第五章　大数定律及中心极限定理

极限定理是概率论的基本理论，在理论研究和应用中起着重要的作用，其中最重要的是称为“大数定律”与“中心极限定理”的一些定理. 大数定律是叙述随机变量序列的前一些项的算术平均值在某种条件下收敛到这些项的均值的算术平均值；中心极限定理则是确定在什么条件下，大量随机变量之和的分布逼近于正态分布. 本章介绍几个大数定理和中心极限定理.

§1　大数定律

第一章曾讲过，大量试验证实，随机事件 A 的频率 $f_n(A)$ 当重复试验的次数 n 增大时总呈现出稳定性，稳定在某一个常数的附近. 频率的稳定性是概率定义的客观基础. 本节我们将对频率的稳定性作出理论的说明.

弱大数定理（辛钦大数定理）　设 $X_1, X_2, \cdots$ 是相互独立[①]，服从同一分布的随机变量序列，且具有数学期望 $E(X_k)=\mu$ $(k=1,2,\cdots)$. 作前 n 个变量的算术平均 $\frac{1}{n}\sum_{k=1}^{n}X_k$，则对于任意 $\varepsilon>0$，有

$$\lim_{n\to\infty}P\left\{\left|\frac{1}{n}\sum_{k=1}^{n}X_k-\mu\right|<\varepsilon\right\}=1. \tag{1.1}$$

证　我们只在随机变量的方差 $D(X_k)=\sigma^2$ $(k=1,2,\cdots)$ 存在这一条件下证明上述结果. 因为

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

又由独立性得

$$D\left(\frac{1}{n}\sum_{k=1}^{n}X_k\right)=\frac{1}{n^2}\sum_{k=1}^{n}D(X_k)=\frac{1}{n^2}(n\sigma^2)=\frac{\sigma^2}{n},$$

由切比雪夫不等式（见第四章(2.9)式）得

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

在上式中令 $n\to\infty$，即得

① 是指对于任意 $n>1$，$X_1, X_2, \cdots, X_n$ 是相互独立的.

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

$\left\{\left|\frac{1}{n}\sum_{k=1}^{n}X_k-\mu\right|<\varepsilon\right\}$是一个随机事件. 等式(1.1)式表明,当$n\to\infty$时这个事件的概率趋于1. 即对于任意正数$\varepsilon$,当$n$充分大时,不等式$\left|\frac{1}{n}\sum_{k=1}^{n}X_k-\mu\right|<\varepsilon$成立的概率很大. 通俗地说,辛钦大数定理是说,对于独立同分布且具有均值μ的随机变量$X_1,\cdots,X_n$,当n很大时它们的算术平均$\frac{1}{n}\sum_{k=1}^{n}X_k$很可能接近于$\mu$.

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$\varepsilon$,有

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

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

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

依概率收敛的序列有以下的性质.

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

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

这样,上述定理又可叙述为:

弱大数定理(辛钦大数定理)　设随机变量$X_1,X_2,\cdots,X_n,\cdots$相互独立,服从同一分布且具有数学期望$E(X_k)=\mu\ (k=1,2,\cdots)$,则序列$\overline{X}=\frac{1}{n}\sum_{k=1}^{n}X_k$依概率收敛于$\mu$,即$\overline{X}\xrightarrow{P}\mu$.

下面介绍辛钦大数定理的一个重要推论.

伯努利大数定理　设f_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

$$\lim_{n\to\infty}P\left\{\left|\frac{f_A}{n}-p\right|<\varepsilon\right\}=1\tag{1.2}$$

或

$$\lim_{n\to\infty}P\left\{\left|\frac{f_A}{n}-p\right|\geqslant\varepsilon\right\}=0.\tag{1.2$'$}$$

证　因为$f_A\sim b(n,p)$,由第四章§2例6,有

$$f_A=X_1+X_2+\cdots+X_n,$$

其中,$X_1,X_2,\cdots,X_n$相互独立,且都服从以p为参数的$(0-1)$分布,因而$E(X_k)=p\ (k=1,2,\cdots,n)$,由(1.1)式即得

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

即

$$\lim_{n\to\infty}P\left\{\left|\frac{f_A}{n}-p\right|\geqslant\varepsilon\right\}=0.\qquad\square$$

伯努利大数定理的结果表明，对于任意 $\varepsilon>0$，只要重复独立试验的次数 n 充分大，事件$\left\{\left|\frac{f_A}{n}-p\right|\geqslant\varepsilon\right\}$是一个小概率事件，由实际推断原理知（见第一章§4），这一事件实际上几乎是不发生的，即在 n 充分大时事件$\left\{\left|\frac{f_A}{n}-p\right|<\varepsilon\right\}$实际上几乎是必定要发生的，亦即对于给定的任意小的正数 ε，在 n 充分大时，事件"频率$\frac{f_A}{n}$与概率 p 的偏差小于 ε"实际上几乎是必定要发生的. 这就是我们所说的频率稳定性的真正含义. 由实际推断原理，在实际应用中，当试验次数很大时，便可以用事件的频率来代替事件的概率.

§2　中心极限定理

在客观实际中有许多随机变量，它们是由大量的相互独立的随机因素的综合影响所形成的. 而其中每一个别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. 这种现象就是中心极限定理的客观背景. 本节只介绍三个常用的中心极限定理.

定理一（独立同分布的中心极限定理）　设随机变量 $X_1,X_2,\cdots,X_n,\cdots$ 相互独立，服从同一分布，且具有数学期望和方差：$E(X_k)=\mu,D(X_k)=\sigma^2>0\ (k=1,2,\cdots)$，则随机变量之和 $\sum\limits_{k=1}^{n}X_k$ 的标准化变量

$$Y_n=\frac{\sum\limits_{k=1}^{n}X_k-E\left(\sum\limits_{k=1}^{n}X_k\right)}{\sqrt{D\left(\sum\limits_{k=1}^{n}X_k\right)}}=\frac{\sum\limits_{k=1}^{n}X_k-n\mu}{\sqrt{n}\sigma}$$

的分布函数 $F_n(x)$ 对于任意 x 满足

$$\begin{aligned}\lim_{n\to\infty}F_n(x)&=\lim_{n\to\infty}P\left\{\frac{\sum\limits_{k=1}^{n}X_k-n\mu}{\sqrt{n}\sigma}\leqslant x\right\}\\&=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-t^2/2}\mathrm{d}t=\Phi(x).\end{aligned}\qquad(2.1)$$

证明略.

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 $X_1,X_2,\cdots,X_n$ 之和 $\sum\limits_{k=1}^{n}X_k$ 的标准化变量,当 n 充分大时,有

$$\frac{\sum\limits_{k=1}^{n}X_k-n\mu}{\sqrt{n}\sigma}\overset{\text{近似地}}{\sim}N(0,1). \tag{2.2}$$

在一般情况下,很难求出 n 个随机变量之和 $\sum\limits_{k=1}^{n}X_k$ 的分布函数,(2.2)式表明,当 n 充分大时,可以通过 $\Phi(x)$ 给出其近似的分布.这样,就可以利用正态分布对 $\sum\limits_{k=1}^{n}X_k$ 作理论分析或作实际计算,其好处是明显的.

将(2.2)式左端改写成 $\dfrac{\frac{1}{n}\sum\limits_{k=1}^{n}X_k-\mu}{\sigma/\sqrt{n}}=\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}$,这样,上述结果可写成:当 n 充分大时,

$$\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\overset{\text{近似地}}{\sim}N(0,1)\quad\text{或}\quad\overline{X}\overset{\text{近似地}}{\sim}N(\mu,\sigma^2/n). \tag{2.3}$$

这是独立同分布中心极限定理结果的另一个形式.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 $X_1,X_2,\cdots,X_n$ 的算术平均 $\overline{X}=\frac{1}{n}\sum\limits_{k=1}^{n}X_k$,当 n 充分大时近似地服从均值为 μ,方差为 σ^2/n 的正态分布.这一结果是数理统计中大样本统计推断的基础.

定理二(李雅普诺夫(Lyapunov)定理)　设随机变量 $X_1,X_2,\cdots,X_n,\cdots$ 相互独立,它们具有数学期望和方差

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

记

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

若存在正数 δ,使得当 $n\to\infty$ 时,

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

则随机变量之和 $\sum\limits_{k=1}^{n}X_k$ 的标准化变量

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

的分布函数 $F_n(x)$ 对于任意 x，满足

$$\begin{aligned}\lim_{n\to\infty}F_n(x) &= \lim_{n\to\infty}P\left\{\frac{\sum_{k=1}^{n}X_k-\sum_{k=1}^{n}\mu_k}{B_n}\leqslant x\right\}\\ &= \int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-t^2/2}\mathrm{d}t=\Phi(x).\end{aligned} \tag{2.4}$$

证明略.

定理二表明，在定理的条件下，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，近似地服从正态分布 $N(0,1)$. 由此，当 n 很大时，$\sum_{k=1}^{n}X_k=B_nZ_n+\sum_{k=1}^{n}\mu_k$ 近似地服从正态分布 $N\left(\sum_{k=1}^{n}\mu_k,B_n^2\right)$. 这就是说，无论各个随机变量 X_k $(k=1,2,\cdots)$ 服从什么分布，只要满足定理的条件，那么它们的和 $\sum_{k=1}^{n}X_k$ 当 n 很大时，就近似地服从正态分布. 这就是为什么正态随机变量在概率论中占有重要地位的一个基本原因. 在很多问题中，所考虑的随机变量可以表示成很多个独立的随机变量之和，例如，在任一指定时刻，一个城市的耗电量是大量用户耗电量的总和；一个物理实验的测量误差是由许多观察不到的、可加的微小误差所合成的，它们往往近似地服从正态分布.

下面介绍另一个中心极限定理，它是定理一的特殊情况.

定理三（棣莫弗—拉普拉斯(De Moivre-Laplace)定理） 设随机变量 η_n $(n=1,2,\cdots)$ 服从参数为 n,p $(0<p<1)$ 的二项分布，则对于任意 x，有

$$\lim_{n\to\infty}P\left\{\frac{\eta_n-np}{\sqrt{np(1-p)}}\leqslant x\right\}=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-t^2/2}\mathrm{d}t=\Phi(x). \tag{2.5}$$

证 由第四章§2例6知可以将 η_n 分解成为 n 个相互独立、服从同一(0—1)分布的诸随机变量 $X_1,X_2,\cdots,X_n$ 之和，即有

$$\eta_n=\sum_{k=1}^{n}X_k,$$

其中 X_k $(k=1,2,\cdots,n)$ 的分布律为

$$P\{X_k=i\}=p^i(1-p)^{1-i},\quad i=0,1.$$

由于 $E(X_k)=p,D(X_k)=p(1-p)$ $(k=1,2,\cdots,n)$，由定理一得

$$\lim_{n\to\infty}P\left\{\frac{\eta_n-np}{\sqrt{np(1-p)}}\leqslant x\right\}=\lim_{n\to\infty}P\left\{\frac{\sum_{k=1}^{n}X_k-np}{\sqrt{np(1-p)}}\leqslant x\right\}$$

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

这个定理表明，正态分布是二项分布的极限分布. 当 n 充分大时，我们可以利用(2.5)式来计算二项分布的概率. 下面举几个关于中心极限定理应用的例子.

例 1　一加法器同时收到 20 个噪声电压 V_k $(k=1,2,\cdots,20)$，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 $E(V_k)=5, D(V_k)=100/12$ $(k=1,2,\cdots,20)$. 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

$$\begin{aligned} P\{V>105\} &= P\left\{\frac{V-20\times 5}{(10/\sqrt{12})\sqrt{20}} > \frac{105-20\times 5}{(10/\sqrt{12})\sqrt{20}}\right\} \\ &= P\left\{\frac{V-100}{(10/\sqrt{12})\sqrt{20}} > 0.387\right\} \\ &= 1 - P\left\{\frac{V-100}{(10/\sqrt{12})\sqrt{20}} \leqslant 0.387\right\} \\ &\approx 1 - \int_{-\infty}^{0.387} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt = 1-\Phi(0.387) = 0.348. \end{aligned}$$

即有

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

例 2　一船舶在某海区航行，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 90 000 次波浪冲击，问其中有 29 500～30 500 次纵摇角度大于 3°的概率是多少?

解　我们将船舶每遭受一次波浪冲击看作是一次试验，并假定各次试验是独立的. 在 90 000 次波浪冲击中纵摇角度大于 3°的次数记为 X，则 X 是一个随机变量，且有 $X\sim b(90\,000, 1/3)$. 其分布律为

$$P\{X=k\} = \binom{90\,000}{k}\left(\frac{1}{3}\right)^k \left(\frac{2}{3}\right)^{90\,000-k}, k=0,1,\cdots,90\,000.$$

所求的概率为

$$P\{29\,500 \leqslant X \leqslant 30\,500\} = \sum_{k=29\,500}^{30\,500} \binom{90\,000}{k}\left(\frac{1}{3}\right)^k \left(\frac{2}{3}\right)^{90\,000-k}.$$

要直接计算是麻烦的，我们利用棣莫弗－拉普拉斯定理来求它的近似值. 即有

$$
\begin{aligned}
&P\{29\ 500 \leqslant X \leqslant 30\ 500\} \\
&= P\left\{\frac{29\ 500 - np}{\sqrt{np(1-p)}} \leqslant \frac{X - np}{\sqrt{np(1-p)}} \leqslant \frac{30\ 500 - np}{\sqrt{np(1-p)}}\right\} \\
&\approx \int_{\frac{29\ 500-np}{\sqrt{np(1-p)}}}^{\frac{30\ 500-np}{\sqrt{np(1-p)}}} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-t^2/2} \mathrm{d}t = \Phi\left(\frac{30\ 500 - np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{29\ 500 - np}{\sqrt{np(1-p)}}\right),
\end{aligned}
$$

其中 $n = 90\ 000, p = 1/3$. 即有

$$
P\{29\ 500 \leqslant X \leqslant 30\ 500\} \approx \Phi\left(\frac{5\sqrt{2}}{2}\right) - \Phi\left(-\frac{5\sqrt{2}}{2}\right) = 0.999\ 5. \qquad \square
$$

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、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.

解 (1) 以 X_k ($k=1,2,\cdots,400$) 记第 k 个学生来参加会议的家长人数，则 X_k 的分布律为

X_k	0	1	2
p_k	0.05	0.8	0.15

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

$$
\begin{aligned}
P\{X > 450\} &= P\left\{\frac{X - 400 \times 1.1}{\sqrt{400}\ \sqrt{0.19}} > \frac{450 - 400 \times 1.1}{\sqrt{400}\ \sqrt{0.19}}\right\} \\
&= 1 - P\left\{\frac{X - 400 \times 1.1}{\sqrt{400}\ \sqrt{0.19}} \leqslant 1.147\right\} \\
&\approx 1 - \Phi(1.147) = 0.125\ 1.
\end{aligned}
$$

(2) 以 Y 记有一名家长参加会议的学生人数，则 $Y \sim b(400, 0.8)$，由定理三

$$
\begin{aligned}
&P\{Y \leqslant 340\} \\
&= P\left\{\frac{Y - 400 \times 0.8}{\sqrt{400 \times 0.8 \times 0.2}} \leqslant \frac{340 - 400 \times 0.8}{\sqrt{400 \times 0.8 \times 0.2}}\right\}
\end{aligned}
$$

$$=P\left\{\frac{Y-400\times0.8}{\sqrt{400\times0.8\times0.2}}\leqslant2.5\right\}\approx\Phi(2.5)=0.9938.$$ □

小结

人们在长期实践中认识到频率具有稳定性,即当试验次数不断增大时,频率稳定在一个数的附近．这一事实显示了可以用一个数来表征事件发生的可能性的大小．这使人们认识到概率是客观存在的,进而由频率的性质的启发和抽象给出了概率的定义,因而频率的稳定性是概率定义的客观基础．伯努利大数定理则以严密的数学形式论证了频率的稳定性．

中心极限定理表明,在相当一般的条件下,当独立随机变量的个数不断增加时,其和的分布趋于正态分布．这一事实阐明了正态分布的重要性．也揭示了为什么在实际应用中会经常遇到正态分布,也就是揭示了产生正态分布变量的源泉．另一方面,它提供了独立同分布随机变量之和 $\sum_{k=1}^{n}X_k$（其中 X_k 的方差存在）的近似分布,只要和式中加项的个数充分大,就可以不必考虑和式中的随机变量服从什么分布,都可以用正态分布来近似,这在应用上是有效和重要的．

中心极限定理的内容包含极限,因而称它为极限定理是很自然的．又由于它在统计中的重要性,称它为中心极限定理,这是波利亚(Polya)在 1920 年取的名字．

■ 重要术语及主题

依概率收敛　伯努利大数定理　辛钦大数定理　独立同分布的中心极限定理　李雅普诺夫中心极限定理　棣莫弗－拉普拉斯中心极限定理

习题

1. 据以往经验,某种电器元件的寿命服从均值为 100 h 的指数分布,现随机地取 16 只,设它们的寿命是相互独立的.求这 16 只元件的寿命的总和大于 1920 h 的概率.

2. (1) 一保险公司有 10 000 个汽车投保人,每个投保人索赔金额的数学期望为 280 美元,标准差为 800 美元,求索赔总金额超过 2 700 000 美元的概率.

(2) 一公司有 50 张签约保险单,各张保险单的索赔金额为 $X_i, i=1,2,\cdots,50$（以千美元计）服从韦布尔(Weibull)分布,均值 $E(X_i)=5$,方差 $D(X_i)=6$,求 50 张保险单索赔的合计金额大于 300 的概率(设各保险单索赔金额是相互独立的).

3. 计算器在进行加法时,将每个加数舍入最靠近它的整数,设所有舍入误差相互独立且在 $(-0.5,0.5)$ 上服从均匀分布.

(1) 将1 500 个数相加,问误差总和的绝对值超过 15 的概率是多少?

(2) 最多可有几个数相加使得误差总和的绝对值小于 10 的概率不小于 0.90?

4. 设各零件的重量都是随机变量,它们相互独立,且服从相同的分布,其数学期望为 0.5 kg,均方差为 0.1 kg,问 5 000 只零件的总重量超过 2 510 kg 的概率是多少?

5. 有一批建筑房屋用的木柱,其中 80% 的长度不小于 3 m,现从这批木柱中随机地取 100 根,求其中至少有 30 根短于 3 m 的概率.

6. 一工人修理一台机器需两个阶段,第一阶段所需时间(小时)服从均值为 0.2 的指数

分布，第二阶段服从均值为 0.3 的指数分布，且与第一阶段独立. 现有 20 台机器需要修理，求他在 8 h 内完成的概率.

7. 一食品店有三种蛋糕出售，由于售出哪一种蛋糕是随机的，因而售出一只蛋糕的价格是一个随机变量，它取 1 元、1.2元、1.5 元各个值的概率分别为 0.3、0.2、0.5. 若售出 300 只蛋糕.

(1) 求收入至少 400 元的概率.

(2) 求售出价格为1.2元的蛋糕多于 60 只的概率.

8. 一复杂的系统由 100 个相互独立起作用的部件所组成，在整个运行期间每个部件损坏的概率为 0.10. 为了使整个系统起作用，至少必须有 85 个部件正常工作，求整个系统起作用的概率.

9. 已知在某十字路口，一周事故发生数的数学期望为 2.2，标准差为1.4.

(1) 以 $\overline{X}$ 表示一年(以 52 周计)此十字路口事故发生数的算术平均，试用中心极限定理求 $\overline{X}$ 的近似分布，并求 $P\{\overline{X}<2\}$.

(2) 求一年事故发生数小于 100 的概率.

10. 某种小汽车氧化氮的排放量的数学期望为 0.9 g/km，标准差为1.9 g/km，某汽车公司有这种小汽车 100 辆，以 $\overline{X}$ 表示这些车辆氧化氮排放量的算术平均，问当 L 为何值时 $\overline{X}>L$ 的概率不超过 0.01.

11. 随机地选取两组学生，每组 80 人，分别在两个实验室里测量某种化合物的 pH. 各人测量的结果是随机变量，它们相互独立，服从同一分布，数学期望为 5，方差为 0.3，以 $\overline{X}$，$\overline{Y}$ 分别表示第一组和第二组所得结果的算术平均.

(1) 求$P\{4.9<\overline{X}<5.1\}$.

(2) 求 $P\{-0.1<\overline{X}-\overline{Y}<0.1\}$.

12. 一公寓有 200 户住户，一户住户拥有汽车辆数 X 的分布律为

X	0	1	2
p_k	0.1	0.6	0.3

问需要多少车位，才能使每辆汽车都具有一个车位的概率至少为 0.95.

13. 某种电子器件的寿命(小时)具有数学期望 μ(未知)，方差 $\sigma^2=400$. 为了估计 μ，随机地取 n 只这种器件，在时刻 $t=0$ 投入测试(测试是相互独立的)直到失效，测得其寿命为 X_1，X_2，…，X_n，以 $\overline{X}=\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i$ 作为μ 的估计，为使 $P\{|\overline{X}-\mu|<1\}\geqslant 0.95$，问 n 至少为多少？

14. 某药厂断言，该厂生产的某种药品对于医治一种疑难血液病的治愈率为 0.8，医院任意抽查 100 个服用此药品的病人，若其中多于 75 人治愈，就接受此断言，否则就拒绝此断言.

(1) 若实际上此药品对这种疾病的治愈率是 0.8. 问接受这一断言的概率是多少？

(2) 若实际上此药品对这种疾病的治愈率为 0.7，问接受这一断言的概率是多少？