

第五章 大数定律与中心极限定理

第一节 大数定律

在第一章中我们已经指出,人们经过长期实践认识到,虽然个别随机事件在某次试验中可能发生也可能不发生,但是在大量重复试验中却呈现明显的规律性,即随着试验次数的增大,一个随机事件发生的频率在某一固定值附近摆动.这就是所谓的频率具有稳定性.同时,人们通过实践发现大量测量值的算术平均值也具有稳定性,也就是说,无论个别随机现象的结果如何,以及它们在进行过程中的个别特征如何,大量随机现象的平均值实际上是一个与个别现象特征无关的常数,几乎不再是随机的了.概率论中用来阐明大量随机现象的平均值的稳定性的一系列定律称为大数定律.

在引入大数定律之前,我们先证明一个重要的不等式——切比雪夫(Chebyshev)不等式.

定理 5.1 (切比雪夫不等式) 设随机变量 X 存在有限方差 $D(X)$, 则对任意 $\varepsilon > 0$, 有

$$P\{|X-E(X)| \geq \varepsilon\} \leq \frac{D(X)}{\varepsilon^2}. \quad (5-1)$$

证 如果 X 是连续型随机变量, 设 X 的概率密度为 $f(x)$, 则有

$$\begin{aligned} P\{|X-E(X)| \geq \varepsilon\} &= \int_{|x-E(X)| \geq \varepsilon} f(x) dx \leq \int_{|x-E(X)| \geq \varepsilon} \frac{|x-E(X)|^2}{\varepsilon^2} f(x) dx \\ &\leq \frac{1}{\varepsilon^2} \int_{-\infty}^{+\infty} [x-E(X)]^2 f(x) dx = \frac{D(X)}{\varepsilon^2}. \end{aligned}$$

请读者自己证明 X 是离散型随机变量的情形.

切比雪夫不等式也可表示成

$$P\{|X-E(X)| < \varepsilon\} \geq 1 - \frac{D(X)}{\varepsilon^2}. \quad (5-2)$$

这个不等式给出了在随机变量 X 的分布未知的情况下事件 $\{|X-E(X)| < \varepsilon\}$ 的概率的下限估计, 例如, 在切比雪夫不等式中, 令 $\varepsilon = 3\sqrt{D(X)}$ 以及 $4\sqrt{D(X)}$ 分别可得到

$$P\{|X-E(X)| < 3\sqrt{D(X)}\} \geq 0.8889,$$

$$P\{|X-E(X)| < 4\sqrt{D(X)}\} \geq 0.9375.$$

例 5.1 设 X 表示掷一颗骰子所出现的点数, 若给定 $\varepsilon = 1, 2$, 请实际计算 $P\{|X-E(X)| \geq \varepsilon\}$, 并验证切比雪夫不等式成立.

解 因为 X 的分布律是 $P\{X=k\} = 1/6 \ (k=1, 2, \dots, 6)$, 所以

$$E(X) = 7/2, \quad D(X) = 35/12,$$

$$P\{|X-7/2| \geq 1\} = P\{X=1\} + P\{X=2\} + P\{X=5\} + P\{X=6\} = 2/3;$$

$$P\{|X-7/2| \geq 2\} = P\{X=1\} + P\{X=6\} = 1/3.$$

$$\begin{aligned}\varepsilon=1: \quad \frac{D(X)}{\varepsilon^2} &= 35/12 > 2/3, \\ \varepsilon=2: \quad \frac{D(X)}{\varepsilon^2} &= 1/4 \times 35/12 = 35/48 > 1/3.\end{aligned}$$

可见切比雪夫不等式成立.

例 5.2 设电站供电网有 10000 盏电灯, 夜晚每一盏灯开灯的概率都是 0.7, 而假定开、关时间彼此独立, 估计夜晚同时开着的灯数在 6800 与 7200 之间的概率.

解 设 X 表示在夜晚同时开着的灯的数目, 它服从参数为 $n=10000, p=0.7$ 的二项分布. 若要准确计算, 应该用伯努利公式:

$$P\{6800 < X < 7200\} = \sum_{k=6801}^{7199} C_{10000}^k \times 0.7^k \times 0.3^{10000-k}.$$

如果用切比雪夫不等式估计:

$$\begin{aligned}E(X) &= np = 10000 \times 0.7 = 7000, \\ D(X) &= npq = 10000 \times 0.7 \times 0.3 = 2100, \\ P\{6800 < X < 7200\} &= P\{|X - 7000| < 200\} \geq 1 - \frac{2100}{200^2} \approx 0.95.\end{aligned}$$

可见, 虽然有 10000 盏灯, 但是只要有供应 7200 盏灯的电力就能够以相当大的概率保证够用. 事实上, 切比雪夫不等式的估计只说明概率大于 0.95, 后面将具体求出这个概率约为 0.99999. 切比雪夫不等式在理论上具有重大意义, 但估计的精确度不高.

切比雪夫不等式作为一个理论工具, 在大数定律证明中, 可使证明非常简洁.

定义 5.1 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 $Y_n \xrightarrow{P} a$.

定理 5.2 (切比雪夫大数定律) 设 X_1, X_2, \dots 是相互独立的随机变量序列, 各有数学期望 $E(X_1), E(X_2), \dots$ 及方差 $D(X_1), D(X_2), \dots$, 并且对于所有 $i=1, 2, \dots$ 都有 $D(X_i) < l$, 其中 l 是与 i 无关的常数, 则对任给 $\varepsilon > 0$, 有

$$\lim_{n \rightarrow \infty} P\left\{\left|\frac{1}{n} \sum_{i=1}^n X_i - \frac{1}{n} \sum_{i=1}^n E(X_i)\right| < \varepsilon\right\} = 1. \quad (5-3)$$

证 因 X_1, X_2, \dots 相互独立, 所以

$$D\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n^2} \sum_{i=1}^n D(X_i) < \frac{1}{n^2} \cdot nl = \frac{l}{n}.$$

又因

$$E\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n} \sum_{i=1}^n E(X_i),$$

由(5-2)式, 对于任意 $\varepsilon > 0$, 有

$$P\left\{\left|\frac{1}{n} \sum_{i=1}^n X_i - \frac{1}{n} \sum_{i=1}^n E(X_i)\right| < \varepsilon\right\} \geq 1 - \frac{l}{n\varepsilon^2},$$

但是任何事件的概率都不超过 1, 即

$$1 - \frac{l}{n\varepsilon^2} \leq P\left\{\left|\frac{1}{n}\sum_{i=1}^n X_i - \frac{1}{n}\sum_{i=1}^n E(X_i)\right| < \varepsilon\right\} \leq 1,$$

故

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

切比雪夫大数定律说明: 在定理的条件下, 当 n 充分大时, n 个独立随机变量的平均数

这个随机变量的离散程度是很小的. 这意味着, 经过算术平均以后得到的随机变量 $\frac{\sum_{i=1}^n X_i}{n}$ 将

比较密地聚集在它的数学期望 $\frac{\sum_{i=1}^n E(X_i)}{n}$ 的附近, 它与数学期望之差依概率收敛到 0.

推论 (切比雪夫大数定律的特殊情况) 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 $Y_n = \frac{1}{n} \sum_{k=1}^n X_k$ 则对于任意正数 ε 有

$$\lim_{n \rightarrow \infty} P\{|Y_n - \mu| < \varepsilon\} = 1. \quad (5-4)$$

定理 5.3 (伯努利大数定律) 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 ε , 有

$$\lim_{n \rightarrow \infty} P\left\{\left|\frac{n_A}{n} - p\right| < \varepsilon\right\} = 1, \quad (5-5)$$

或

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

证 引入随机变量

$$X_k = \begin{cases} 0, & \text{若在第 } k \text{ 次试验中 } A \text{ 不发生,} \\ 1, & \text{若在第 } k \text{ 次试验中 } A \text{ 发生,} \end{cases} \quad k=1, 2, \dots,$$

显然

$$n_A = \sum_{k=1}^n X_k.$$

由于 X_k 只依赖于第 k 次试验, 而各次试验是相互独立的. 因此 X_1, X_2, \dots 是相互独立的; 又由于 X_k 服从 (0-1) 分布, 因此有

$$E(X_k) = p, \quad D(X_k) = p(1-p), \quad k=1, 2, \dots.$$

由定理 5.2 地推论, 有

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

即

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

伯努利大数定律告诉我们, 事件 A 发生的频率 $\frac{n_A}{n}$ 依概率收敛于事件 A 发生的概率 p , 因此, 本定律从理论上证明了大量重复独立试验中, 事件 A 发生的频率具有稳定性, 正因为这种稳定性, 概率的概念才有实际意义. 伯努利大数定律还提供了通过试验来确定事件的概率的方法, 即既然频率 $\frac{n_A}{n}$ 与概率 p 有较大偏差的可能性很小, 于是我们就可以通过做试验确定某事件发生的频率, 并把它作为相应概率的估计. 因此, 在实际应用中, 如果试验的次数很大时, 就可以用事件发生的频率代替事件发生的概率.

定理 5.2 及其推论中要求随机变量 X_k ($k=1,2,\dots,n$) 的方差存在. 但在随机变量服从同一分布的情形下, 并不需要这一要求, 我们有以下定理.

定理 5.4 (辛钦 (Khinchin) 大数定律) 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

$$\lim_{n \rightarrow \infty} P \left\{ \left| \frac{1}{n} \sum_{k=1}^n X_k - \mu \right| < \varepsilon \right\} = 1. \quad (5-6)$$

显然, 伯努利大数定律是辛钦大数定律的特殊情况, 辛钦大数定律在实际中应用很广泛. 这一定律使算术平均值的法则有了理论根据. 如要测定某一物理量 a , 在不变的条件下重

复测量 n 次, 得观测值 X_1, X_2, \dots, X_n , 求得实测值的算术平均值 $\frac{1}{n} \sum_{i=1}^n X_i$, 根据此定理,

当 n 足够大时, 取 $\frac{1}{n} \sum_{i=1}^n X_i$ 作为 a 的近似值, 可以认为所发生的误差是很小的, 所以实用上

往往用某物体的某一指标值的一系列实测值的算术平均值来作为该指标值的近似值.

第二节 中心极限定理

在客观实际中有许多随机变量, 它们是由大量相互独立的偶然因素的综合影响所形成的, 而每一个因素在总的影响中所起的作用是很小的, 但总的来说, 却对总和有显著影响, 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. 概率论中有关论证独立随机变量之和的极限分布是正态分布的一系列定理称为中心极限定理 (Central limit theorem), 现介绍几个常用的中心极限定理.

定理 5.5 (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:

$$E(X_k) = \mu, D(X_k) = \sigma^2 \neq 0 (k=1,2,\dots).$$

则随机变量

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

的分布函数 $F_n(x)$ 对于任意 x 满足

$$\lim_{n \rightarrow \infty} F_n(x) = \lim_{n \rightarrow \infty} P\left\{\frac{\sum_{k=1}^n X_k - n\mu}{\sqrt{n}\sigma} \leq x\right\} = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt. \quad (5-7)$$

从定理 5.5 的结论可知, 当 n 充分大时, 近似地, 有

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

或者说, 当 n 充分大时, 近似地, 有

$$\sum_{k=1}^n X_k \sim N(n\mu, n\sigma^2). \quad (5-8)$$

如果用 X_1, X_2, \dots, X_n 表示相互独立的各随机因素. 假定它们都服从相同的分布 (不论服从什么分布), 且都有有限的期望与方差 (每个因素的影响有一定限度). 则 (5-8) 式说

明, 作为总和 $\sum_{k=1}^n X_k$ 这个随机变量, 当 n 充分大时, 便近似地服从正态分布.

例 5.3 一个螺丝钉重量是一个随机变量, 期望值是 100 克, 标准差是 10 克. 求一盒 (100 个) 同型号螺丝钉的重量超过 10.2 斤的概率.

解 设一盒重量为 X , 盒中第 i 个螺丝钉的重量为 $X_i, i=1, 2, \dots, 100$. X_1, X_2, \dots, X_{100} 相互独立, $E(X_i) = 100, \sqrt{D(X_i)} = 10$, 则有

$$X = \sum_{i=1}^{100} X_i, \text{ 且 } E(X) = 100 \cdot E(X_i) = 10000 \text{ (克)}, \sqrt{D(X)} = 100 \text{ (克)}.$$

根据定理 5.5, 有

$$P\{X > 10200\} = P\left\{\frac{X - 10000}{100} > \frac{10200 - 10000}{100}\right\} = 1 - P\left\{\frac{X - 10000}{100} \leq 2\right\} \\ \approx 1 - \Phi(2) = 1 - 0.9772 = 0.0228.$$

例 5.4 对敌人的防御地进行 100 次轰炸, 每次轰炸命中目标的炸弹数目是一个随机变量, 其期望值是 2, 方差是 1.69. 求在 100 次轰炸中有 180 颗到 220 颗炸弹命中目标的概率.

解 令第 i 次轰炸命中目标的炸弹数为 X_i , 100 次轰炸中命中目标炸弹数 $X = \sum_{i=1}^{100} X_i$,

应用定理 5.5, X 渐近服从正态分布, 期望值为 200, 方差为 169, 标准差为 13. 所以

$$P\{180 \leq X \leq 220\} = P\{|X-200| \leq 20\} = P\left\{\left|\frac{X-200}{13}\right| \leq \frac{20}{13}\right\}$$

$$\approx 2\Phi(1.54) - 1 = 0.8764.$$

定理 5.6(李雅普诺夫 (Liapunov) 定理) 设随机变量 X_1, X_2, \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

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

的分布函数 $F_n(x)$ 对于任意 x , 满足

$$\lim_{n \rightarrow \infty} F_n(x) = \lim_{n \rightarrow \infty} P\left\{\frac{\sum_{k=1}^n X_k - \sum_{k=1}^n \mu_k}{B_n} \leq x\right\} = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt. \quad (5-9)$$

这个定理说明, 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 $N(0, 1)$. 因此, 当 n 很大时,

$$\sum_{k=1}^n X_k = B_n Z_n + \sum_{k=1}^n \mu_k$$

近似地服从正态分布 $N\left(\sum_{k=1}^n \mu_k, B_n^2\right)$. 这表明无论随机变量 X_k ($k=1, 2, \dots$) 具有怎样的分

布, 只要满足定理条件, 则它们的和 $\sum_{k=1}^n X_k$ 当 n 很大时, 就近似地服从正态分布. 而在许多

实际问题中, 所考虑的随机变量往往可以表示为多个独立的随机变量之和, 因而它们常常近似服从正态分布. 这就是为什么正态随机变量在概率论与数理统计中占有重要地位的主要原因.

在数理统计中我们将看到, 中心极限定理是大样本统计推断的理论基础.

下面介绍另一个中心极限定理.

定理 5.7 设随机变量 X 服从参数为 n, p ($0 < p < 1$) 的二项分布, 则

(1) (拉普拉斯(Laplace)定理) 局部极限定理: 当 $n \rightarrow \infty$ 时

$$P\{X=k\} \approx \frac{1}{\sqrt{2\pi npq}} e^{-\frac{(k-np)^2}{2npq}} = \frac{1}{\sqrt{npq}} \phi\left(\frac{k-np}{\sqrt{npq}}\right), \quad (5-10)$$

其中 $p+q=1, k=0,1,2,\dots,n$, $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$.

(2) (德莫佛-拉普拉斯(De Moivre Laplace)定理) 积分极限定理: 对于任意的 x , 恒有

$$\lim_{n \rightarrow \infty} P\left\{\frac{X-np}{\sqrt{np(1-p)}} \leq x\right\} = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt. \quad (5-11)$$

这个定理表明, 二项分布以正态分布为极限. 当 n 充分大时, 我们可以利用上两式来计算二项分布的概率.

例 5.5 10 部机器独立工作, 每部停机的概率为 0.2, 求 3 部机器同时停机的概率.

解 10 部机器中同时停机的数目 X 服从二项分布, $n=10, p=0.2$, 则 $np=2, \sqrt{npq} \approx 1.265$.

(1) 直接计算: $P\{X=3\} = C_{10}^3 \times 0.2^3 \times 0.8^7 \approx 0.2013$;

(2) 若用局部极限定理近似计算:

$$P\{X=3\} = \frac{1}{\sqrt{npq}} \phi\left(\frac{k-np}{\sqrt{npq}}\right) = \frac{1}{1.265} \phi\left(\frac{3-2}{1.265}\right) \approx \frac{1}{1.265} \phi(0.79) \approx 0.2308.$$

(2) 的计算结果与 (1) 相差较大, 这是由于 n 不够大.

例 5.6 应用定理 5.7 计算第五章第一节中例 5.2 的概率.

解 $np=7000, \sqrt{npq} \approx 45.83$.

$$\begin{aligned} P\{6800 < X < 7200\} &= P\{|X-7000| < 200\} \\ &\approx P\left\{\left|\frac{X-7000}{45.83}\right| < 4.36\right\} \approx 2\Phi(4.36) - 1 \\ &= 0.99999. \end{aligned}$$

例 5.7 产品为废品的概率为 $p=0.005$, 求 10000 件产品中废品数不大于 70 的概率.

解 10000 件产品中的废品数 X 服从二项分布, $n=10000, p=0.005, np=50, \sqrt{npq} \approx 7.053$.

$$P\{X \leq 70\} \approx \Phi\left(\frac{70-50}{7.053}\right) \approx \Phi(2.84) = 0.9977.$$

正态分布和泊松分布虽然都是二项分布的极限分布, 但后者以 $n \rightarrow \infty$, 同时 $p \rightarrow 0$, $np \rightarrow \lambda$ 为条件, 而前者则只要求 $n \rightarrow \infty$ 这一条件. 一般说来, 对于 n 很大, p (或 q) 很小的二项分布 ($np \leq 5$) 用正态分布来近似计算不如用泊松分布计算精确.

例 5.8 每颗炮弹命中飞机的概率为 0.01, 求 500 发炮弹中命中 5 发的概率.

解 500 发炮弹中命中飞机的炮弹数目 X 服从二项分布, $n=500, p=0.01, np=5, \sqrt{npq} \approx 2.2$.

下面用三种方法计算并加以比较:

(1) 用二项分布公式计算:

$$P\{X=5\} = C_{500}^5 \times 0.01^5 \times 0.99^{495} = 0.17635.$$

(2) 用泊松公式计算, 直接查表可得:

$$np = \lambda = 5, k = 5, P_5(5) \approx 0.175467.$$

(3) 用拉普拉斯局部极限定理计算:

$$P\{X=5\} \approx \frac{1}{\sqrt{npq}} \varphi\left(\frac{5-np}{\sqrt{npq}}\right) \approx 0.1793.$$

可见后者不如前者精确.

小 结

本章介绍了切比雪夫不等式、四个大数定律和三个中心极限定理.

切比雪夫不等式给出了随机变量 X 的分布未知, 只知道 $E(X)$ 和 $D(X)$ 的情况下, 对事件 $\{|X-E(X)| \leq \varepsilon\}$ 概率的下限估计.

人们在长期实践中认识到频率具有稳定性, 即当试验次数增大时, 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. 这使人们认识到概率是客观存在的, 进而由频率的 3 条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. 伯努利大数定律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增加时, 其和的分布趋于正态分布. 它阐明了正态分布的重要性. 中心极限定理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.

本章要求读者理解大数定律和中心极限定理的概率意义, 并要求会使用中心极限定理估算有关事件的概率.

重要术语及主题

切比雪夫不等式

切比雪夫大数定律及特殊情况

辛钦大数定律

李雅普诺夫中心极限定理

依概率收敛

伯努利大数定律

独立同分布中心极限定律

德莫佛-拉普拉斯中心极限定理.

习 题 五

1. 一颗骰子连续掷 4 次, 点数总和记为 X . 估计 $P\{10 < X < 18\}$.
2. 假设一条生产线生产的产品合格率是 0.8. 要使一批产品的合格率达到在 76% 与 84% 之间的概率不小于 90%, 问这批产品至少要生产多少件?
3. 某车间有同型号机床 200 部, 每部机床开动的概率为 0.7, 假定各机床开动与否互不影响, 开动时每部机床消耗电能 15 个单位. 问至少供应多少单位电能才可以 95% 的概率保证不致因供电不足而影响生产.
4. 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.

5. 有一批建筑房屋用的木柱, 其中 80% 的长度不小于 3m. 现从这批木柱中随机地取出 100 根, 问其中至少有 30 根短于 3m 的概率是多少?
6. 某药厂断言, 该厂生产的某种药品对于医治一种疑难的血液病的治愈率为 0.8. 医院检验员任意抽查 100 个服用此药品的病人, 如果其中多于 75 人治愈, 就接受这一断言, 否则就拒绝这一断言.
 - (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是 0.7, 问接受这一断言的概率是多少?
7. 用拉普拉斯中心极限定理近似计算从一批废品率为 0.05 的产品中, 任取 1000 件, 其中有 20 件废品的概率.
8. 设有 30 个电子器件. 它们的使用寿命 T_1, \dots, T_{30} 服从参数 $\lambda=0.1$ [单位: (小时)⁻¹] 的指数分布, 其使用情况是第一个损坏第二个立即使用, 以此类推. 令 T 为 30 个器件使用的总计时间, 求 T 超过 350 小时的概率.
9. 上题中的电子器件若每件为 a 元, 那么在年计划中一年至少需多少元才能以 95% 的概率保证够用 (假定一年有 306 个工作日, 每个工作日为 8 小时).
10. 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与独立, 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.
11. 设男孩出生率为 0.515, 求在 10000 个新生儿中女孩不少于男孩的概率?

12. 设有 1000 个人独立行动, 每个人能够按时进入掩蔽体的概率为 0.9. 以 95% 概率估计, 在一次行动中:

- (1) 至少有多少个人能够进入?
- (2) 至多有多少个人能够进入?

13. 在一定保险公司里有 10000 人参加保险, 每人每年付 12 元保险费, 在一年内一个人死亡的概率为 0.006, 死亡者其家属可向保险公司领得 1000 元赔偿费. 求:

- (1) 保险公司没有利润的概率为多大;
- (2) 保险公司一年的利润不少于 60000 元的概率为多大?

14. 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 试根据切比雪夫不等式给出 $P\{|X-Y| \geq 6\}$ 的估计. (2001 研考)

15. 某保险公司多年统计资料表明, 在索赔户中, 被盗索赔户占 20%, 以 X 表示在随机抽查的 100 个索赔户中, 因被盗向保险公司索赔的户数.

(1) 写出 X 的概率分布;

(2) 利用中心极限定理, 求被盗索赔户不少于 14 户且不多于 30 户的概率近似值.

16. 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 50kg, 标准差为 5kg, 若用最大载重量为 5t 的汽车承运, 试利用中心极限定理说明每辆车最多可以装多少箱, 才能保障不超载的概率大于 0.977. (2001 研考)