第五章　母函数与特征函数及极限定理

§1　母 函 数

1. 数列的母函数

对于任一数列 $a_0,a_1,\cdots,a_n,\cdots$,可以引进一函数:

$$g(z)\equiv\sum_{n=0}^{\infty}a_nz^n.$$

它称为该数列$\{a_n,n\geqslant0\}$的母函数。

比如对组合数数列 $C_n^0,C_n^1,\cdots,C_n^n$,其母函数为

$$g(z)=\sum_{k=0}^{n}C_n^kz^k=(1+z)^n.$$

又如对可重复组合数列 $RC_n^0,RC_n^1,\cdots,RC_n^m$,…(记号 RC_n^m 表示从 n 个元素中取 m 个的可重复组合总数),其母函数应该是

$$g(x)=(1+x+x^2+\cdots)^n$$

(可利用第一章(2.4)式给出的 $RC_n^m=C_{n+m-1}^m$ 验证；更具启发性的方法是，从 RC_n^m 的定义直接看出$(1+x+x^2+\cdots)^n$ 展开式中 x^m 的系数就应该是 RC_n^m 。为此，可以先分析$(1+x+x^2+\cdots+x^m)^n$ 展开式中 x^m 的系数)。

2. 非负整值随机变量的母函数

设X为非负整值随机变量，它的概率分布列为

$$P(X=k)=p_k,\quad k=0,1,2,\cdots,$$

其相应的母函数为

$$g(z)=\sum_{k=0}^{\infty}p_kz^k. \tag{1.1}$$

称为X的概率母函数(简称母函数)。

由于 $p_k\geqslant0,k=0,1,2,\cdots$ 且$\sum p_k=1$，因此，(1.1)右端的幂级数的收敛半径$\geqslant1$。

引进母函数的好处，在于它有很好的分析性质(由数学分析课程中对幂级数的讨论知，它至少在$(-1,1)$中无穷次可微)，而且一旦知道了X的母函数，那么X的分布列就找出来了：

$$p_k=\frac{g^{(k)}(0)}{k!},\quad k=0,1,\cdots.$$

注　若已知X的概率母函数 $g(z)$，我们不一定按上式给的方式来找出分布列$\{p_k,k=0,1,\cdots\}$，而直接将 $g(z)$在 $z=0$ 附近 Taylor 展开即可。上式的意义在于：由分布列到母函数的对应是一一对应的。

3. 几个常见分布的母函数

(1) 二项分布

设 $X\sim B(n,p)$，则它的母函数为

$$\begin{aligned}g(z)&=\sum_{k=0}^{n}C_n^kp^kq^{n-k}z^k\\&=(q+pz)^n.\end{aligned}$$

(2) 泊松分布

设 $X\sim P(\lambda)$，则它的母函数为

$$g(z)=\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}e^{-\lambda}z^k=e^{\lambda(z-1)}. \tag{1.2}$$

(3) 几何分布

设X服从参数为p的几何分布，即

$$P(X=k)=q^{k-1}p, \quad k=1,2,\cdots,$$

则它的母函数

$$g(z)=\sum_{k=1}^{\infty}pq^{k-1}z^k=\frac{pz}{1-qz}. \tag{1.3}$$

4. 母函数性质

定理 设非负整值随机变量$X_1,X_2,\cdots,X_n$相互独立，而$g_1,g_2,\cdots,g_n$分别为它们的母函数，则$Y=\sum\limits_1^n X_k$的母函数

$$g(z)=g_1(z)g_2(z)\cdots g_n(z). \tag{1.4}$$

证 显然，只需证$n=2$的情形．设X_1,X_2的母函数分别为

$$g_1(z)=\sum_{k=0}^{\infty}a_kz^k,$$

$$g_2(z)=\sum_{k=0}^{\infty}b_kz^k.$$

则$Y=X_1+X_2$的概率分布列

$$\begin{aligned}P(Y=r)&=P(X_1+X_2=r)\\&=\sum_{k=0}^{r}P(X_1=k,X_1+X_2=r)\\&=\sum_{k=0}^{r}P(X_1=k,X_2=r-k)\\&=\sum_{k=0}^{r}P(X_1=k)P(X_2=r-k)\\&=\sum_{k=0}^{r}a_kb_{r-k}.\end{aligned}$$

因此，Y的母函数为

$$\begin{aligned}\sum_{r=0}^{\infty}P(Y=r)z^r&=\sum_{r=0}^{\infty}\left(\sum_{k=0}^{r}a_kb_{r-k}\right)z^r\\&=g_1(z)g_2(z).\end{aligned}$$

该定理为寻找独立随机变量之和的概率分布带来极大的方便．

例1.1 设$X_1,\cdots,X_n$独立同分布，

$$P(X_1=1)=p, \quad P(X_1=0)=q,$$

求$Y=\sum\limits_1^n X_k$的分布．

解 首先，X_1的母函数为

$$g_1(z)=q+pz,$$

由于$X_1,\cdots,X_n$独立同分布，由定理知，$Y=\sum X_k$的母函数为

$$(q+pz)^n.$$

由此可得Y的概率分布列．

注 这给出了成功次数服从二项分布的另一个证明，而且是“机器处理”的．

例1.2 设$T_1,\cdots,T_n$是相互独立的几何分布(参数同为p)，求$S_n=\sum\limits_1^n T_k$的分布．

解 T_1的母函数为

$$g_1(z)=\frac{pz}{1-qz},$$

于是，按定理知，S_n的母函数为

$$\left(\frac{pz}{1-qz}\right)^n.$$

下面，我们将它展为幂级数，从而找出S_n的分布列．

$$\left(\frac{pz}{1-qz}\right)^n = p^n z^n (1-qz)^{-n}.$$

$$= p^n z^n \sum_{k=0}^{\infty} \frac{n(n+1)\cdots(n+k-1)}{k!}(qz)^k$$

$$= \sum_{k=0}^{\infty}\binom{n+k-1}{k} p^n q^k z^{n+k}$$

$$= \sum_{k=0}^{\infty}\binom{n+k-1}{n-1} p^n q^{(n+k)-n} z^{n+k}$$

$$= \sum_{k=n}^{\infty}\binom{k-1}{n-1} p^n q^{k-n} z^k.$$

这表明 S_n 的分布列为

$$P(S_n = k) = \binom{k-1}{n-1} p^n q^{k-n}, \quad k = n, n+1, \cdots.$$

注 回顾第二章的(2.11)式可知，这里的 S_n，可看作在独立重复试验中直到事件 A 发生 n 次时所需的试验总数，而我们用母函数的方法再一次得到了它的分布列(参见第二章(2.9)式).

例1.3 掷三颗骰子，求总点数为9的概率.

解 设 X_1, X_2, X_3 分别为编号为1,2,3的三颗骰子的点数，则总点数 Y 为

$$Y = X_1 + X_2 + X_3.$$

按定义，X_1 的母函数为

$$g(z) = \frac{1}{6}(z + z^2 + \cdots + z^6)$$

$$= \frac{1}{6}\frac{z(1-z^6)}{1-z},$$

所以，Y 的母函数为

$$(g(z))^3 = \frac{z^3(1-z^6)^3}{6^3(1-z)^3}$$

$$= \frac{z^3}{6^3}(1 - 3z^6 + 3z^{12} - z^{18}) \cdot \sum_{k=0}^{\infty}\binom{k+2}{2} z^k.$$

容易看出上式右端中 z^9 的系数为

$$\frac{1}{6^3}\left(\binom{6+2}{2} - 3\right) = \frac{25}{216},$$

它就是 $P(Y=9)$.

说明 在这里使用的公式

$$(1-z)^{-n} = \sum_{k=0}^{\infty}\frac{n(n+1)\cdots(n+k-1)}{k!} z^k$$

$$= \sum_{k=0}^{\infty}\binom{n+k-1}{n-1} z^k.$$

在本书中多次出现.

5. 其他

(1) 母函数与数字特征的关系

定理 母函数与期望和方差间有两个直接的关系式（设 X 的母函数为 $g(z)$）：

$$E(X) = g'(1), \tag{1.5}$$

$$\mathrm{Var}(X) = g''(1) + g'(1) - (g'(1))^2. \tag{1.6}$$

证 首先有

$$g(z) = \sum_{k=0}^{\infty} p_k z^k.$$

两边求微商，得

$$g'(z) = \sum_{k=1}^{\infty} k p_k z^{k-1}.$$

因而

$$g'(1)=\sum_{k=1}^{\infty}kp_k,$$

此即(1.5)。而(1.6)的证明是类似的，留给读者。

说明 上面的证明过程中，有一个逐项求微商的问题。由数学分析课程中关于幂级数的讨论可知，如果 $\sum_{k=0}^{\infty}p_kz^k$ 的收敛半径 $R>1$，一切畅通无阻；在收敛半径 $R=1$ 时，只需 $\sum_{k=1}^{\infty}kp_k<+\infty$ 即可，此即 $E(X)<+\infty$。对(1.6)式也可类似讨论。

(2) 母函数的期望表示

对于我们讨论的非负整值随机变量的概率母函数

$$g(z)=\sum_{k=0}^{\infty}p_kz^k,$$

由随机变量函数的期望公式(2.1′)，有

$$E(z^X)=\sum_{k=0}^{\infty}z^kp_k=g(z),\tag{1.7}$$

即母函数 $g(z)$ 实际上是 z^X 的期望。

利用(1.7)很容易证明(1.4)式。这是因为Y的母函数

$$\begin{aligned}E(z^Y)&=E(z^{X_1+\cdots+X_n})\\&=E(z^{X_1}\cdots z^{X_n}).\end{aligned}$$

由于 $X_1,\cdots,X_n$ 相互独立，所以 $z^{X_1},\cdots,z^{X_n}$ 相互独立，于是由期望性质得

$$\begin{aligned}E(z^Y)&=E(z^{X_1})\cdots E(z^{X_n})\\&=g_1(z)\cdots g_n(z).\end{aligned}$$

例1.4 在一次核反应中，某个粒子可能分裂为 2 或 3 个粒子或不分裂，这三种可能性相应的概率分别为 p_2,p_3 与 p_1；新粒

子的形态是相同的(与老粒子也相同)，并且行为彼此独立。求两次反应以后粒子总数的分布。

解 做过习题的读者会发现，这就是第二章习题 9。这里用母函数与数学期望来讨论。

记 X_1 为一次反应后的粒子数，X_2 为两次反应后的粒子数。于是按题设，X_1 的母函数

$$E(z^{X_1})=g(z)=p_1z+p_2z^2+p_3z^3.$$

下面我们来求 X_2 的母函数 $E(z^{X_2})$。

注意到两次反应后的粒子总数，是由一次反应后的各个粒子分别独立的分裂(或不分裂)产生粒子数的总和，所以有分解式

$$X_2=Y_1+\cdots+Y_{X_1}.$$

这里 Y_1,Y_2,Y_3 分别表示第一次反应后的第 1,2,3 个粒子分裂产生的粒子数。

注意，Y_1,Y_2,Y_3 应与 X_1 有相同的分布，因此有相同形式的母函数；而且它们之间是相互独立的。于是，X_2 的母函数

$$E(z^{X_2})=E(z^{Y_1+\cdots+Y_{X_1}}).$$

注意，由于这里的 X_1 是随机变量，通常对常量适用的公式不再成立。因此我们用条件期望，具体说，用上章(5.6)式，

$$E(z^{X_2})=E(E(z^{Y_1+\cdots+Y_{X_1}}\mid X_1)).$$

对于内层 $E(z^{Y_1+\cdots+Y_{X_1}}\mid X_1)$，我们先求

$$\begin{aligned}E(z^{Y_1+\cdots+Y_{X_1}}\mid X_1=k)&=E(z^{Y_1+\cdots+Y_k}\mid X_1=k)\\&=E(z^{Y_1+\cdots+Y_k})\\&=E(z^{Y_1})\cdots E(z^{Y_k})\\&=(g(z))^k,\end{aligned}$$

于是，按条件期望定义((5.2))，

$$E(z^{Y_1+\cdots+Y_{X_1}}\mid X_1)=(g(z))^{X_1}.$$

因此有

$$E(z^{X_2}) = E((g(z))^{X_1}).$$

由于

$$E(z^{X_1}) = g(z),$$

所以 X_2 的母函数

$$E(z^{X_2}) = g(g(z)). \tag{1.8}$$

注 (i) 由于 $g(z) = p_1z + p_2z^2 + p_3z^3$，因此

$$g(g(z)) = p_1(p_1z + p_2z^2 + p_3z^3) + p_2(p_1z + p_2z^2 + p_3z^3)^2 + p_3(p_1z + p_2z^2 + p_3z^3)^3.$$

(ii) 用本例的方法不难导出 m 次反应后粒子总数的母函数的一般公式。

§2 特征函数

概率母函数为非负整值随机变量的研究带来很大方便，对于一般的随机变量是否有类似的东西呢？这就是特征函数。

1. 定义

对任一随机变量 X，称

$$f(t) \equiv E(e^{itX}), \quad -\infty < t < +\infty \tag{2.1}$$

为 X 的特征函数。

说明 (i) 规定

$$E(\xi + i\eta) \equiv E(\xi) + iE(\eta),$$

因此有

$$E(e^{itX}) = E(\cos tX + i\sin tX) = E(\cos tX) + iE(\sin tX).$$

(ii) 由于

$$|e^{itX}| \leqslant 1$$

(或者说，由于 $|\cos tX| \leqslant 1$，$|\sin tX| \leqslant 1$)，因此，对任一随机变量 X，都有特征函数。

(iii) 特征函数 $f(t)$ 是实变复值的。

(iv) 若 X 的特征函数为 $f(t)$，则容易看出 $a + bX$ 的特征函数为

$$e^{ita}f(bt). \tag{2.2}$$

(v) $f(0) = 1$. (2.3)

2. 几种常见分布的特征函数

(1) 二点分布

按定义式(2.1)，二点分布的特征函数

$$f(t) = pe^{it} + q.$$

(2) 泊松分布

$$f(t) = \sum_{k=0}^{\infty} e^{itk} \cdot \frac{\lambda^k}{k!}e^{-\lambda} = \sum_{k=0}^{\infty} \frac{(\lambda e^{it})^k}{k!}e^{-\lambda} = e^{\lambda(e^{it}-1)}.$$

(3) 正态分布

先求标准正态分布的特征函数。

因为 $\int_{-\infty}^{\infty} \sin tx \cdot \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx = 0$，所以

$$f(t) = \int_{-\infty}^{\infty} e^{itx} \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx = \int_{-\infty}^{\infty} \cos tx \cdot \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx.$$

为找出 $f(t)$，我们来建立一微分方程.

$$f'(t) = -\int_{-\infty}^{\infty} x\sin tx \cdot \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx$$

$$=\int_{-\infty}^{\infty}\sin tx\cdot\frac{1}{\sqrt{2\pi}}\cdot\left(e^{-\frac{x^2}{2}}\right)'dx$$

$$=-\int_{-\infty}^{\infty}t\cos tx\cdot\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx$$

$$=-tf(t),$$

即

$$f'(t)+tf(t)=0。$$

再连同初值条件(按(2.3))

$$f(0)=1,$$

得

$$f(t)=e^{-\frac{t^2}{2}}。$$

对一般的正态分布，即 $X\sim N(\mu,\sigma^2)$，我们知道

$$Y=\frac{X-\mu}{\sigma}\sim N(0,1)。$$

上面已导出了 Y 的特征函数，于是由(2.2)，

$$X=\mu+\sigma Y$$

的特征函数为

$$e^{i\mu t-\frac{1}{2}\sigma^2t^2}。\tag{2.4}$$

3. 特征函数的性质

概率母函数有很好的分析性质，而且作为一种变换（从分布列到某类实变实值函数）还具有一一对应性；另外它与随机变量的数字特征还有直接的联系。对于一般随机变量都有定义的特征函数，也有很好的分析性质，作为一种变换（从分布函数到某类实变复值函数）不仅具有一一对应性，而且还有某种连续性，此外，它与数字特征也有直接的联系。

这里，我们不加证明地列出有关结果，以便后面引用。

性质 1 设 $f(t)$ 是 X 的特征函数，则

(i) $|f(t)|\leqslant 1$.

(ii) $f(t)$ 在 $(-\infty,+\infty)$ 上一致连续.

(iii) 若 $E(X^k)$ 存在，则对任 $t\in(-\infty,+\infty)$，$f(t)$ k 阶可导，且

$$f^{(k)}(t)=i^kE(X^ke^{itX}).$$

特别有

$$f^{(k)}(0)=i^kE(X^k).\tag{2.5}$$

(iv) $f(t)$ 是“非负定”的，即对任意的正整数 n 及任意实数 $t_1,\cdots,t_n$ 与复数 $\lambda_1,\cdots,\lambda_n$，总有

$$\sum_{k=1}^{n}\sum_{j=1}^{n}f(t_k-t_j)\lambda_k\bar{\lambda}_j\geqslant 0.$$

性质 2 设 $X_1,\cdots,X_n$ 相互独立，$f_1,\cdots,f_n$ 分别为它们的特征函数，则

$$Y=X_1+\cdots+X_n$$

的特征函数

$$f(t)=f_1(t)f_2(t)\cdots f_n(t).\tag{2.6}$$

性质 3 由(2.1)所建立的由分布函数 $F(x)$ 到特征函数 $f(t)$ 的变换是一一对应的。

注意，$E(e^{itX})=\int_{-\infty}^{\infty}e^{itx}dF(x)$，因此，$f(t)$ 实质上是由 X 的分布函数 $F(x)$ 决定的。所谓一一对应的，这里指：若 $F_1(x)\neq F_2(x)$，则 $f_1(t)\neq f_2(t)$。从 $f(t)$ 到 $F(x)$ 的具体公式如下：

$$F(b)-F(a)=\frac{1}{2\pi}\lim_{T\to+\infty}\int_{-T}^{T}\frac{e^{-ita}-e^{-itb}}{it}f(t)dt,$$

其中 $a<b$，皆是 $F(x)$ 的连续点。

性质 4 分布函数列 $\{F_n(x),n\geqslant 1\}$ 与分布函数 $F(x)$ 具有关系(这种关系称为分布函数列 $\{F_n(x),n\geqslant 1\}$ 以 $F(x)$ 为极限分布)

$$\lim_{n\to\infty}F_n(x)=F(x)\quad（对任 F(x) 的连续点），$$

当且仅当

$$\lim_{n\to\infty} f_n(t) = f(t), \tag{2.7}$$

其中 $f(t)$是 $F(x)$的特征函数，而 $f_n(t)$是 $F_n(x)$的.

§3 二项分布的正态逼近

1. 斯梯林(Stirling)公式

数学分析课程中已证明

$$n! \sim \left(\frac{n}{e}\right)^n \sqrt{2\pi n},$$

即

$$\frac{n!}{\left(\frac{n}{e}\right)^n \sqrt{2\pi n}} \to 1 \quad (n\to\infty), \tag{3.1}$$

这就是著名的斯梯林公式。本节中我们将用它来讨论二项分布

$$b_k(n,p) = C_n^k p^k (1-p)^{n-k}, \quad 0\leqslant k\leqslant n$$

的逼近问题(对一个固定的 p 值和很大的 n)；比如，对$\left(n=100,\right.$ $\left.k=50,\ p=\frac{1}{2}\right)$,

$$b_{50}\left(100, \frac{1}{2}\right) = \frac{100!}{50!50!} \cdot \frac{1}{2^{100}}.$$

一般地，我们有

$$C_{2n}^n \frac{1}{2^{2n}} = \frac{(2n)!}{n!n!} \cdot \frac{1}{2^{2n}}$$

$$\sim \frac{\left(\frac{2n}{e}\right)^{2n} \sqrt{4\pi n}}{\left(\frac{n}{e}\right)^{2n} 2\pi n} \cdot \frac{1}{2^{2n}}$$

$$= \frac{1}{\sqrt{\pi n}}.$$

因此

$$b_{50}\left(100, \frac{1}{2}\right) \doteq \frac{1}{\sqrt{50\pi}} \doteq 0.08.$$

2. 局部极限定理

定理 设 $0<p<1$，$q=1-p$，记

$$x_k = \frac{k-np}{\sqrt{npq}}, \quad 0\leqslant k\leqslant n, \tag{3.2}$$

则对满足$|x_k|\leqslant A$ 的 k，一致的有

$$C_n^k p^k q^{n-k} \sim \frac{1}{\sqrt{2\pi npq}} e^{-x_k^2/2} \quad (n\to\infty). \tag{3.3}$$

证 首先，对 $|x_k|\leqslant A$ 的 k，由(3.2)，有

$$k = np + \sqrt{npq}\,x_k,$$

$$n-k = nq - \sqrt{npq}\,x_k.$$

因此，一致的有

$$\left.\begin{aligned} k &\sim np, \\ n-k &\sim nq. \end{aligned}\right\} \quad (n\to\infty) \tag{3.4}$$

于是

$$C_n^k p^k q^{n-k} = \frac{n!}{k!(n-k)!} p^k q^{n-k}$$

$$\sim \frac{\left(\frac{n}{e}\right)^n \sqrt{2\pi n}\, p^k q^{n-k}}{\left(\frac{k}{e}\right)^k \sqrt{2\pi k} \left(\frac{n-k}{e}\right)^{n-k} \sqrt{2\pi(n-k)}}$$

$$= \sqrt{\frac{n}{2\pi k(n-k)}} \left(\frac{np}{k}\right)^k \left(\frac{nq}{n-k}\right)^{n-k}$$

$$\sim\sqrt{\frac{1}{2\pi npq}}\cdot\varphi(n,k),$$

其中

$$\varphi(n,k)=\left(\frac{np}{k}\right)^k\left(\frac{nq}{n-k}\right)^{n-k}.$$

取对数

$$\ln\varphi(n,k)=\ln\left(\frac{np}{k}\right)^k+\ln\left(\frac{nq}{n-k}\right)^{n-k},$$

而

$$\ln\left(\frac{np}{k}\right)^k=k\ln\left(1-\frac{\sqrt{npq}x_k}{k}\right)$$

$$=k\left(-\frac{\sqrt{npq}x_k}{k}-\frac{npqx_k^2}{2k^2}+\alpha(n,k)\right),$$

$$\ln\left(\frac{nq}{n-k}\right)^{n-k}=(n-k)\ln\left(1+\frac{\sqrt{npq}x_k}{n-k}\right)$$

$$=(n-k)\left(\frac{\sqrt{npq}x_k}{n-k}-\frac{npqx_k^2}{2(n-k)^2}+\beta(n,k)\right),$$

于是有

$$\ln\varphi(n,k)=-\frac{npq}{2k}x_k^2-\frac{npq}{2(n-k)}x_k^2+k\alpha(n,k)+(n-k)\beta(n,k).$$

$$=-\frac{n^2pq}{2k(n-k)}x_k^2+k\alpha(n,k)+(n-k)\beta(n,k).$$

由(3.4)，对$|x_k|\leqslant A$的k，一致的有

$$-\frac{n^2pq}{2k(n-k)}x_k^2\sim-\frac{x_k^2}{2}\quad(n\to\infty).$$

对于第二、三项，我们先导出一个事实，再据此来证明它们当$n\to\infty$时，对$|x_k|\leqslant A$的k，一致的趋于0.

引理　对$|x|\leqslant\frac{2}{3}$，有

$$\ln(1+x)=x-\frac{x^2}{2}+\theta(x),$$

其中

$$|\theta(x)|\leqslant|x|^3.$$

证　由Taylor展式，

$$\ln(1+x)=x-\frac{x^2}{2}+\cdots+(-1)^{n-1}\frac{x^n}{n}+\cdots,\qquad|x|<1.$$

记

$$\theta(x)=\frac{x^3}{3}-\frac{x^4}{4}+\cdots,$$

则

$$\ln(1+x)=x-\frac{x^2}{2}+\theta(x).$$

但

$$|\theta(x)|=\left|\frac{x^3}{3}-\frac{x^4}{4}+\cdots\right|$$

$$\leqslant\frac{|x|^3}{3}(1+|x|+|x|^2+\cdots)$$

$$=\frac{|x|^3}{3}\frac{1}{1-|x|},$$

因而，当$|x|\leqslant\frac{2}{3}$时，有

$$|\theta(x)|\leqslant|x|^3.$$

下面我们来分析第二项，即$k\alpha(n,k)$。由引理

$$|k\cdot\alpha(n,k)|\leqslant k\cdot\left|\frac{\sqrt{npq}x_k}{k}\right|^3=\frac{(npq)^{\frac{3}{2}}}{k^2}|x_k|^3,$$

于是由(3.4)得，对$|x_k|\leqslant A$的k，一致的趋于0. 通过类似的讨论知，第三项亦一致趋于0.

综上所述，我们有：对$|x_k|\leqslant A$的k，一致的有

$$\ln\varphi(n,k)\sim-\frac{x_k^2}{2}\quad(n\to\infty),$$

因而，一致的有

$$\varphi(n,k)\sim e^{-x_k^2/2}\quad(n\to\infty).$$

定理得证。

3. 积分极限定理(De Moivre-Laplace 定理)

在局部极限定理的基础上，我们来导出更具理论与应用价值的积分极限定理。

定理 对任 $-\infty<a<b<+\infty$，有

$$\lim_{n\to\infty}P\left(a<\frac{S_n-np}{\sqrt{npq}}\leqslant b\right)=\frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{x^2}{2}}dx,\tag{3.5}$$

其中随机变量列 S_n 服从参数为(n,p)的二项分布。

证 记

$$x_k=\frac{k-np}{\sqrt{npq}},$$

则

$$P\left(a<\frac{S_n-np}{\sqrt{npq}}\leqslant b\right)$$

$$=\sum_{a<x_k\leqslant b}P(S_n=k)①$$

$$=\sum_{a<x_k\leqslant b}C_n^kp^kq^{n-k}.$$

由(3.3)，并注意到“一致的有”

$$P\left(a<\frac{S_n-np}{\sqrt{npq}}\leqslant b\right)\sim\sum_{a<x_k\leqslant b}\frac{1}{\sqrt{2\pi npq}}e^{-x_k^2/2}.$$

① 记号 $\sum\limits_{a<x_k\leqslant b}$ 表示对所有满足“$a<x_k\leqslant b$”的k项求和。

由于 $x_{k+1}-x_k=\frac{1}{\sqrt{npq}}$，所以

$$P\left(a<\frac{S_n-np}{\sqrt{npq}}\leqslant b\right)=\frac{1}{\sqrt{2\pi}}\sum_{a<x_k\leqslant b}e^{-x_k^2/2}(x_{k+1}-x_k).$$

上面的和式实际上是定积分

$$\frac{1}{\sqrt{2\pi}}\int_a^b e^{-x^2/2}dx$$

的黎曼和(确切地说，与黎曼和差一个无穷小量)。

为看清这个事实，我们只需注意：

(i) n固定时，k与x_k是一一对应的；

(ii) $x_0,x_1,\cdots,x_n$ 在区间$[-\sqrt{np/q},\ \sqrt{nq/p}]$中等间隔（间隔为$1/\sqrt{npq}$），因此当$n\to\infty$时，该区间总能包含$(a,b]$，且间隔$\to0$。于是，我们有

$$\frac{1}{\sqrt{2\pi}}\sum_{a<x_k\leqslant b}e^{-x_k^2/2}(x_{k+1}-x_k)\to\frac{1}{\sqrt{2\pi}}\int_a^b e^{-x^2/2}dx\quad(n\to\infty).$$

因而有(3.5)。

注 (i) 在局部极限定理的证明中，为(3.3)式中的“一致的有”，我们多付了不少精力；而从本定理的证明过程中看出，这是该付的。

(ii) (3.5)式中，考虑随机变量

$$\frac{S_n-np}{\sqrt{npq}}$$

是自然的，这是一种将随机变量“标准化”的处理方式。所谓将随机变量Y“标准化”，是指将Y转换为

$$Y^*=\frac{Y-E(Y)}{\sqrt{\mathrm{Var}(Y)}}$$

来考虑问题。这时，$E(Y^*)=0$，$\mathrm{Var}(Y^*)=1$。

例3.1 设船舶在某海区航行，已知每遭受一次波浪的冲击，纵摇角度大于6°的概率为 $p=\frac{1}{3}$，若船舶遭受了 90000 次波浪冲击，问其中有29500—30500次纵摇角度大于6°的概率是多少？

解 将 9 万次波浪中纵摇角度大于6°的次数记为X，则

$$P(X=k)=C_{90000}^{k}\left(\frac{1}{3}\right)^{k}\left(\frac{2}{3}\right)^{90000-k},\quad k=0,1,\cdots,90000.$$

所求的概率为

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

而

$$np=90000\cdot\frac{1}{3}=30000,$$

$$\sqrt{np(1-p)}=\sqrt{20000},$$

所以

$$a\equiv\frac{29500-np}{\sqrt{np(1-p)}}=-\frac{5}{2}\sqrt{2},$$

$$b\equiv\frac{30500-np}{\sqrt{np(1-p)}}=\frac{5}{2}\sqrt{2}.$$

按积分极限定理(3.5)式，

$$P(29500<X\leqslant 30500)$$
$$=P\left(a<\frac{X-np}{\sqrt{np(1-p)}}\leqslant b\right)$$
$$\doteq\int_{-\frac{5}{2}\sqrt{2}}^{\frac{5}{2}\sqrt{2}}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx\doteq 0.9995.$$

§4 中心极限定理

1. 中心极限定理

利用§2所讨论的特征函数这个有力工具，可以得到比上一节的积分极限定理一般得多的极限定理。

定理(Lindeberg-Levy) 设 $X_1,X_2,\cdots,X_n,\cdots$是独立同分布的随机变量列，$E(X_1)$，$\mathrm{Var}(X_1)$都存在，且$\mathrm{Var}(X_1)\neq 0$。记$\mu\equiv E(X_1)$，$\sigma^2\equiv\mathrm{Var}(X_1)$，则对任$x$，有

$$\lim_{n\to\infty}P\left(\frac{S_n-n\mu}{\sqrt{n\sigma^2}}\leqslant x\right)=\Phi(x),\tag{4.1}$$

其中

$$S_n=X_1+X_2+\cdots+X_n.$$

证 记

$$Z_n\equiv\frac{S_n-n\mu}{\sqrt{n\sigma^2}},$$

于是(4.1)化为

$$F_{Z_n}(x)\to\Phi(x)\quad(n\to\infty).\tag{4.1'}$$

为证(4.1′)，我们考虑 Z_n 的特征函数$f_n(t)$。令

$$Y=\frac{X_1-\mu}{\sigma},$$

并记Y的特征函数为$f(t)$。由于

$$Z_n=\frac{1}{\sqrt{n}}\left(\frac{X_1-\mu}{\sigma}+\frac{X_2-\mu}{\sigma}+\cdots+\frac{X_n-\mu}{\sigma}\right),$$

且$X_1,\cdots,X_n$独立同分布，所以由特征函数的性质((2.6)式)知

$$\sum_{k=1}^{n}\frac{X_k-\mu}{\sigma}$$

的特征函数为$(f(t))^n$。再由(2.2)式知Z_n的特征函数为

$$f_n(t)=\left(f\left(\frac{t}{\sqrt{n}}\right)\right)^n. \qquad (4.2)$$

现在再分析$f(t)$，它是$Y=\dfrac{X_1-\mu}{\sigma}$的特征函数(由于$Y$是标准化的)我们有$E(Y)=0$，$\mathrm{Var}(Y)=1$。据特征函数有关性质((2.5)式)得

$$f'(0)=\mathrm{i}E(Y)=0,$$
$$f''(0)=\mathrm{i}^2E(Y^2)=-1.$$

由 Taylor 公式，

$$f(t)=1-\frac{t^2}{2}+o(t^2),$$

或写成

$$f(t)=1-\frac{t^2}{2}+\alpha(t)\cdot t^2, \qquad (4.3)$$

其中

$$\alpha(t)\to 0\ (t\to 0).$$

将(4.3)代入(4.2)，并取对数得

$$\ln f_n(t)=n\ln\left(1-\frac{t^2}{2n}+\alpha\left(\frac{t}{\sqrt{n}}\right)\cdot\frac{t^2}{n}\right)$$

$$\to -\frac{t^2}{2}\ (n\to\infty),$$

即

$$f_n(t)\to e^{-\frac{t^2}{2}}\ (n\to\infty).$$

注意到上式右端是标准正态分布的特征函数(参见(2.4)式)，于是由特征函数的性质知(4.1′)式成立。

注 积分极限定理仅是本定理的一个推论(请读者说明之)但本书还是详细介绍了它，原因有三：(i) 积分极限定理推动人们讨论更一般的形式；(ii) 关于积分极限定理的证明，是完整的，摸得着的；(iii) 通过积分极限定理可以从反面体会特征函数这

个工具的锐利性，而就特征函数的实质，仅是一种变换而已。

例4.1 一加法器同时收到20个噪声电压$V_k(k=1,2,\cdots,20)$，假设它们是相互独立的，且都在区间$(0,10)$上均匀分布，记$V=\sum\limits_1^{20}V_k$，求$P(V>105)$。

解 由题设知

$$E(V_1)=5,\quad \mathrm{Var}(V_1)=\frac{100}{12}.$$

由中心极限定理，

$$\frac{V-20\times 5}{\sqrt{20\times 100/12}}$$

近似服从标准正态分布，于是

$$P(V>105)=P\left(\frac{V-20\times 5}{\sqrt{20\times 100/12}}>\frac{105-20\times 5}{\sqrt{20\times 100/12}}\right)$$

$$=P\left(\frac{V-20\times 5}{\sqrt{20\times 100/12}}>0.387\right)$$

$$\doteq 1-\Phi(0.387)=0.348.$$

注 在随机模拟中，产生正态随机数的方法之一是，产生12个独立的$(0,1)$均匀分布随机数，然后求和。

2. Lindeberg 条件

这小节介绍较一般的古典中心极限定理，它并不要求“同分布”。

对随机变量列$X_1,X_2,\cdots,X_n,\cdots$，记

$$a_k\equiv E(X_k),\quad b_k\equiv\sqrt{\mathrm{Var}(X_k)}>0,$$

并记$F_k(x)$，$k=1,2,\cdots$为X_k的分布函数，称下列条件为 Lindeberg 条件：

对任$\varepsilon>0$，有

$$\lim_{n\to\infty}\frac{1}{B_n^2}\sum_{k=1}^{n}\int_{|x-a_k|>\varepsilon B_n}(x-a_k)^2\mathrm{d}F_k(x)=0,\qquad(4.4)$$

其中

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

定理(Lindeberg-Feller) 设 $X_1,X_2,\cdots$是相互独立的随机变量列，满足条件(4.4)，则对任意x，有

$$\lim_{n\to\infty}P\left(\frac{1}{B_n}\sum_{k=1}^{n}(X_k-a_k)\leqslant x\right)=\Phi(x).\qquad(4.5)$$

这个定理可用特征函数来证。但由于超出本书范围，故我们只给出定理，不加证明。

注 (i) 关于条件(4.4)的概率意义，由下列不等式能看到一个侧面：

$$\begin{aligned}&\frac{1}{B_n^2}\sum_{k=1}^{n}\int_{|x-a_k|>\varepsilon B_n}(x-a_k)^2\mathrm{d}F_k(x)\\&\geqslant\varepsilon^2\sum_{k=1}^{n}\int_{|x-a_k|>\varepsilon B_n}\mathrm{d}F_k(x)\\&=\varepsilon^2\sum_{k=1}^{n}P\left(\frac{|X_k-a_k|}{B_n}>\varepsilon\right)\\&\geqslant\varepsilon^2P\left(\bigcup_1^n\left(\frac{|X_k-a_k|}{B_n}>\varepsilon\right)\right)\\&=\varepsilon^2P\left(\max_{1\leqslant k\leqslant n}\frac{|X_k-a_k|}{B_n}>\varepsilon\right),\end{aligned}$$

于是，由(4.4)成立可得

$$P\left(\max_{1\leqslant k\leqslant n}\frac{|X_k-a_k|}{B_n}>\varepsilon\right)\to0\ (n\to\infty).$$

上式表明，相对偏差$\left(\text{即}\left|\frac{X_k-a_k}{B_n}\right|\right)$一致的小的概率接近1。

(ii) 在$\{X_n,n\geqslant1\}$为相互独立随机变量列的前提下，条件(4.4)不仅是(4.5)成立的充分条件，而且也差不多是必要条件；确切的说，只需再添加两个很普通的条件，(4.4)就是(4.5)成立的必要条件。这两个条件是：

$$\lim_{n\to\infty}B_n=\infty,\qquad\lim_{n\to\infty}\frac{b_n}{B_n}=0.$$

§5 大数定律与强大数定律

1. 大数定律

定理(切比雪夫) 设$\{X_n,n\geqslant1\}$是相互独立的随机变量列，若对一切k有

$$\mathrm{Var}(X_k)\leqslant C$$

则对任$\varepsilon>0$，有

$$P\left(\left|\frac{S_n-E(S_n)}{n}\right|\geqslant\varepsilon\right)\to0\ (n\to\infty),\qquad(5.1)$$

其中

$$S_n=\sum_1^nX_k.$$

证 由切比雪夫不等式得

$$\begin{aligned}P\left(\left|\frac{S_n}{n}-E\left(\frac{S_n}{n}\right)\right|\geqslant\varepsilon\right)&\leqslant\frac{1}{\varepsilon^2}\mathrm{Var}\left(\frac{S_n}{n}\right)\\&=\frac{1}{n^2\varepsilon^2}\sum_1^n\mathrm{Var}(X_k)\leqslant\frac{1}{n\varepsilon^2}.\end{aligned}$$

由此立得(5.1)。

系(Bernoulli) 设每次试验中事件A出现的概率为p，ν_n是

A在n次重复独立试验中出现的次数，则对任意$\varepsilon>0$，有

$$\lim_{n\to\infty} P\left(\left|\frac{\nu_n}{n}-p\right|\geqslant\varepsilon\right)=0. \tag{5.2}$$

证 令

$$X_k=\begin{cases}1, & \text{第}k\text{次试验中}A\text{发生,}\\ 0, & \text{第}k\text{次试验中}A\text{未发生,}\end{cases} \quad k=1,2,\cdots,$$

于是对一切k有

$$\mathrm{Var}(X_k)=pq\leqslant\frac{1}{4}.$$

由(5.1)即得(5.2)。

说明 (5.2)式的概率意义是，n次重复独立试验中，A出现的频率$\left(\text{即}\frac{\nu_n}{n}\right)$有稳定值$p$，即频率在$p$附近摆动(偏差很小的概率趋于1)。该式从一个侧面刻划了频率与概率的关系；然而，更深刻的结果由下面的定理给出。

2. 强大数定律

定理(柯尔莫哥洛夫) 设$\{X_n,n\geqslant1\}$是相互独立的随机变量列，满足

$$\sum_{n=1}^{\infty}\frac{\mathrm{Var}(X_n)}{n^2}<\infty,$$

则

$$P\left(\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}(X_k-E(X_k))=0\right)=1. \tag{5.3}$$

说明 (i) 本定理的证明已超出本书的范围。

(ii) 若用记号$S_n=\sum_1^n X_k$，(5.3)式可改写为

$$P\left(\lim_{n\to\infty}\frac{S_n-E(S_n)}{n}=0\right)=1. \tag{5.3'}$$

(iii) 可以证明由(5.3′)可得(5.1)。

(iv) 显然，由$\mathrm{Var}(X_k)\leqslant C$(对一切$k$)，可导出$\sum\frac{\mathrm{Var}(X_n)}{n^2}<\infty$，因此切比雪夫大数定律仅是本定理的一个推论。或者说，本定理较为深刻。

系 设每次试验中事件A出现的概率为p，ν_n是A在n次重复独立试验中出现的次数，则有

$$P\left(\lim_{n\to\infty}\frac{\nu_n}{n}=p\right)=1. \tag{5.4}$$

说明 (5.4)式表明，n次重复独立试验中，A出现的频率“几乎必然”以p为极限。

注 (i) 对相互独立而且有相同分布的随机变量列$\{X_n,n\geqslant1\}$，有一个更完美的结果(也是由柯尔莫哥洛夫给出的)：

设$\{X_n,n\geqslant1\}$独立同分布，则

$$P\left(\lim_{n\to\infty}\frac{1}{n}\sum_1^n X_k=a\right)=1 \tag{5.5}$$

的充分必要条件是：$E(X_1)$存在且等于a。

(ii) 随机变量列$\{X_n,n\geqslant1\}$收敛于一个随机变量$X(n\to\infty)$，或说以X为极限$(n\to\infty)$，有多种形式，本章涉及到三种最基本的形式：

(a) $F_{X_n}(x)\to F_X(x)\quad(n\to\infty)$.

这就是所谓$\{X_n,n\geqslant1\}$依分布收敛到X。比如，(4.1′)表示$\{Z_n,n\geqslant1\}$依分布收敛到标准正态随机变量。

(b) 任给$\varepsilon>0$，

$$P(|X_n - X|>\varepsilon)\to 0\ (n\to\infty).$$

这就是所谓$\{X_n, n\geqslant 1\}$依概率(测度)收敛到X。比如,(5.1)表示随机变量列$\left\{\frac{S_n - E(S_n)}{n},\ n\geqslant 1\right\}$依概率收敛到0。又如,(5.2)表示随机变量列$\left\{\frac{\nu_n}{n},\ n\geqslant 1\right\}$依概率收敛到$p$。

(c) $P(\lim\limits_{n\to\infty} X_n = X) = 1$.

这就是所谓$\{X_n, n\geqslant 1\}$几乎必然(或说概率为1的)收敛到X。比如,(5.3′)表示随机变量列$\left\{\frac{S_n - E(S_n)}{n};\ n\geqslant 1\right\}$几乎必然收敛到0。又如,(5.5)表示随机变量列$\left\{\frac{1}{n}\sum_1^n X_k, n\geqslant 1\right\}$几乎必然收敛到$a$。

这三种形式的收敛,我们分别用下列较为统一的记号来表示:

$X_n\to X$ (依分布),

$X_n\to X$ (依概率),

$X_n\to X$ (几乎必然)。

可以证明这三种收敛性一个比一个强。比如,若$X_n\to X$(几乎必然),则$X_n\to X$(依概率)。

习 题 五

1. 求负二项分布

$$P(X=k)=\binom{k+r-1}{r-1}p^r q^k,\quad k=0,1,\cdots$$

的母函数

2. 求分布列,其母函数为:

(1) $\frac{1}{4}(1+z)^2$;

(2) $\frac{1}{2-z}$;

(3) e^{z-1}.

3. 设非负整值随机变量X的母函数为$g(z)$,求$X+1$及$2X$的母函数。

4. 设非负整值随机变量X的母函数为$g(z)$,求:

(1) $\sum\limits_{k=0}^{\infty}P(X\leqslant k)z^k$;

(2) $\sum\limits_{k=0}^{\infty}P(X=2k)z^k$.

5. 掷四个正20面体骰子,求总点数为10的概率。

6. 设甲、乙二人各掷均匀硬币n次,试利用母函数求甲得正面的次数比乙得正面的次数多k次的概率。

7. 在重复独立试验中,每次试验成功的概率为p,令X为首次在成功之后即遇到失败的试验次数(即“$X=n$”意味着第$n-1$次试验成功,第n次试验失败,但在前$n-2$次试验中没有一次失败是紧跟在成功之后的)。求X的母函数,并找出$E(X)$,$\mathrm{Var}(X)$。

8. 求下列分布的特征函数

(1) $(-a,a)$上的均匀分布,

(2) 哥西分布,即密度函数为

$$p(x)=\frac{a}{\pi}\cdot\frac{1}{(x-b)^2+a^2},\quad -\infty<x<\infty;$$

(3) 伽玛分布,即密度函数为

$$p(x)=\begin{cases}\frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}, & x>0;\\ 0, & x\leqslant 0.\end{cases}$$

9. 直接找出下列特征函数的分布函数

(1) $\cos t$;