工科学生

概率论课外读物

叶尔骅 编

南京航空学院

1981.9.

目 录

一、加法定理

设 $A_1, A_2, \cdots\cdots, A_n$ 是随机事件，则有

$$P\left(\bigcup_{i=1}^{n} A_i\right)=\sum_{i_1=1}^{n} P(A_{i_1})-\sum_{1\leqslant i_1<i_2\leqslant n} P(A_{i_1}A_{i_2})+$$

$$+\sum_{1\leqslant i_1<i_2<i_3\leqslant n} P(A_{i_1}A_{i_2}A_{i_3})-+\cdots+(-1)^{n-1}P(A_1A_2\cdots A_n)。 \quad ①$$

记

$$S_k=\sum_{1\leqslant i_1<i_2<\cdots<i_k\leqslant n} P(A_{i_1}A_{i_2}\cdots A_{i_k}),\quad k=1,2,\cdots,n, \quad ②$$

①式可写为

$$P\left(\bigcup_{i=1}^{n} A_i\right)=S_1-S_2+S_3-S_4+-\cdots+(-1)^{n-1}S_n。 \text{———} \quad ③$$

用数学归纳法证明③式成立。

当 $n=2$ 时，显然成立。设 $n=k$ 时③式也成立，今证当 $n=k+1$
③式也成立，即证下式成立：

$$P\left(\bigcup_{i=1}^{k+1} A_i\right)=\tilde{S}_1-\tilde{S}_2+\tilde{S}_3-\tilde{S}_4+-\cdots+(-1)^{k}\tilde{S}_{k+1}, \text{———} \quad ④$$

中

$$\tilde{S}_l=\sum_{1\leqslant i_1<i_2<\cdots<i_l\leqslant k+1} P(A_{i_1}A_{i_2}\cdots A_{i_l}),\quad l=1,2,\cdots,k+1。 \quad ⑤$$

$$P\left(\bigcup_{i=1}^{k+1} A_i\right)=P\left[\left(\bigcup_{i=1}^{k} A_i\right)\cup A_{k+1}\right]$$

$$=P\left(\bigcup_{i=1}^{k} A_i\right)+P(A_{k+1})-P\left(\bigcup_{i=1}^{k} A_i A_{k+1}\right),$$

$$P\left(\bigcup_{i=1}^{k} A_i\right)=S_1-S_2+S_3-S_4+-\cdots+(-1)^{k-1}S_k,$$

其中

$$S_l=\sum_{1\leqslant i_1<i_2<\cdots<i_l\leqslant k} P(A_{i_1}A_{i_2}\cdots A_{i_l}),\quad l=1,2,\cdots,k。$$

$$P\left(\bigcup_{i=1}^{k} A_i A_{k+1}\right)=S_1'-S_2'+S_3'-S_4'+-\cdots+(-1)^{k-1}S_k',$$

其中

$$S_l'=\sum_{1\leqslant i_1<i_2<\cdots<i_l\leqslant k} P(A_{i_1}A_{i_2}\cdots A_{i_l}A_{k+1}),\ l=1,2,\cdots,k。$$

故有

$$\begin{aligned}P\left(\bigcup_{i=1}^{k+1} A_i\right)&=S_1-S_2+S_3-S_4+-\cdots+(-1)^{k-1}S_k+P(A_{k+1})\\&\quad-\left[S_1'-S_2'+S_3'-S_4'+-\cdots+(-1)^{k-1}S_k'\right]\\&=\left[S_1+P(A_{k+1})\right]-(S_2+S_1')+(S_3+S_2')-(S_4+S_3')\\&\quad+-\cdots+(-1)^{l-1}(S_l+S_{l-1}')+\cdots+(-1)^{k-1}(S_k+S_{k-1}')\\&\quad+(-1)^k S_k'\ 。\qquad\text{———} ⑥\end{aligned}$$

但 $S_1+P(A_{k+1})=\sum_{1\leqslant i_1\leqslant k} P(A_{i_1})+P(A_{k+1})=\sum_{1\leqslant i_1\leqslant k+1} P(A_{i_1})=\widetilde{S}_1,$

$$\begin{aligned}S_l+S_{l-1}'&=\sum_{1\leqslant i_1<i_2<\cdots<i_l\leqslant k} P(A_{i_1}A_{i_2}\cdots A_{i_l})+\\&\quad+\sum_{1\leqslant i_1<i_2<\cdots<i_{l-1}\leqslant k} P(A_{i_1}A_{i_2}\cdots A_{i_{l-1}}A_{k+1}),\end{aligned}$$

$$\widetilde{S}_l=\sum_{1\leqslant i_1<i_2<\cdots<i_l\leqslant k+1} P(A_{i_1}A_{i_2}\cdots A_{i_l})$$

（注意到 i_l 要么 $\leqslant k$；要么 $=k+1$，此时 $i_{l-1}\leqslant k$）

$$=\sum_{1\le i_1<i_2<\cdots<i_l\le k} P(A_{i_1}A_{i_2}\cdots A_{i_l})+$$

$$+\sum_{1\le i_1<i_2<\cdots<i_{l-1}\le k} P(A_{i_1}A_{i_2}\cdots A_{i_{l-1}}A_{k+1})$$

$$=S_l+S_{l-1}' \quad (l=2,3,\cdots,k),$$

$$(-1)^k S_k'=(-1)^k\sum_{1\le i_1<i_2<\cdots<i_k\le k} P(A_{i_1}A_{i_2}\cdots A_{i_k}A_{k+1})=$$

$$=(-1)^k P(A_1A_2\cdots A_kA_{k+1})$$

$$=(-1)^k\tilde{S}_{k+1}。$$

于是⑥式就变为

$$P(\bigcup_{i=1}^{k+1}A_i)=\tilde{S}_1-\tilde{S}_2+\tilde{S}_3-\tilde{S}_4+-\cdots+(-1)^k\tilde{S}_{k+1},$$

这就证得了④式成立，至此就证明了加法定理③式成立。

二、概率方法证明恒等式

用概率论的方法来证明一些恒等式或解决其它一些数学分析中的问题，是概率论的重要研究方向之一。这里仅举几例来说明。

[例 1]　证明：$C_n^r=C_{n-1}^{r-1}+C_{n-1}^r$　（n，r 为正整数且 $n\ge r$）。

证明：用初等代数证明上式是很容易的，也是大家所熟知的。现用概率方法证明。

设盒子中装有 n 个球（可辨），其中一个是白球，其余 $n-1$ 个是红球。从中任取 r 个球。记事件 A 为"取到白球"，则

$$P(A)+P(\bar{A})=1,$$

即

$$\frac{C_{n-1}^{r-1}}{C_n^r}+\frac{C_{n-1}^r}{C_n^r}=1,$$

故 $$C_n^r = C_{n-1}^{r-1} + C_n^r$$ 。

用概率方法证明这种恒等式的思路是首先构造一个概率模型，再利用事件之间的一些关系导出各个概率之间的关系，化简整理便可最终得到所要的结果。

［例2］ 证明：$(C_n^0)^2+(C_n^1)^2+(C_n^2)^2+\cdots+(C_n^n)^2=C_{2n}^n$ （n为正整数）。

证明： 设盒子中装有$2n$个球（可辨），其中n个为白球，其余n个为红球，从中任意取出n个球。记事件A_i为"抽到i个白球"（$i=0,1,2,\cdots,n$）．则

$$A_0 \cup A_1 \cup A_2 \cup \cdots \cup A_n = S \quad (S\text{为必然事件}),$$

且$\{A_i\}$两两互不相容，故由概率的有限可加性得：

$$P(A_0)+P(A_1)+P(A_2)+\cdots+P(A_n)=1,$$

但 $$P(A_i)=\frac{C_n^i C_n^{n-i}}{C_{2n}^n}=\frac{(C_n^i)^2}{C_{2n}^n} \quad (i=0,1,2,\cdots,n),$$

代入前式经整理得

$$(C_n^0)^2+(C_n^1)^2+(C_n^2)^2+\cdots+(C_n^n)^2=C_{2n}^n \quad .$$

［例3］ 证明：

$$1+\frac{n-r}{n-1}+\frac{(n-r)(n-r-1)}{(n-1)(n-2)}+\cdots+\frac{(n-r)(n-r-1)\cdots 2\cdot 1}{(n-1)(n-2)\cdots(r+1)r}=\frac{n}{r}$$ （n, r为正整数，$n>r$）。

证明：设盒子中装有n个球（可辨），其中r个为白球，任意地将球一个一个地取出，每次取后不放回，直到出现白球为止。设事件B_k为"第k次才取得白球"，则

$$\bigcup_{k=1}^{n-r+1} B_k = B_1 \cup B_2 \cup \cdots \cup B_{n-r+1} = S ,$$

又 $\{B_k\}$ 两两互不相容，故由概率有限可加性：

$$\sum_{k=1}^{n-r+1} P(B_k)=1 \text{。} \qquad (*)$$

但

$$P(B_1)=\frac{r}{n},$$

$$P(B_2)=P(\bar{B}_1B_2)=P(\bar{B}_1)P(B_2|\bar{B}_1)=\frac{n-r}{n}\cdot\frac{r}{n-1},$$

$$\vdots$$

$$P(B_k)=P(\bar{B}_1\bar{B}_2\cdots\bar{B}_{k-1}B_k)=P(\bar{B}_1)P(\bar{B}_2|\bar{B}_1)\cdots P(\bar{B}_{k-1}|\bar{B}_1\bar{B}_2\cdots\bar{B}_{k-2})\cdot$$

$$\cdot P(B_k|\bar{B}_1\bar{B}_2\cdots\bar{B}_{k-1})=\frac{n-r}{n}\cdot\frac{n-r-1}{n-1}\cdots\frac{n-r-(k-2)}{n-(k-2)}\cdot\frac{r}{n-(k-1)},$$

$$(k=2,3,\cdots,n-r+1)\text{。}$$

代入(*)式得

$$\frac{r}{n}+\frac{n-r}{n}\cdot\frac{r}{n-1}+\frac{n-r}{n}\cdot\frac{n-r-1}{n-1}\cdot\frac{r}{n-2}+\cdots+\frac{n-r}{n}\cdot\frac{n-r-1}{n-1}\cdots\frac{1}{r+1}\cdot\frac{r}{r}=1,$$

上式两端同乘 $\frac{n}{r}$ 得

$$1+\frac{n-r}{n-1}+\frac{(n-r)(n-r-1)}{(n-1)(n-2)}+\cdots+\frac{(n-r)(n-r-1)\cdots 2\cdot 1}{(n-1)(n-2)\cdots(r+1)r}=\frac{n}{r}\text{。}$$

三、几个概率分布

1. 几何分布

考虑贝努里试验，每次试验只有两个可能结果：A 和 $\bar{A}$ 记 $P(A)=p,(0<p<1)$，$P(\bar{A})=1-p=q$。设 X 为首次出现 A 所需的试验次数，则 X 为离散型随机变量且具有概率分布：

$$P\{X=k\}=q^{k-1}p, \qquad (k=1,2,\cdots)\text{。} \quad \text{——}(1.1)$$

称 X 服从参数为 p 的几何分布。

一种等待分布，这类概率在许多问题中会遇到。例如，若每次试验的可能结果为成功和失败，则首次成功所需的试验次数服从几何分布。在射击中，每次射击有两个可能结果：击中目标和不击中目标，则首先击中目标所需的射击次数也服从几何分布等等。

下面计算X的数学期望和方差。

$$EX=\sum_{k=0}^{\infty}kP\{X=k\}=\sum_{k=1}^{\infty}kq^{k-1}p=p\sum_{k=1}^{\infty}kq^{k-1}$$

$$=p\Big(\sum_{k=1}^{\infty}x^{k}\Big)'_{x=q}=p\Big(\frac{x}{1-x}\Big)'_{x=q}=p\cdot\frac{1}{(1-q)^2}=\frac{1}{p}\ . \quad \text{——} (1.2)$$

$$E(X^2)=\sum_{k=0}^{\infty}k^2P\{X=k\}=\sum_{k=1}^{\infty}k^2q^{k-1}p=pq\sum_{k=2}^{\infty}k(k-1)q^{k-2}+$$

$$+\sum_{k=1}^{\infty}kq^{k-1}p=pq\Big(\sum_{k=1}^{\infty}x^{k}\Big)''_{x=q}+\frac{1}{p}$$

$$=Rq\Big(\frac{x}{1-x}\Big)''_{x=q}+\frac{1}{p}=pq\cdot\frac{2}{(1-q)^3}+\frac{1}{p}=\frac{2q}{p^2}+\frac{1}{p},$$

$$DX=E(X^2)-(EX)^2=\frac{2q}{p^2}+\frac{1}{p}-\frac{1}{p^2}=\frac{q}{p^2}\ . \quad \text{——} (1.3)$$

几何分布有一个很重要的性质，叫做"无记忆性"。为了阐明这个性质，先计算条件概率

$$P\{X=m+l\,|\,X>m\}=\frac{P\{X>m,\,X=m+l\}}{P\{X>m\}}=\frac{P\{X=m+l\}}{P\{X>m\}}$$

$$=\frac{q^{m+l-1}p}{\sum_{k=m+1}^{\infty}q^{k-1}p}=\frac{q^{m+l-1}p}{\frac{q^{m}}{q}p}=q^{l-1}p=P\{X=l\}\quad(m,l\text{为正整数})。$$

$$\text{——} (1.4)$$

这说明如果已知在前m次试验中没有出现成功，那么为了达到首次成功再需要做l次试验的概率与前面的失败次数m无关。形象地说就是把过去的经历完全忘记了，这个性质就是"无记忆

性”。但是更有意思的是，在离散型随机变量的概率分布中，也只有几何分布才具有这样一种特性。下面我们来证明这个事实。

设 X 是取正整数值的随机变量，并且具有无记忆性，即下式成立

$$P\{X=m+l \mid X>m\}=P\{X=l\}, \qquad (1.5)$$

要证明 X 服从几何分布。因为

$$P\{X=k+1 \mid X>k\}=P\{X=1\}, \qquad (1.6)$$

记

$$p=P\{X=1\}, \quad p_k=P\{X=k\}, \quad q_k=P\{X>k\}, \qquad (1.7)$$

则
$$p_{k+1}=P\{X=k+1\}=P\{X>k\}-P\{X>k+1\}=q_k-q_{k+1}, \qquad (1.8)$$

再由（1.6）、（1.7）式得

$$p=P\{X=1\}=P\{X=k+1 \mid X>k\}=\frac{P\{X>k, X=k+1\}}{P\{X>k\}}$$

$$=\frac{P\{X=k+1\}}{P\{X>k\}}=\frac{p_{k+1}}{q_k}=\frac{q_k-q_{k+1}}{q_k}=1-\frac{q_{k+1}}{q_k},$$

所以

$$\frac{q_{k+1}}{q_k}=1=p \quad 。 \qquad (1.9)$$

注意到 $q_0=1$，故由（1.9）式得

$$q_k=(1-p)^k,$$

再由（1.8）式得

$$p_k=P\{X=k\}=(1-p)^{k-1}p, \quad (k=1, 2, \cdots)。 \qquad (1.10)$$

这就证明了 X 服从参数为 p 的几何分布。

2. 负二项分布［巴斯卡（Pascal）分布］

设贝努里试验进行到第 r 次成功为止，每次试验中成功的概率为 p（$0<p<1$），$q=1-p$。令 X 为试验进行的次数。则 X 为离散型随机变量。因事件“$X=k$”等价于“第 k 次试验时出现成功，且在前 $k-1$ 次试验中成功 $r-1$ 次”，故 X 的概率分布为

$$P\{X=k\}=C_{k-1}^{r-1}p^{r-1}q^{k-r}\cdot p=C_{k-1}^{r-1}p^{r}q^{k-r},\quad k=r,r+1,\cdots。 \qquad (2.1)$$

称 X 服从参数为 r，p 的负二项分布，也叫巴斯卡分布。当 $r=1$ 时，就是几何分布。

下面来计算 X 的数学期望和方差。

方法Ⅰ：注意到（2.1）式给出了 X 的概率分布，故得恒等式

$$\sum_{k=r}^{\infty}C_{k-1}^{r-1}p^{r}q^{k-r}=1 \quad 。 \qquad (2.2)$$

（注：上式也可从 $(1-q)^{-r}$ 对 q 展开推得）

（2.2）式中以 $r+1$ 代 r，$k+1$ 代 k 得

$$\sum_{k=r}^{\infty}C_{k}^{r}p^{r+1}q^{k-r}=1 \quad 。 \qquad (2.3)$$

再在（2.3）式中以 $r+1$ 代 r，$k+1$ 代 k 得

$$\sum_{k=r}^{\infty}C_{k+1}^{r+1}p^{r+2}q^{k-r}=1 \quad 。 \qquad (2.4)$$

利用这些等式便可计算 X 的数学期望和方差。

$$EX=\sum_{k=r}^{\infty}kC_{k-1}^{r-1}p^{r}q^{k-r}=\sum_{k=r}^{\infty}\frac{k(k-1)!}{(r-1)!(k-r)!}p^{r}q^{k-r}$$

$$=\frac{r}{p}\sum_{k=r}^{\infty}\frac{k!}{r!(k-r)!}p^{r+1}q^{k-r}=\frac{r}{p}\sum_{k=r}^{\infty}C_{k}^{r}p^{r+1}q^{k-r}=\frac{r}{p} \qquad (2.5)$$

$$E(X^{2})=\sum_{k=r}^{\infty}k^{2}C_{k-1}^{r-1}p^{r}q^{k-r}=\sum_{k=r}^{\infty}\left[k(k+1)-k\right]C_{k-1}^{r-1}p^{r}q^{k-r}$$

$$=\sum_{k=r}^{\infty}k(k+1)\frac{(k-1)!}{(r-1)!(k-r)!}p^{r}q^{k-r}-EX$$

$$=\frac{r(r+1)}{p^2}\sum_{k=r}^{\infty}\frac{(k+1)!}{(r+1)![(k+1)-(r+1)]!}p^{r+2}q^{k-r}-\frac{r}{p}$$

$$=\frac{r(r+1)}{p^2}\sum_{k=r}^{\infty}C_{k+1}^{r+1}p^{r+2}q^{k-r}-\frac{r}{p}$$

$$=\frac{r(r+1)}{p^2}-\frac{r}{p}\ ,$$

$$DX=E(X^2)-(EX)^2=\frac{r(r+1)}{p^2}-\frac{r}{p}-\frac{r^2}{p^2}=\frac{rq}{p^2}\ 。\qquad\text{——— (2.6)}$$

方法Ⅱ：记 X_1 为第 1 次出现成功所需的试验次数，X_2 为第 1 次成功后到第 2 次成功之间的试验次数，……，X_r 为第 $r-1$ 次成功后到第 r 次成功之间的试验次数，则 X_1，X_2，……，X_r 为相互独立的随机变量，且都服从有相同参数 p 的几何分布，而

$$X=X_1+X_2+\cdots\cdots+X_r\ ,$$

$$\therefore\qquad EX=\sum_{i=1}^{r}EX_i=\frac{r}{p}\ ,\qquad\text{——— (2.7)}$$

$$DX=\sum_{i=1}^{r}DX_i=\frac{rq}{p^2}\ 。\qquad\text{——— (2.8)}$$

方法Ⅱ简便得多，这种把一个随机变量 X 化为几个相互独立的随机变量之和来计算 X 的数字特征的方法在概率论中是很有用的。

负二项分布和二项分布有一定的联系，即所谓的对偶关系，这种关系表明这两种分布可以互为表示。

仍考虑贝努里试验，假定 p 固定，设 X 服从参数为 r，p 的负二项分布，Y 服从参数为 n，p 的二项分布。记

$$F_{nb}(n|r,p)=P\{X\leqslant n\}\ ,\qquad\text{——— (2.9)}$$

$$F_b(r|n,p)=P\{Y\leqslant r\}\ ,\qquad\text{——— (2.10)}$$

则因为

$$P\{r\text{次成功所需的试验次数}\leqslant n\}=P\{\text{在}n\text{次试验中，成功次数}\geqslant r\},$$

所以得

$$F_{nb}(n|r,p)=1-F_b(r-1|n,p)$$

或

$$F_{nb}(n|r+1,p)=1-F_b(r|n,p) \qquad \text{——— (2.11)}$$

上式表明对于负二项分布的概率计算可以转化为二项分布来计算．

负二项分布在抽样检验，可靠性等问题中应用广泛。

3，超几何分布

一批产品共有 N 件，其中有 M 件次品（ $M<N$ ）．现从中无放回地任取 n 件（不失一般性假定 $n\leqslant N-M$），记 X 为这 n 件产品中所含有的次品件数，则 X 为离散型随机变量，其概率分布为

$$P\{X=k\}=\frac{C_M^k C_{N-M}^{n-k}}{C_N^n}, \quad k=0,1,2,\cdots,l,\ l=\min(M,n)。 \qquad \text{——— (3.1)}$$

称 X 服从超几何分布．

下面计算 X 的数学期望和方差．

方法Ⅰ：由于（3.1）式给出了概率分布，故有

$$\sum_{k=0}^{l}\frac{C_M^k\, C_{N-M}^{n-k}}{C_N^n}=1 \quad 。 \qquad \text{——— (3.2)}$$

在求 EX 与 DX 时要用到这个恒等式。

$$EX=\sum_{k=0}^{l}k\cdot\frac{C_M^k C_{N-M}^{n-k}}{C_N^n}=\sum_{k=1}^{l}\frac{C_{M-1}^{k-1}C_{N-M}^{n-k}\cdot M}{C_{N-1}^{n-1}\cdot\frac{N}{n}} \qquad (\text{令 }k-1=i)$$

$$=\frac{nM}{N}\sum_{i=0}^{l-1}\frac{C_{M-1}^{i}C_{(N-1)-(M-1)}^{n-1-i}}{C_{N-1}^{n-1}}=\frac{nM}{N} \quad 。 \qquad \text{——— (3.3)}$$

上式最后一步利用了（3.2）式的结果。

$$E(X^2)=\sum_{k=0}^{l} k^2\cdot\frac{C_M^k C_{N-M}^{n-k}}{C_N^n}=\sum_{k=1}^{l}\left[k(k-1)+k\right]\frac{C_M^k C_{N-M}^{n-k}}{C_N^n}$$

$$=\sum_{k=2}^{l} k(k-1)\frac{C_M^k C_{N-M}^{n-k}}{C_N^n}+EX$$

$$=\sum_{k=2}^{l}\frac{C_{M-2}^{k-2} C_{N-M}^{n-k}}{C_{N-2}^{n-2}}\cdot\frac{M(M-1)n(n-1)}{N(N-1)}+\frac{nM}{N} \qquad (\text{令 } k-2=i)$$

$$=\frac{M(M-1)n(n-1)}{N(N-1)}\sum_{i=0}^{l-2}\frac{C_{M-2}^{i} C_{(N-2)-(M-2)}^{n-2-i}}{C_{N-2}^{n-2}}+\frac{nM}{N}$$

$$=\frac{nM(n-1)(M-1)}{N(N-1)}+\frac{nM}{N},$$

$$DX=E(X^2)-(EX)^2=\frac{nM(n-1)(M-1)}{N(N-1)}+\frac{nM}{N}-\frac{n^2M^2}{N^2}$$

$$=\frac{nM(N-n)(N-M)}{N^2(N-1)} \qquad \text{——— (3.4)}$$

方法Ⅱ：先看一个具体例子．

［例1］设有100件产品，其中有5件次品，95件正品，现从中任取20件，求抽到的次品件数X的数学期望和方差。

解：显然X服从超几何分布，其概率分布为

$$P\{X=k\}=\frac{C_5^k C_{95}^{20-k}}{C_{100}^{20}}, \quad k=0,1,2,\cdots,5 \text{。} \qquad \text{—— (3.5)}$$

$$EX=\sum_{k=0}^{5} k\cdot\frac{C_5^k C_{95}^{20-k}}{C_{100}^{20}}=\frac{1}{C_{100}^{20}}\sum_{k=1}^{5} k C_5^k C_{95}^{20-k} \text{。} \qquad \text{—— (3.6)}$$

关键要求上百的和式。因为

$$(1+x)^5(1+x)^{95}=\cdots\cdots+\left(\sum_{k=0}^{5} C_5^k C_{95}^{20-k}\right)x^{20}+\cdots\cdots,$$

$$(1+tx)^5(1+x)^{95}=\cdots\cdots+\left(\sum_{k=0}^{5} t^k C_5^k C_{95}^{20-k}\right)x^{20}+\cdots\cdots,$$

$$\frac{\partial}{\partial t}\left[(1+tx)^5(1+x)^{95}\right]=\cdots\cdots+\left(\sum_{k=1}^{5} k t^{k-1} C_5^k C_{95}^{20-k}\right)x^{20}+\cdots\cdots,$$

$$\frac{\partial}{\partial t}\left[(1+tx)^5(1+x)^{95}\right]_{t=1}=\cdots\cdots+\left(\sum_{k=1}^{5} k C_5^k C_{95}^{20-k}\right)x^{20}+\cdots\cdots,$$

所以

$$\sum_{k=1}^{5} k C_5^k C_{95}^{20-k}=\frac{\partial}{\partial t}\left[(1+tx)^5(1+x)^{95}\right]_{t=1}\text{ 展开式中 } x^{20} \text{ 的系数}$$

$$=5x(1+x)^{99}\text{ 展开式中 } x^{20} \text{ 的系数}$$

$$=5C_{99}^{19}=C_{100}^{20}, \qquad \text{——— (3.7)}$$

代入（3.6）式得

$$EX=\frac{C_{100}^{20}}{C_{100}^{20}}=1\text{。} \qquad \text{——— (3.8)}$$

为求 DX，先求 $E(X^2)$。

$$E(X^2)=\sum_{k=0}^{5} k^2\cdot\frac{C_5^k C_{95}^{20-k}}{C_{100}^{20}}=\frac{1}{C_{100}^{20}}\sum_{k=1}^{5} k^2 C_5^k C_{95}^{20-k}\text{。} \qquad \text{——— (3.9)}$$

与上面类似地有

$$\frac{\partial^2}{\partial t^2}\left[(1+tx)^5(1+x)^{95}\right]=\cdots\cdots+\left(\sum_{k=2}^{5} k(k-1)t^{k-2} C_5^k C_{95}^{20-k}\right)x^{20}+\cdots$$

$$\frac{\partial^2}{\partial t^2}\left[(1+tx)^5(1+x)^{95}\right]_{t=1}=\cdots\cdots+\left(\sum_{k=2}^{5} k(k-1) C_5^k C_{95}^{20-k}\right)x^{20}+\cdots\cdots$$

所以，注意到（3.7）式有

$$\sum_{k=1}^{5} k^2 C_5^k C_{95}^{20-k}=\sum_{k=2}^{5} k(k-1) C_5^k C_{95}^{20-k}+\sum_{k=1}^{5} k C_5^k C_{95}^{20-k}$$

$$=\frac{\partial^2}{\partial t^2}\left[(1+tx)^5(1+x)^{95}\right]_{t=1}\text{ 展开式中 } x^{20} \text{ 的系数}+C_{100}^{20}$$

$$=20x^2(1+x)^{98}\text{ 展开式中 }x^{20}\text{ 的系数}+C_{100}^{20}$$

$$=20\,C_{98}^{18}+C_{100}^{20}=\frac{76}{99}\times C_{100}^{20}+C_{100}^{20}\ , \qquad (3.10)$$

代入（3.9）式得

$$E(X^2)=\frac{1}{C_{100}^{20}}\left(\frac{76}{99}\times C_{100}^{20}+C_{100}^{20}\right)=\frac{76}{99}+1\ , \qquad (3.11)$$

所以

$$DX=E(X^2)-(EX)^2=\frac{76}{99}+1-1=\frac{76}{99}=0.768\ . \qquad (3.12)$$

上述计算过程中的要领是把计算和式 $\sum_{k=0}^{5}k\,C_5^k C_{95}^{20-k}$ 与 $\sum_{k=0}^{5}k^2 C_5^k C_{95}^{20-k}$ 归结为计算某个多项式的系数，这种方法在概率论中也常用。

现在应用这种方法来计算超几何分布（3.1）的数学期望和方差。不失一般性，设 $M\leqslant n$，故 $l=M$。而

$$EX=\sum_{k=0}^{M}k\cdot\frac{C_M^k C_{N-M}^{n-k}}{C_N^n}=\frac{1}{C_N^n}\sum_{k=1}^{M}k\,C_M^k C_{N-M}^{n-k}$$

$$=\frac{1}{C_N^n}\cdot\frac{\partial}{\partial t}\left[(1+tx)^M(1+x)^{N-M}\right]_{t=1}\text{ 展开式中 }x^n\text{ 的系数}$$

$$=\frac{1}{C_N^n}\cdot Mx(1+x)^{N-1}\text{ 中 }x^n\text{ 的系数}$$

$$=\frac{1}{C_N^n}\cdot M\,C_{N-1}^{n-1}=\frac{nM}{N} \qquad (3.13)$$

$$E(X^2)=\sum_{k=0}^{M}k^2\cdot\frac{C_M^k C_{N-M}^{n-k}}{C_N^n}=\frac{1}{C_N^n}\sum_{k=1}^{M}k^2 C_M^k C_{N-M}^{n-k}$$

$$=\frac{1}{C_N^n}\left[\sum_{k=2}^{M}k(k-1)C_M^k C_{N-M}^{n-k}+\sum_{k=1}^{M}k\,C_M^k C_{N-M}^{n-k}\right]$$

$$= \frac{1}{C_N^n}\left\{\frac{\partial^2}{\partial t^2}\left[(1+tx)^M(1+x)^{N-M}\right]_{t=1}+\frac{\partial}{\partial t}\left[(1+tx)^M(1+x)^{N-M}\right]\right.$$

展开式中 x^n 的系数

$$= \frac{1}{C_N^n}\left[M(M-1)x^2(1+x)^{N-2}+Mx(1+x)^{N-1}\right] \text{ 展开式中 } x^n \text{ 的系数}$$

$$= \frac{1}{C_N^n}\left[M(M-1)C_{N-2}^{n-2}+MC_{N-1}^{n-1}\right]$$

$$= \frac{nM(n-1)(M-1)}{N(N-1)}+\frac{nM}{N}, \qquad (3.14$$

所以

$$DX = E(X^2)-(EX)^2 = \frac{nM(n-1)(M-1)}{N(N-1)}+\frac{nM}{N}-\frac{n^2M^2}{N^2}$$

$$= \frac{nM(N-n)(N-M)}{N^2(N-1)}. \qquad (3.1$$

可见超几何分布的数字特征计算是比较麻烦的。

应当指出，导出超几何分布时，我们是采用了无放回抽样。如果采用放回抽样，那么就是一个贝努里试验，这时次品件数X的概率分布为二项分布，即

$$P\{X=k\}=C_n^k p^k(1-p)^{n-k}=C_n^k\left(\frac{M}{N}\right)^k\left(1-\frac{M}{N}\right)^{n-k}, \quad k=0,1,2,\cdots \quad l=\min(n,M)$$

$$(3.16$$

其中 $p=P(A)=\frac{M}{N}$，A为事件"每次抽到次品"。这说明，随着抽样方式的改变，会引起X的概率分布的改变。但是从直观上看，当N很大时，可以认为两种抽样方式所对应的结果差别应很小，即超几何分布与二项分布差别很小。这个直观的认识确实可以从理论上加以阐明。可以证明：当 $N\to\infty$，$\frac{M}{N}\to p$（n，k 不变）时，超几何分布的极限分布为二项分布，即

$$\lim_{N\to\infty}\frac{C_M^k C_{N-M}^{n-k}}{C_N^n}=C_n^k p^k(1-p)^{n-k}. \qquad (3.1$$

下面证明（3.17）式。

$$\frac{C_M^k C_{N-M}^{n-k}}{C_N^n}=\frac{M!}{k!(M-k)!}\cdot\frac{(N-M)!}{(n-k)![N-M-(n-k)]!}\cdot\frac{n!(N-n)!}{N!}$$

$$=\frac{n!}{k!(n-k)!}\cdot\left[\frac{M(M-1)\cdots(M-k+1)}{\underbrace{N\cdot N\cdots N}_{k个}}\right]\cdot\left[\frac{(N-M)(N-M-1)\cdots(N-M-(n-k)+1)}{\underbrace{N\cdot N\cdots N}_{(n-k)个}}\right]\cdot$$

$$\cdot\left[\frac{\overbrace{N\cdot N\cdots N}^{n个}}{N(N-1)\cdots(N-n+1)}\right],$$

当 $N\to\infty$，$\frac{M}{N}\to p$ 时，不难看出：第一个方括号 $\longrightarrow p^k$；第二个方括号 $\longrightarrow(1-p)^{n-k}$；第三个方括号 $\longrightarrow 1$，因此（3.17）式成立。

4．Γ－分布（Gamma分布）

若连续型随机变量 X 的概率密度为

$$p(x)=\begin{cases}\frac{\lambda^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\lambda x}, & x>0\\ 0, & x\leqslant 0\end{cases}\quad(\alpha>0,\lambda>0 之常数)——(4.1)$$

则称 X 服从参数为 λ，α 的Γ－分布。这里 Γ(α) 是参数为 α 的Γ－函数，即

$$\Gamma(\alpha)=\int_0^\infty t^{\alpha-1}e^{-t}dt\quad(\alpha>0),\qquad ——(4.2)$$

Γ－函数有下列简单性质：

$$\Gamma(\alpha+1)=\alpha\Gamma(\alpha),\quad\Gamma(n+1)=n!\qquad ——(4.3$$

当 $\alpha=1$ 时，因 $\Gamma(1)=1$，故 X 的分布退化为参数为 λ 的指数分布，这表明参数为 λ 的指数分布可看参数为 λ，$\alpha=1$ 的Γ－分布。

下面计算 X 的均值和方差。