努利大数定理

设 Y_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 $\forall\varepsilon>0$，有

$\boxed{\lim\limits_{n\to\infty}P\left\{\left|\frac{Y_A}{n}-p\right|<\varepsilon\right\}=1}$ 或 $\boxed{\lim\limits_{n\to\infty}P\left\{\left|\frac{Y_A}{n}-p\right|\ge\varepsilon\right\}=0}$

评 注　$(a)\ Y_A=X_1+X_2+\ldots+X_n$， $X_1,X_2,\ldots,X_n$ 相互独立，且都服从概率为 p 的同一 $(0-1)$ 分布；当然

$(0-1)$分布数学期望和方差也存在。

(b)当n很大（一般要求大于 45）时，事件发生的频率$\frac{Y_A}{n}$具有稳定性，且逼近于其概率，这也是为什么在实际应用中，常用频率来代替事件发生概率的原因。

(c)显然，它要求的条件最高：不仅期望、方差存在，而且同分布$(0-1)$。

6．三个大数定理的应用选择原则

大数定理提供了算术平均代替加权平均（数学期望）的理论根据，适应于事件发生的平均值依概率收敛情形。如果能已知EX，DX都存在，则使用**切比雪夫大数定理**；如果仅知道EX存在，而未知DX是否存在，则使用**辛钦大数定理**；如果是伯努利试验，则使用**伯努利大数定理**。

■大数定理题型题法

【例 6】$X_i \sim E(2)$，$\{X_i\}$独立同，求$\lim\limits_{n\to+\infty}\frac{1}{n}\sum\limits_{i=1}^{n}X_i^2$。

解：注意随机变量的极限是指依概率收敛情形。本题知道了具体分布，求随机变量**平均值**的极限，故使用大数定理，又能够确定EX, DX，故使用**切比雪夫大数定理**。

$$\lim_{n\to\infty}P\{\left|\frac{1}{n}\sum_{i=1}^{n}X_i-\mu\right|<\varepsilon\}=1\Rightarrow\lim_{n\to\infty}P\{\left|\frac{1}{n}\sum_{i=1}^{n}X_i^2-EX_i^2\right|<\varepsilon\}=1$$

$$EX_i^2=DX_i+\left(EX_i\right)^2=\frac{1}{\lambda^2}+\frac{1}{\lambda^2}=\frac{2}{\lambda^2}=\frac{1}{2}$$

$$\Rightarrow\lim_{n\to\infty}P\{\left|\frac{1}{n}\sum_{i=1}^{n}X_i^2-\frac{1}{2}\right|<\varepsilon\}=1\Rightarrow\lim_{n\to+\infty}\frac{1}{n}\sum_{i=1}^{n}X_i^2=\frac{1}{2}(P).$$

【例 7】设$\{X_n\}$独立同分布，$F(X)=a+\frac{1}{\pi}arctg\frac{x}{b}(b\neq0)$，问辛钦大数定理可否适应。

解：$f(x)=F'(x)=\frac{|b|}{\pi(b^2+x^2)}$，　$EX=\int_{-\infty}^{+\infty}xf(x)dx=\frac{2|b|}{\pi}\int_0^{+\infty}\frac{x}{b^2+x^2}dx=\frac{|b|}{\pi}m(b^2+X^2)\Big|_0^{+\infty}=+\infty$

数学期望不存在，故不可适用辛钦大数定理。

【例 8】设$\{X_n\}$独立同分布，且$EX_n=0$，$n=1,2,\cdots$，求$\lim\limits_{n\to\infty}P\left(\sum\limits_{i=1}^{n}X_i<n\right)$。

解：不知道方差是否存在，使用辛钦大数定理

$$\lim_{n\to\infty}P\left\{\left|\overline{X}-\mu\right|<\varepsilon\right\}=1\Rightarrow\lim_{n\to\infty}P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}X_i\right|<\varepsilon\right\}=1\xrightarrow{let\ \ \varepsilon=1}\lim_{n\to\infty}P\left\{\left|\sum_{i=1}^{n}X_i\right|<n\right\}=1$$

$$\text{而 }1\geq\lim_{n\to\infty}P\left\{\sum_{i=1}^{n}X_i<n\right\}\geq\lim_{n\to\infty}P\left\{\left|\sum_{i=1}^{n}X_i\right|<n\right\}=1\Rightarrow\lim_{n\to\infty}P\left\{\sum_{i=1}^{n}X_i<n\right\}=1$$

【例 9】设$\{X_n\}$相互独立，$X_n\sim E(n)$ $n=1,2,\cdots$。则下列哪个不符合切比雪夫大数定理。

$(A)X_1,\ X_2,\cdots X_n$　　　　$(B)X_1,\ 2^2X_2,\cdots n^2X_n$

$(C)X_1,\ \frac{1}{2}X_2,\cdots\ \frac{1}{n}X_n$　　　　$(D)X_1,\ 2X_2,\cdots nX_n$

解：选(B)。

$(A)EX_n=\frac{1}{n};\ DX_n=\frac{1}{n^2}$,符合。

$(B)E\left(n^2X_n\right)=n^2\cdot\frac{1}{n}=n;\ D\left(n^2X_n\right)=n^4\cdot\frac{1}{n^2}=n^2$无界，即不存在,不符合。

$(C)E\left(\frac{1}{n}X_n\right)=\frac{1}{n}\cdot\frac{1}{n}=\frac{1}{n^2};\ D\left(\frac{1}{n}X_n\right)=\frac{1}{n}\cdot\frac{1}{n^2}=\frac{1}{n^3}$符合。

$(D)E\left(nX_n\right)=n\cdot\frac{1}{n}=1;\ D\left(\frac{1}{n}X_n\right)=n\cdot\frac{1}{n^2}=\frac{1}{n}$符合。

三、2 个中心极限定理

●**中心极限的应用范围：**

$(1)X_1,X_2,\cdots,X_n\ (n>45)$独立同分布；　$(2)EX_n$ 和 DX_n都存在，且 $DX_n\neq 0\ (n=1,\ 2,\cdots)$。

●**中心极限的特征：**体现一个“和”字

中心极限中的随机变量、数学期望、方差（标准差）均是对“和”而言。如$\sum_{k=1}^{n}X_k$ 和$n\mu$ 及$\sqrt{n}\sigma$。

3.1 列维—林德伯格中心极限定理（又称独立同分布的中心极限定理）

设$X_1,X_2,\cdots,X_n\cdots$独立同分布，且具有数学期望和大于零的方差，设$E(X_i)=\mu,\ D(X_i)=\sigma^2\neq 0$，则随机变量$\sum_{i=1}^{n}X_k$ 的标准化量$Y_n=\frac{\sum_{k=1}^{n}X_k-E\left(\sum_{k=1}^{n}X_k\right)}{\sqrt{D\left(\sum_{k=1}^{n}X_k\right)}}=\frac{\sum_{k=1}^{n}X_k-n\mu}{\sqrt{n}\sigma}$的分布函数$F_n(x)$满足

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

评 注 ①$\sum_{k=1}^{n}X_k\sim N(n\mu,\ n\sigma^2)$，　$n\to\infty$；

②此处Y_n表达式中，分子与分母可同乘以$\frac{1}{n}\to Y_n=\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$，正好对应标准化$N(0,\ 1)$。

3.2 棣莫佛—拉普拉斯中心极限定理

设随机变量$\eta_n\left(n=1,2,\ldots\right)$服从参数为$n,\ p$的二项分布（二项分布也是要求$X_1,\cdots X_n$相互独立，$\eta_n=\sum_{k=1}^{n}X_k$，同时隐含$DX_k=\sigma^2=np(1-p)\neq 0$），则$\forall x,$ 随机变量$\eta_n=\sum_{i=1}^{n}X_i$的标准化量$\frac{\eta_n-np}{\sqrt{np(1-p)}}$满足

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

评 注 ①正态分布是二项分布的极限分布；

②$\eta_n=\sum_{k=1}^{n}X_k\sim N\left[np,\ np\left(1-p\right)\right]$，$n\to\infty$。

③ 一般来说，$\boxed{x \ge 3 \Rightarrow \Phi(x) \approx 1;\quad x \le -3 \Rightarrow \Phi(x) \approx 0}$

智轩第 9 技 2 个中心极限定理的应用选择方法

中心极限定理提供了任何备选事件发生的标准化量依概率收敛于 $N(0,\ 1)$ 的理论根据。

当 EX，DX 都存在，且 $DX \ne 0$ 时，如果是伯努利试验（离散型），则使用**莫佛—拉普拉斯中心极限定理**；一般型使用**列维一林德伯格中心极限定理**。

■中心极限定理题型题法

【例 10】设随机变量序列 $\{X_n\}$ 相互独立，$n = 1, 2, \cdots$。则下列哪个条件下，序列 $\{X_n^2\}$ 符合列维一林德伯格中心极限定理。

$(A) P\{X_i = m\} = p^m q^{1-m} \quad m = 0,\ 1$

$(B) P\{X_i \le x\} = \int_{-\infty}^{x} \frac{1}{\pi(1+t^2)} dt$

$(C) P\{|X_i| = m\} = \frac{c}{m^2}$，$m = 0,\ 1, \cdots.\ c = \left(\sum_{i=1}^{n} \frac{2}{m^2}\right)^{-1} = \frac{3}{\pi^2}$

$(D) X_i$ 服从 $E(i)$，$i = 0,\ 1, \cdots$

解：选 (A)。

$(A) X_i \sim B(1,\ p) \Rightarrow X_i^2 \sim B(1,\ p)$,期望和方差都存在，符合。

$(B) P\{X_i \le x\} = \int_{-\infty}^{x} \frac{1}{\pi(1+t^2)} dt$

$$\Rightarrow \int_{-\infty}^{x} \frac{|x|}{\pi(1+x^2)} dx = \frac{1}{\pi} \lim_{a \to -\infty} \int_{-a}^{x} \frac{|t|}{1+t^2} dt = \frac{1}{\pi} \lim_{a \to -\infty} \int_{-a}^{x} \frac{|t|}{1+t^2} dt$$

$$= \frac{1}{\pi} \lim_{a \to -\infty} \left[-\int_{-a}^{0} \frac{t}{1+t^2} dt + \int_{0}^{x} \frac{t}{1+t^2} dt \right]$$

$$= \frac{2}{\pi} \lim_{a \to -\infty} \left[-\ln(1+t^2)\Big|_{-a}^{0} + \ln(1+t^2)\Big|_{0}^{x} \right]$$

$$= \frac{2}{\pi} \lim_{a \to -\infty} \left[\ln(1+a^2) + \ln(1+x^2) \right] \to +\infty \Rightarrow X_i \text{期望不存在, 不符合。}$$

$(C) P\{|X_i| = m\} = \frac{c}{m^2} \quad \Rightarrow \sum_{m=1}^{+\infty} |x| P\{x = \pm m\} = 2\sum_{m=1}^{+\infty} m \cdot \frac{c}{m^2} = 2c\sum_{m=1}^{+\infty} \frac{1}{m}$不收敛，故不符合。

$(D) X_i$ 随 i 的变化而变化，不同分布，故不符合。

【例 11】设随机变量 $X_1,\ X_2, \cdots,\ X_n$ 相互独立，且都服从均匀分布 $U(0,\ 2)$，求 $\sum_{i=1}^{\infty} X_i$ 的分布。

解：$X_1,\ X_2, \cdots,\ X_n$ 独立同分布，且 $EX_i = \frac{a+b}{2} = 1;\ DX_i = \frac{(b-a)}{12} = \frac{1}{3}$，满足列维一林德伯格中心极限定理，则 $\lim_{n \to \infty} P\left\{ \frac{\sum_{i=1}^{n} X_i - n \times 1}{\sqrt{\frac{n}{3}}} \le x \right\} = \Phi(x) \Rightarrow \frac{\sum_{i=1}^{n} X_i - n \times 1}{\sqrt{\frac{n}{3}}}$ 近似服从 $N(0,\ 1) \Rightarrow \sum_{i=1}^{\infty} X_i \sim N\left(n,\ \frac{n}{3}\right)$。

【例 12】 一个供电网内共有 10000 盏功率相同的灯，夜晚每一盏灯开着的概率是 0.7，假设各盏灯开、关彼此独立，求夜晚同时开着的灯数在 6800 到 7200 之间的概率。

解：设 X 表示夜晚同时开着的灯的数目，依题意，X 服从 n=10000，P=0.7 的二项分布。

$$EX = np = 10000\times 0.7 = 7000,\ DX = np(1-p) = 2100$$

由 n 较大，根据棣莫佛—拉普拉斯定理 X 近似服从正态分布 N（7000，2100）

$$P\{6800 < X < 7200\} = P\left\{\left|\frac{x-7000}{\sqrt{2100}}\right| < \frac{200}{\sqrt{2100}}\right\} = \int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt = 2\Phi\left(x\right)-1$$

$$= 2\phi\left(\frac{200}{\sqrt{2100}}\right)-1 = 0.9999 \qquad \text{其中}\ \Phi\left(x\right) = \int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt$$

（题中：$np \to 7000;\ \ \eta_n \to x$）

【例 13】 多次重复观测一个物理量，假设每次测量产生的随机误差都服从正态分布 $N\left(0,\ 0.3^2\right)$，如果取 n 次测量的算术平均值作为测量结果，试计算：

（1）测量结果与真值之差的绝对值小于一个小正数 δ 的概率 P；

（2）给定 $\delta = 0.05,$ 使 P 不小于 0.95，至少应进行的测量次数 n。

解：令随机变量 X_i，ε_i 分别表示第 i 次的测量结果与测量误差，μ 表示真值，所以

$$X_i = \mu + \varepsilon_i，\ i = 1,2,\cdots,n$$

$$\varepsilon_i \sim N(0,\ 0.3^2) \quad \text{易见}\ \varepsilon_i\ \text{相互独立}$$

所以，$\{X_i\}$ 服从相互独立的正态分布 $N(\mu, 0.3^2)$

$$\sum_{i=1}^{n} X_i\ \text{服从}\ \ N(n\mu, 0.09n)$$

$$\overline{X} = \frac{1}{n}\Sigma X_i \quad \text{服从}\ \ N(\mu, \frac{0.09}{n})$$

（1） 根据独立同中心极限定理

$$P\{\left|\overline{X}-\mu\right| < \delta\} = P\left\{\left|\frac{n\overline{X}-n\mu}{0.3\sqrt{n}}\right| = \left|\frac{\Sigma X_i - n\mu}{\sqrt{n}\sigma}\right| < \frac{\delta\sqrt{n}}{0.3}\right\} = 2\phi\left(\frac{\delta\sqrt{n}}{0.3}\right)-1$$

* 注意中心极限定义中积分下限为 $-\infty$ 时，对应 $\Phi(x)$。

（2） $P\left\{\left|\overline{X}-\mu\right| < 0.05\right\} = 2\phi\left(\frac{0.05\sqrt{n}}{0.3}\right)-1 \ge 0.95 \Rightarrow n \ge 138.2976$，故观测次数至少为 139。

【例 14】设随机变量 $\{X_n\}$ 相互独立，$S_n = \sum_{k=1}^{n} X_k$，则根据列维一林德柏格中心极限定理，当 n 充分大 S_n 似服从正态分布，只要 $\{X_n\}$ （　　）

（A）有相同的数学期望　　　　（B）有相同的方差

（C）服从同一指数方布　　　　（D）服从同一离散分布

解：不管是哪一种中心极限定理，其共同的条件是$\{X_n\}$存在数学期望（可以为 0）和方差（不能为 0），上述中只有（C）满足。

【例 15】一生产线生产的产品成箱包装，每箱的重量是随机的，假设每箱平均重量 50kg，标准差为 5kg，若用最大载重量为 5 吨的汽车承运，试用中心极限定理说明每辆车最多可装多少箱，才能保障不超载的概率大于 0.977 $(\phi(2)=0.977)$。

解：设X_i是装运的第 i 的重量，n 表示装运箱数。

$$EX_i=50,\ \ DX_i=5^2=25$$

且装运的总重量$Y=X_1+X_2+\cdots+X_n$，$\{X_n\}$独立同分布

$$EY=50n\quad DY=25n$$

由刘维一林德伯格中心极限定理知　$Y\sim N(50n,\ 25n)$

于是 $P(Y\le 5000)=P\left\{\dfrac{Y-50n}{5\sqrt{n}}\le\dfrac{5000-50n}{5\sqrt{n}}\right\}=\phi\left(\dfrac{1000-10n}{\sqrt{n}}\right)>0.977=\phi(2)$

故 $\dfrac{1000-10n}{\sqrt{n}}>2\Rightarrow n<98.0199$，也就是最多可以装 98 箱。

评 注　本题是求n，题目中隐含条件为$X\ge 0$，若写出此条件，则求解过程就变为

$$P\{0\le Y\le 5000\}=P\left\{\frac{0-50n}{5\sqrt{n}}\le\frac{Y-50n}{5\sqrt{n}}\le\frac{5000-50n}{5\sqrt{n}}\right\}\approx\Phi\left(-10\sqrt{n}\right)-\Phi\left(\frac{1000-10n}{\sqrt{n}}\right)$$

$\Rightarrow\Phi\left(-10\sqrt{n}\right)-\Phi\left(\dfrac{1000-10n}{\sqrt{n}}\right)>0.977\Rightarrow$无法解出结果。

但我们可以判断 $\Phi\left(-10\sqrt{n}\right)=1-\Phi\left(10\sqrt{n}\right)=0$，所以原解答没必要写出此隐含条件。

【例 16】利用中心极限定理证明　$\lim\limits_{n\to\infty}\sum\limits_{k=1}^{n}\dfrac{n^k}{k!}e^{-n}=\dfrac{1}{2}$

证明：设$\{X_i\}$独立同分布，均服从参数$\lambda=1$的泊松分布$\left(P\{X=k\}=\dfrac{\lambda^k e^{-\lambda}}{k!}=\dfrac{e^{-1}}{k!},\ k=0,1,\cdots\right)$，则由泊松分布的可加性知

$Y=\sum\limits_{i=1}^{n}X_i$ 服从$\lambda=n$的泊松分布，且$EY=n,\quad DY=n\Rightarrow Y\sim\dfrac{n^k e^{-n}}{k!}$

于是由列维—林德伯格中心极限定理知

$$\lim_{n\to\infty}\sum_{k=1}^{n}\frac{n^k}{k!}e^{-n}=\lim_{n\to\infty}P(Y\le n)=\lim_{n\to\infty}P(\sum_{i=1}^{n}X_i\le n)=\lim_{x\to\infty}P\left\{\frac{\sum\limits_{i=1}^{n}X_i-n}{\sqrt{n}}\le 0\right\}=\phi(0)=\frac{1}{2}$$

$\boxed{\text{评 注}}$ 在应用中心极限定理时可采用下列步骤：

第一步 根据题意选取独立同分布的随机变量序列$\{X_i\}$，求出EX_i，DX_i

第二步 弄清$Y=\sum_{i=1}^{n}X_i$所表示的意义，求出$EY,\ DY$，重新写出新的分布函数。

第三步 应用中心极限定理（独立同分布，林维--林德柏格），计算$P(a\le Y\le b)$

若 $E(X_i)=\mu, D(X_i)=\sigma^2$ 则

$$P(a\le Y\le b)=P\left(\frac{a-EY}{\sqrt{DY}}\le\frac{Y-EY}{\sqrt{DY}}\le\frac{b-EY}{\sqrt{DY}}\right)$$

$$P\left(\frac{a-n\mu}{\sigma\sqrt{n}}\le\frac{Y-EY}{\sqrt{n}\sqrt{DY}}\le\frac{b-n\mu}{\sigma\sqrt{n}}\right)\approx\Phi\left(\frac{b-n\mu}{\sigma\sqrt{n}}\right)-\Phi\left(\frac{a-n\mu}{\sigma\sqrt{n}}\right)=\Phi\left(\frac{b-n\mu}{\sigma\sqrt{n}}\right)+\Phi\left(-\frac{a-n\mu}{\sigma\sqrt{n}}\right)-1$$

【例 17】试利用切比雪夫不等式和中心极限定理，分别确定投掷一均匀硬币的次数，使得出现“正面向上”的频率在 0.4 和 0.6 之间的概率不少于 0.9。

解：设X表示投掷一枚均匀硬币n次“正面向上”的次数

$$X\sim B(n,\ 0.5)\Rightarrow EX=np=0.5n;\ \ DX=np(1-p)=0.25n$$

$$(1)\ \ P\left\{0.4<\frac{X}{n}<0.6\right\}=P\{0.4n<X<0.6n\}$$

$$=P\{0.4n-0.5n<X-0.5n<0.6n-0.5n\}$$

$$=P\{|X-0.5n|<0.1n\}\ge1-\frac{0.25n}{(0.1n)^2}=1-\frac{0.25n}{0.01n^2}=1-\frac{25}{n}$$

$$\Rightarrow1-\frac{25}{n}\ge0.9\Rightarrow n\ge250$$

$$(2)\ P\left\{0.4<\frac{X}{n}<0.6\right\}=P\left\{\frac{0.4n-0.5n}{\sqrt{0.25n}}<\frac{X-0.5n}{\sqrt{0.25n}}<\frac{0.6n-0.5n}{\sqrt{0.25n}}\right\}$$

$$=P\left\{-0.2\sqrt{n}<\frac{X-0.5n}{\sqrt{0.25n}}<0.2\sqrt{n}\right\}$$

$$=\Phi(0.2\sqrt{n})-\Phi(-0.2\sqrt{n})=2\Phi(0.2\sqrt{n})-1$$

$$2\Phi(0.2\sqrt{n})-1\ge0.9\Rightarrow0.2\sqrt{n}\ge1.65\Rightarrow n\ge67.5\Rightarrow n=68$$

$\boxed{\text{评 注}}$ 本题计算结果告诉我们，在能够确定具体分布的情形下，虽然切比雪夫不等式和中心极限定理都能求解同类问题，但利用切比雪夫不等式要粗糙得多，故一般不采用。

【例 18】(1)一个复杂系统由 100 个相互独立的元件组成，在系统运行期间每个元件损坏的概率为 0.1，若系统要正常运转，至少需要 85 个元件工作，求系统的可靠性（即系统正常运转的概率）。

(2)加入上述系统由n个相互独立的元件组成，而且至少要有 80%的元件正常工作才能使整个系统正常运转，问n至少多大时才能保证系统正常运行的可靠度为 0.95。

解：(1) 令 $X_k=\begin{cases}1, & 第k个元件没有损坏\\ 0, & 第k个元件损坏\end{cases}(k=1,2,\cdots,100)$，显然 $X_k\sim B(1,\ p)$

$X=\sum_{k=1}^{100}X_k\sim B(100,\ 0.9)$为系统正常运行时没有损坏的元件数。

$\Rightarrow EX=np=100\times0.9=90;\ DX=np(1-p)=100\times0.9\times0.1=9$

根据棣莫弗-拉普拉斯中心极限定理，所求可靠度为

$$P\{85\le X\le100\}=P\left\{\frac{85-90}{\sqrt{9}}\le\frac{X-90}{\sqrt{9}}\le\frac{100-90}{\sqrt{9}}\right\}\approx\Phi\left(\frac{10}{3}\right)-\Phi\left(-\frac{5}{3}\right)\approx1-\Phi\left(-\frac{5}{3}\right)=0.952。$$

(2) $X=\sum_{k=1}^{n}X_k\sim B(n,\ 0.9)$为系统正常运行时没有损坏的元件数。

$\Rightarrow EX=np=n\times0.9=0.9n;\ DX=np(1-p)=n\times0.9\times0.1=0.09n$

根据棣莫弗-拉普拉斯中心极限定理

$$P\{0.8n\le X\le100\}=P\left\{\frac{0.8n-0.9n}{\sqrt{0.09n}}\le\frac{X-0.9n}{\sqrt{0.09n}}\le\frac{100-0.9n}{\sqrt{0.09n}}\right\}$$
$$\approx\Phi\left(\frac{100-0.9n}{\sqrt{0.09n}}\right)-\Phi\left(-\frac{1}{3}\sqrt{n}\right)\approx1-\Phi\left(-\frac{1}{3}\sqrt{n}\right)=\Phi\left(\frac{1}{3}\sqrt{n}\right)=0.95$$

$$\Rightarrow\frac{1}{3}\sqrt{n}=1.65\Rightarrow n=25.$$

第五章 大数定律和中心极限定理模拟题

一．填空题

1. 设随机变量 X 的数学期望 E(X)=11，方差 D(X)＝9,则根据切比雪夫不等式估计，$P\{5<X<17\}\geq$_______.

2. 设 $X_1,X_2,\cdots,X_n$ 是 n 个相互独立的随机变量，且 $E(X_i)=\mu,D(X)=4(i=1,2,\cdots,n)$,

对于$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$，则由切比雪夫不等式估计有$P\{\mu-2<\overline{X}<\mu+2\}\geq$________.

3. 设随机变量 X 的数学期望 E(X)＝13,方差 D(X)＝4,用切比雪夫不等式估计得$P\{|X-13|\geq c\}\leq 0.01$，则 c=_________.

4. 设随机变量 X 和 Y 的数学期望都是 2,方差分别为 1 和 4，而相关系数为 0.5，则根据切比雪夫不等式估计 $P\{|X-Y|\geq 6\}\leq$_________.

二．选择题

1. 设随机变量 X 的方差存在，并且满足不等式$P\{|X-E(X)|\geq 3\}\leq\frac{2}{9}$，则一定有

（A） D(X)＝2　　（B）D(X)≠2

（C）$P\{|X-E(X)|<3\}\leq\frac{7}{9}$　　（D）$P\{|X-E(X)|<3\}\geq\frac{7}{9}$　　[　]

2. 设随机变量 X ～B(n, p), 对任意 0<p<1，利用切比雪夫不等式估计有$P\{|X-np|\geq\sqrt{2n}\}\leq$

（A）$\frac{1}{2}$　（B）$\frac{1}{4}$　（C）$\frac{1}{8}$　（D）$\frac{1}{16}$　　[　]

3. 设随机变量$X_1,\cdots,X_{16}$相互独立同分布，$E(X_i)=1,D(X_i)=1,i=1,\cdots,16.$

令$S_{16}=\sum_{i=1}^{16}X_i$, 则对任意$\varepsilon>0$, 从切比雪夫不等式直接可得

（A）$P\{|\frac{1}{16}S_{16}-1|<\varepsilon\}\geq 1-\frac{16}{\varepsilon^2}$　　（B）$P\{|S_{16}-16|<\varepsilon\}\geq 1-\frac{16}{\varepsilon^2}$

（C）$P\{|\frac{1}{16}S_{16}-1|<\varepsilon\}\geq 1-\frac{1}{\varepsilon^2}$　　（D）$P\{|S_{16}-16|<\varepsilon\}\geq 1-\frac{1}{\varepsilon^2}$　　[　]

4. 设随机变量$X_1,X_2,\cdots,X_n$是独立同分布，且X_i的密度函数为 f(x)，记$p=P\{\sum_{i=1}^{n}X_i\leq x_i\}$，当 n 充分大时，则有

（A）p 可以根据 f(x)进行计算　　（B）p 不可以根据 f(x)进行计算

（C）p 一定可以用中心极限定理近似计算　　（D）p 一定不能用中心极限定理近似计算 [ ]

三．解答题

1. 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，设各参加家长会的家长数相互独立，且服

从同一分布。

（1）求参加会议的家长数超过 450 的概率；

（2）求仅有 1 名家长来参加会议的学生数不多于 340 的概率。

2. 一生产线生产的产品成箱包装，每箱重量是随机的。假设每箱平均重 50 千克，标准差 5 千克。若用最大载重量为 5 吨的汽车承运，试利用中心极限定理说明每辆车最多装多少箱，才能保证不超载的概率大于 0.977.（$\phi(2)=0.977,$其中$\phi(x)$是标准正态分布函数）。

3. 设有 2500 个同一年龄段和同一社会阶层的人参加了某保险公司的人寿保险，假设在一年中每个人死亡的概率为 0.002，每个人在年初向保险公司交纳保费 120 元，而死亡时家属可以从保险公司领到 20000 元，问：

（1）保险公司亏本的概率是多少？

（2）保险公司获利不少于 100000 元的概率是多少？

（3）如果保险公司希望 99.9%可能性保证获利不少于 500000 元，问公司至少要发展多少个客户？

4. 分别用切比雪夫不等式和中心极限定理估计，当掷一枚均匀硬币时，需掷多少次，才能保证使得出现正面的频率在 0.45 和 0.55 之间的概率不小于 90%。

第五章 大数定律和中心极限定理模拟题答案

一．填空题

1. $\frac{3}{4}$　　2. $P\{|X-\mu|\}<2\ge 1-\frac{1}{n}$　　3. 20　　4. $\frac{1}{12}$

二．选择题

1. （D）　2. （C）　3. （B）　4. （A）

三．解答题

1．（1）0.1251；　（2）0.9938

2．最多可装 98 箱。

3. （1）0.000069；　（2）0.9874；　（3）公司至少要发展 4771 个客户。

4．切比雪夫不等式$n\ge 1000$, 中心极限定理估计$n\ge 136$。