第一章　随机变量概述

读者极有可能已不同程度地对本章的绝大部分内容有所了解. 本章在重温这些内容时, 定位在中等水平: 一方面, 将回避诸如 “样本空间” 或 “测度” 这些概念 (因为在本书中并不需要它们); 另一方面, 将经常使用 Stieltjes 积分, 读者最先会在公式 (1.5) 中遇到这一积分. 利用 Stieltjes 积分 (通常借助随机变量的累积分布函数表述), 便可避免冗烦地区分 “离散的” 与 “连续的” 情形. 关于 Stieltjes 积分的一个优秀的入门性介绍可在 Bühlmann 著作 [12] 的附录中找到.

1. 一维随机变量

只要一个试验的结果是一个数字, 而且带有偶然性, 这一数值结果便可命之为随机变量. 一个随机变量的结果 (即数值) 以 X 记之, 它是不能预知的. 能够预知的 (或至少预先能肯定它是存在的) 乃是它的分布, 它可以由累积 **分布函数**(cdf)$F_X(\cdot)$ 表示, 这里

$$F_X(x) = P(X \le x), \quad -\infty < x < \infty \tag{1.1}$$

等于 X 在区间 $(-\infty, x]$ 中取值的概率. 作为通常的概率公理的推论, 由上述定义知 cdf 是右连续的非减函数, 而且随着自变量自 $-\infty$ 增加到 ∞, 其函数值则由 0 增至 1. 如果清楚地知道所谈及的是哪一个随机变量, 也可略去下标, 而以 $F(\cdot)$ 简记 $F_X(\cdot)$.

【例 1.1】 一个自区间 $[0,1]$ 中随机抽取的数 U 是一随机变量，且有

$$F_U(x)=\begin{cases}0 & x<0\\ x & 0\le x<1\\ 1 & x\ge 1\end{cases} \tag{1.2}$$

【例 1.2】 掷一颗骰子，并以 N 表示观察到的点数，则它也是一随机变量. 如果骰子是“均匀”的， N 的 *cdf* 为

$$F_N(x)=\begin{cases}0 & x<1\\ \dfrac{i}{6} & i\le x<i+1 \quad (i=1,\cdots,5)\\ 1 & x\ge 6\end{cases} \tag{1.3}$$

连续分布 指的是 *cdf* 为连续函数的情形. 如果这一函数是可微的，则称该分布是 *绝对连续的*, 并称 $f(x)=F'(x)$ 为 **概率密度函数** (*pdf*). 如 *cdf* F 是一阶梯函数，则称相应的分布为 **离散分布**. 进一步，如果 F 的跳跃均发生在某一正数 d 的整数倍处，则称该分布是 *以 d 为跨度的算术分布*. 此时，

$$f_i=F_X(id)-F_X(id\text{–}0)=\mathrm{P}\ [X=id] \tag{1.4}$$

称为是 X 的 **概率频率函数**(pff).

对于一个具有 *cdf* F 的随机变量 X 来说，有若干特征量是有兴趣的，其中有它的 **均值**(或 **期望**)

$$E[X]=\int xdF(x) \tag{1.5}$$

与 *方差*

$$\mathrm{Var}\ [X]=\mathrm{E}\ [(X-\mathrm{E}\ X)^2]=\int (x-\mathrm{E}\ [X])^2dF(x) \tag{1.6}$$

在将平方式展开后，便得下述熟知的公式:

$$\mathrm{Var}\ [X]=\mathrm{E}\ [X^2]-(\mathrm{E}\ [X])^2 \tag{1.7}$$

X 的关于零点的 k 阶矩可如下定义:

$$\mathrm{E}\left[X^k\right]=\int x^k dF(x), \quad \forall k \geq 1 \tag{1.8}$$

X 的值域是一个可能为无限的区间 (l_x, r_x), 其中

$$l_x=\inf\{x|F(x)>0\}$$
$$r_x=\sup\{x|F(x)<1\}$$

如 $l_x \geq 0$, 则 X 是一个非负随机变量, 可用不等式 $X \geq 0$ 示之. 鉴于不同的分布可能具有以上提及的任一种相等的特征量, 譬如说均值, 我们就只能认为这些特征量仅仅刻画了 X 的部分特征.

与随机变量 X 有联系的还有一个如下定义的函数 $M(t)=M_X(t)$:

$$M(t)=\mathrm{E}\left[e^{tX}\right]=\int e^{tx} dF(x) \tag{1.9}$$

它称为是 X 的矩母函数 (mgf). 注意, $M(t)$ 仅定义在使得方程 (1.9) 中的积分收敛的区间上. 本专著主要讨论这样一些随机变量, 它们的分布函数在“尾部”(即在 $\pm\infty$ 处) 锐减, 从而使得 mgf 在某一具有正长度的区间中存在. 这种区间自然总是包含点 $t=0$ 的, 这是因为 $M(0)=1$.

mgf 具有完全特征的意义: 如果两个随机变量具有相同的 mgf, 它们的分布必恒同. 鉴于这种一一对应关系, mgf 便成为研究随机变量的有用工具: 以 mgf 表述的任何结论都可转述成关于分布的相应结论.

矩母函数 (mgf) 的得名起因于下述公式:

$$\mathrm{E}\left[X^k\right]=M^{(k)}(0), \quad \forall k \geq 1 \tag{1.10}$$

上述等式可如下推得: 首先, 针对变量 t 将方程微商 k 次, 然后在方程的右端交换积分与求导的顺序, 最后再令 $t=0$ 即得.

对于非负随机变量 $X \geq 0$ 来说，当 $t \leq 0$ 时，$M_X(t)$ 总是存在的. 这样, 习惯上可作一变量替换, 令 $s = -t$, 并记 $L_X(s) = M_X(t)$, 即有

$$L_X(s) = \mathrm{E}\left[e^{-sX}\right] = \int e^{-sx} dF(x), \quad \forall s \geq 0 \tag{1.11}$$

通常称 $L_X(s)$ 是 X 的 Laplace变换. 下述 (属于 S. Bernstein 的) 著名结果刻画了可表为 Laplace 变换的所有函数的特征.

定理 一函数 $\mathcal{L}(s)$ $s \geq 0$ 是某一 *cdf* 的 Laplace 变换的充要条件为 $\mathcal{L}(0) = 1$, 无穷次可导，且满足

$$(-1)^n \mathcal{L}^{(n)}(s) \geq 0, \quad \forall s \geq 0, \quad n \geq 0$$

关于这一定理的证明可参阅 Feller[22] 中的 13.4 节. 定理中叙述的条件显然是必要的，其充分性之证明则远非平凡.

2. 交换函数与期望的次序

若 $u(x)$, $-\infty < x < \infty$, 是一函数， X 是一随机变量，人们通常会问 $u(\mathrm{E}[X])$ 与 $\mathrm{E}[u(X)]$ 之间有什么联系 (如 u 是线性函数，这二个量是相等的). 本节将介绍二种简单 (但很有用) 的处理方法： Jensen 不等式与 Taylor 级数方法.

为了给出 Jensen 不等式，需假定函数 u 是 凹 的，即对任一点 $(x_0, u(x_0))$ 而言，通过该点的切线总是在 u 的图形之上. 这表明，对任一 x_0,

$$u(x) \leq u(x_0) + u'(x_0)(x - x_0), \quad \forall -\infty < x < \infty \tag{2.1}$$

若 u 是二次可导的，函数的凹性等价于对一切 x, $u''(x) \leq 0$. 在方程 (2.1) 中取 $x_0 = \mathrm{E}[X]$, 可得

$$u(X) \leq u(\mathrm{E}[X]) + u'(\mathrm{E}[X])(X - \mathrm{E}[X]) \tag{2.2}$$

再在上式两端取期望，即得

$$\mathrm{E}\,[u(X)] \le u(\mathrm{E}\,[X]) \tag{2.3}$$

这便是熟知的 Jensen 不等式.

【例 2.1】 假设某人已活到年龄 y. 现考虑在其剩余寿命期间每年可付 1 元的连续年金. 按国际通用的精算学上的记法 *, 这些付款的 (以 δ 为利息力度的) 现值之期望值为 $\bar{a}_y$. 于是，$\bar{a}_y = \mathrm{E}\,[u(X)]$, 其中

$$u(x) = \bar{a}_{\overline{x}|} = \frac{1-e^{-\delta x}}{\delta}$$

它恰等于期限为 x 的定期年金的现值；而 X 则为已存活到年龄 y 的某人的剩余寿命，即有

$$\mathrm{P}\,(X \ge x) = {}_x p_y = \frac{l_{x+y}}{l_y}, \quad \forall x \ge 0 \tag{2.4}$$

与 $\mathrm{E}\,[X] = \overset{\circ}{e}_y$. 由于 $u''(x) < 0$, Jensen 不等式适用，从而可得

$$\bar{a}_y \le \bar{a}_{\overline{\overset{\circ}{e}_y}|}$$

为了运用 Taylor 级数法，需假定 u 是二次可导的. 将 u 在 $\mathrm{E}\,[X]$ 附近展成 Taylor 级数，可得如下近似式：

$$\begin{aligned} u(X) \approx & u(\mathrm{E}\,[X]) + u'(\mathrm{E}\,[X])(X - \mathrm{E}\,[X]) \\ & + \frac{1}{2}u''(\mathrm{E}\,[X])(X - \mathrm{E}\,[X])^2 \end{aligned} \tag{2.5}$$

再在上式的两端取期望，即得

$$\mathrm{E}\,[u(X)] \approx u(\mathrm{E}\,[X]) + \frac{1}{2}u''(\mathrm{E}\,[X])\ \mathrm{Var}\,[X] \tag{2.6}$$

---

* 本例中提及的诸精算符号之详义可参阅 Gerber 的另一译著《人寿保险数学》(世界图书出版公司， 1996). —— 译著注

若公式 (2.5) 是合适的，上述近似便令人满意. 譬如说，当 u 接近于是一个二次函数，或当 X 的分布密集于它的均值附近时，近似式 (2.6) 就相当不错.

3. 若 干 例 子

本节将列出若干重要的分布与它们的数字特征. 表 1 给出了 (均以 1 为跨度的) 算术分布的例子. 注意，几何分布是负二项分布 ($\alpha = 1$) 的特例. 鉴于这些公式在很多其他的概率论教材中都曾讨论过，本节仅推导关于 Poisson 分布的有关公式. 首先，计算 mgf,

$$
\begin{aligned}
M(t) = \mathrm{E}\left[e^{tX}\right] &= \sum_{x=0}^{\infty} e^{tx}\frac{\lambda^x}{x!}e^{-\lambda} \\
&= \sum_{x=0}^{\infty} \frac{(\lambda e^t)^x}{x!}e^{-\lambda} = \exp\left(\lambda(e^t - 1)\right) \qquad (3.1)
\end{aligned}
$$

(注意，最后一步利用了指数函数的 Taylor 级数). 由此可得

$$
M'(t) = \lambda e^t \exp(\lambda(e^t - 1)) \qquad (3.2)
$$

与

$$
\mathrm{E}[X] = M'(0) = \lambda
$$

再将 (3.2) 式微分一次，可得

$$
M''(t) = M'(t) + \lambda e^t M'(t)
$$

于是,

$$
\mathrm{E}[X^2] = M''(0) = \lambda + \lambda^2
$$

最后，利用方程 (1.7) 可得 $\mathrm{Var}[X] = \lambda$.

表 1　　某些重要的算术分布

名　称	概率密度函数	参数	均值	方差	矩母函数
二项分布	$\binom{N}{x}p^x q^{N-x},\ x=0,1,\cdots,N$	$N=1,2,\cdots$ $0<(p=1-q)<1$	Np	Npq	$(pe^t+q)^N,\ -\infty<t<\infty$
Poisson 分布	$e^{-\lambda}\dfrac{\lambda^x}{x!},\ x=0,1,\cdots$	$\lambda>0$	λ	λ	$e^{\lambda(e^t-1)},\ -\infty<t<\infty$
负二项分布	$\binom{\alpha+x-1}{x}p^\alpha q^x,\ x=0,1,\cdots$	$\alpha>0$ $0<(p=1-q)<1$	$\dfrac{\alpha q}{p}$	$\dfrac{\alpha q}{p^2}$	$\left(\dfrac{p}{1-qe^t}\right)^\alpha,\ t<-\log q$
几何分布	$pq^x,\ x=0,1,\cdots$	$0<(p=1-q)<1$	$\dfrac{q}{p}$	$\dfrac{q}{p^2}$	$\dfrac{p}{1-qe^t},t<-\log q$

表 2　　某些重要的绝对连续分布

名　称	概率密度函数	参数	均值	方差	矩母函数
正态分布	$\dfrac{1}{\sqrt{2\pi}\sigma}\cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$, $-\infty<x<\infty$	$-\infty<\mu<\infty$ $\sigma>0$	μ	σ^2	$e^{(1/2)\sigma^2t^2+\mu t},\ -\infty<t<\infty$
gamma 分布	$\dfrac{\lambda}{\Gamma(\alpha)}(\lambda x)^{\alpha-1}e^{-\lambda x},x>0$	$\lambda>0$ $\alpha>0$	$\dfrac{\alpha}{\lambda}$	$\dfrac{\alpha}{\lambda^2}$	$\left(\dfrac{\lambda}{\lambda-t}\right)^\alpha,t<\lambda$
指数分布	$\lambda e^{-\lambda x},x>0$	$\lambda>0$	$\dfrac{1}{\lambda}$	$\dfrac{1}{\lambda^2}$	$\dfrac{\lambda}{\lambda-t},t<\lambda$
均匀分布	$\dfrac{1}{b-a},a<x<b$	$-\infty<a<b<\infty$	$\dfrac{a+b}{2}$	$\dfrac{(b-a)^2}{12}$	$\dfrac{e^{bt}-e^{at}}{t(b-a)},t\neq 0$

表 2 列出了某些绝对连续的分布与它们的性质. 尽管指数分布是 gamma 分布的特例 ($\alpha = 1$), 单独将其列出则是考虑到它的特殊重要性. 注意, gamma 分布是负二项分布的连续类比, 特别地, 指数分布是几何分布的连续类比. 作为说明, 考虑表 2 中关于 gamma 分布一栏的推导. 首先, 计算它的 mgf.

$$\begin{aligned} M(t) = \mathrm{E}\left[e^{tX}\right] &= \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty x^{\alpha-1} e^{-(\lambda-t)x} dx \\ &= \frac{\lambda^\alpha}{\Gamma(\alpha)} \cdot \frac{\Gamma(\alpha)}{(\lambda-t)^\alpha} = \left(\frac{\lambda}{\lambda-t}\right)^\alpha, \quad \forall t < \lambda \end{aligned} \tag{3.3}$$

注意, 仅当 $t < \lambda$ 时, $M(t)$ 才存在. 它的 k 阶导数为

$$M^{(k)}(t) = (\alpha + k - 1)^{(k)} \frac{\lambda^\alpha}{(\lambda - t)^{\alpha+k}}, \quad \forall k \geq 1 \tag{3.4}$$

(以上符号 $x^{(k)}$ 表示 $x(x-1)\cdots(x-k+1)$). 这样,

$$\mathrm{E}\left[X^k\right] = M^{(k)}(0) = \frac{(\alpha+k-1)^{(k)}}{\lambda^k} \tag{3.5}$$

特别地,

$$\mathrm{E}[X] = \frac{\alpha}{\lambda}, \quad \mathrm{E}\left[X^2\right] = \frac{\alpha(\alpha+1)}{\lambda^2}$$

由此推知

$$\mathrm{Var}\,[X] = \frac{\alpha}{\lambda^2}$$

4. 随机变量族

经常要同时考虑若干个随机变量, 如 $X_1, \cdots, X_n$. 这样, 模型便需由它们的联合分布描述. 联合 cdf $F = F_{X_1,\cdots,X_n}$ 是 n 元函数, 它由下式给出:

$$F(x_1, \cdots, x_n) = \mathrm{P}\left[X_1 \leq x_1, \cdots, X_n \leq x_n\right] \tag{4.1}$$

其中 $-\infty < x_i < \infty$ $(i=1,\cdots,n)$. 在更为一般的情形，如考虑随机过程时，要涉及随机变量的无限族. 此时还需要引入若干附加的技巧，但基本的原理则是相同的.

X_i 的 边缘 *cdf* 可通过在 (4.1) 式中令

$$x_1 = \cdots = x_{i-1} = x_{i+1} = \cdots = x_n = \infty$$

而获得. 若联合 *cdf* 等于它们的边缘 *cdf* 的乘积，则称随机变量 $X_1,\cdots,X_n$ 是相互独立的.

如 $h(x_1,\cdots,x_n)$ 是 n 元函数， $Y=h(X_1,\cdots,X_n)$ 便是一随机变量. 它的期望可通过两种途径求得，或先求出 Y 的边缘分布 $F_Y(y)$, 或可直接由 n 重积分算出:

$$\mathrm{E}[Y] = \int h(x_1,\cdots,x_n)dF(x_1,\cdots,x_n) \tag{4.2}$$

通常，第二种途径更方便些.

特别有兴趣的是形如 $(X_i-\mathrm{E}[X_i])(X_j-\mathrm{E}[X_j])$ 的随机变量的期望:

$$\mathrm{Cov}\,[X_i,X_j] = \mathrm{E}\left[(X_i-\mathrm{E}[X_i])(X_j-\mathrm{E}[X_j])\right] \tag{4.3}$$

在将方括号内的乘积展开后，可得

$$\mathrm{Cov}\,[X_i,X_j] = \mathrm{E}\left[X_i\ X_j\right] - \mathrm{E}[X_i]\mathrm{E}[X_j] \tag{4.4}$$

上式推广了公式 (1.7). 与此有关，重温如下公式

$$\mathrm{Var}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n}\mathrm{Var}\,[X_i] + 2\sum\sum_{i<j}\mathrm{Cov}\,[X_i,X_j] \tag{4.5}$$

该式在首二阶矩存在的假定下是成立的. 注意，上式与下式

$$\mathrm{E}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n}\mathrm{E}[X_i] \tag{4.6}$$

一样，它们的成立并不要求诸 X_i 的相互独立性. 若诸 X_i 相互独立，它们自然也是“不相关的”，从而由公式 (4.5) 可推知，相互独立的诸随机变量的和之方差恰等于各自的方差之和.

条件分布 与 *条件期望* 的概念是特别重要的. 设 X 与 Y 是两个随机变量，(给定 X 后)Y 的条件期望以 $\mathrm{E}[Y|X]$ 记之. 这一条件期望是随机变量 X 的函数，从而其自身也是一随机变量.

如果条件期望 $\mathrm{E}[Y|X]$ 以及 X 的边缘分布已知， $\mathrm{E}[Y]$ 可借助下述公式求得

$$\mathrm{E}[Y] = \mathrm{E}\left[\mathrm{E}[Y|X]\right] \tag{4.7}$$

这一公式称为是 *期望的累次法则*(以下简称为累次法则)，它可视为是 *全概律* 的引伸. 为了针对 $\mathrm{Var}[Y]$ 觅得一种相似的法则，首先可对随机变量 Y^2 运用累次法则，随之再利用公式 (1.7) 以求得

$$\begin{aligned}\mathrm{E}[Y^2] &= \mathrm{E}\left[\mathrm{E}[Y^2|X]\right] \\ &= \mathrm{E}\left[\mathrm{Var}\,[Y|X]\right] + \mathrm{E}\left[(\mathrm{E}[Y|X])^2\right]\end{aligned} \tag{4.8}$$

再分别在上式左端减去 $\mathrm{E}[Y]$ 的平方，在右端减去 $\mathrm{E}\left[\mathrm{E}[Y|X]\right]$ 的平方，便得所需要的公式：

$$\mathrm{Var}\,[Y] = \mathrm{E}\left[\mathrm{Var}\,[Y|X]\right] + \mathrm{Var}\,\left[\mathrm{E}[Y|X]\right] \tag{4.9}$$

上式表明，总的方差可分解成 *条件方差的期望* 与 *条件期望的方差* 之和.

【例 4.1】 先捻动一颗骰子，再根据骰子所示的点数，将一钱币抛掷若干次. 以 X 表示骰子所示之点数，以 Y 表示投掷钱币时“正面”出现的次数. 试求 $\mathrm{E}[Y]$ 与 $\mathrm{Var}[Y]$.

解　如骰子是均匀的，便有

$$\mathrm{P}[X=i]=\frac{1}{6},\quad i=1,\cdots,6$$

(给定 X 后)Y 的条件分布是以 X 与 $\frac{1}{2}$ 为参数的二项分布，于是有

$$\mathrm{E}[Y|X]=\frac{1}{2}X,\quad \mathrm{Var}\,[Y|X]=\frac{1}{4}X$$

再由公式 (4.7) 知

$$\mathrm{E}[Y]=\frac{1}{2}\mathrm{E}[X]=\frac{7}{4}$$

而由公式 (4.9) 便得

$$\mathrm{Var}\,[Y]=\frac{1}{4}\mathrm{E}[X]+\frac{1}{4}\mathrm{Var}\,[X]$$

稍经计算，知 $\mathrm{Var}\,[X]=\frac{35}{12}$, 从而

$$\mathrm{Var}\,[Y]=\frac{7}{8}+\frac{35}{48}=\frac{77}{48}$$

【例 4.2】　以 N 表示某一保单组合在一给定的时间区间内的索赔次数. 假设 N 服从 Poisson 分布. 如果关于 Poisson 分布的参数 Θ 也存在不确定性，便可把它也视为是一随机变量. 上述假定可更为严格地叙述如下：(给定 Θ 后)N 的条件分布是以 Θ 为参数的 Poisson 分布，而 Θ 的 *cdf* 为 $U(\theta),\theta\geq 0$. 为了求得 N 的概率频率函数，由 (4.7) 式可得

$$\begin{aligned}p_n=\mathrm{P}[N=n]&=\mathrm{E}\left[\mathrm{P}[N=n|\Theta]\right]=\mathrm{E}\left[e^{-\Theta}\frac{\Theta^n}{n!}\right]\\&=\frac{1}{n!}\int_0^{\infty}e^{-\theta}\theta^n dU(\theta),\quad \forall n\geq 0\end{aligned}\tag{4.10}$$

这样的分布称为混合 Poisson 分布. 特别地，当 U 服从以 α 与 λ 为参数的 gamma 分布时，可确切地算出 (4.10) 式中的积分，从而推知 N 事实上服从以 α 与 $p=\lambda/(\lambda+1)$ 为参数的负二项分布.

5. 相互独立的随机变量之和

首先考虑仅有两个相互独立的随机变量的情形，并以 S 表示它们的和： $S=X+Y$. 下述计算表明， S 的 mgf 是 X 与 Y 的 mgf 的乘积.

$$\begin{aligned} M_S(t) =& \mathrm{E}\left[e^{t(X+Y)}\right] = \mathrm{E}\left[e^{tX}e^{tY}\right] \\ &= \mathrm{E}\left[e^{tX}\right]\mathrm{E}\left[e^{tY}\right] = M_X(t)M_Y(t) \end{aligned} \tag{5.1}$$

(上述证明中，哪一步用到了独立性的假定？) 这一公式在许多场合都是很有用的.

【例 5.1】 假设 X 和 Y 分别服从参数为 λ 与 ν 的 Poisson 分布，再假定它们是相互独立的. 试问它们的和服从什么分布？

解　由表 1 知，

$$M_X(t)=\exp\left\{\lambda(e^t-1)\right\},\quad M_Y(t)=\exp\left\{\nu(e^t-1)\right\}$$

于是，由公式 (5.1) 知，

$$M_{X+Y}(t)=\exp\left\{(\lambda+\nu)(e^t-1)\right\}$$

这表明 $X+Y$ 仍服从 Poisson 分布，相应的参数为 $\lambda+\nu$.

为了自 X 与 Y 的 *cdf* 求得 S 的 *cdf*, 首先可利用累次法则求得

$$\begin{aligned} F_S(s) &= \mathrm{P}\left[S\le s\right] = \mathrm{E}\left[\mathrm{P}\left[S\le s|X\right]\right] \\ &= \mathrm{E}\left[\mathrm{P}\left[Y\le s-X|X\right]\right] \end{aligned} \tag{5.2}$$

如 X 和 Y 相互独立，则 $\mathrm{P}[Y \le s-X|X]=F_Y(s-X)$, 从而有

$$F_S(s)=\mathrm{E}[F_Y(s-X)]=\int F_Y(s-x)dF_X(x) \tag{5.3}$$

$F_S(s)$ 称为是 F_X 与 F_Y 的 **卷积**, 简记为 $F_S=F_X*F_Y$. 一般地说，(5.3) 中的积分限自 $-\infty$ 扩展到 ∞; 若 $X\ge 0, Y\ge 0$, 则积分可限制于 $0\le x\le s$.

若 X 与 Y 均具有绝对连续的分布，则它们的和也是绝对连续的. 在 (5.3) 式中对 s 求导，即得借助于概率密度函数表示的卷积公式

$$f_S(s)=\int f_Y(s-x)f_X(x)dx \tag{5.4}$$

如 X 与 Y 均具有跨度相同的算术分布，也可得到相似的公式；不过上述公式中的积分必须改成求和号.

【例 5.2】 假设 X 与 Y 是相互独立的 Poisson 随机变量，相应的参数分别为 λ 与 ν, 试求它们的和的概率频率函数.

解　对于 $s=0,1,\cdots$, 恒有

$$\begin{aligned}
f_S(s)&=\sum_{x=0}^{s} f_Y(s-x)f_X(x)\\
&=\sum_{x=0}^{s} e^{-\nu}\frac{\nu^{s-x}}{(s-x)!}e^{-\lambda}\frac{\lambda^x}{x!}\\
&=\frac{e^{-(\lambda+\nu)}}{s!}\sum_{x=0}^{s}\binom{s}{x}\nu^{s-x}\lambda^x\\
&=e^{-(\lambda+\nu)}\frac{(\lambda+\nu)^s}{s!}
\end{aligned} \tag{5.5}$$

上述证明中的最后一步用到了二项定理. 自然，利用 mgf 也可得到相同的结果 (见例 5.1). 如果产生的 mgf 可予识别，后一种方法

显然是优美的. 遗憾的是, 仅在一些有选择的例子中方可能识别所求得的 mgf. 这使得卷积公式更为重要.

前面叙述的概念可以明显地推广至多于两个的独立随机变量的情形. 假设 $X_1,\cdots,X_n$ 是相互独立的随机变量, 并分别具有 $cdf\,F_i$ $(i=1,\cdots,n)$. 它们的和的 cdf 以 $F_1*\cdots*F_n$ 记之, 它可递推地算出. 如诸 X_i 同分布, 即具有相同的 $cdf\,F$, 它们的和的分布简记为 F^{*n}. 今后为了记法上的方便, 对于 $n=0$, 也可定义 F^{*n}. 由于这对应于空和, 它必是常数 0 的 cdf. 于是可令

$$F^{*0}(x)=\begin{cases}0, & x<0\\ 1, & x\geq 0\end{cases}$$

6. 随 机 和

设 $X_1,\cdots,X_n$ 是独立同分布的随机变量, 并具有相同的 $cdf\,F$ 与相同的 $mgf\,M(t)$. 我们在前一节已证明, F^{*n} 是它们的和的分布, $(M(t))^n$ 则得它们的和的 mgf. 本节要把这一结果推广到求和次数也为随机变量的情形.

设 N 是一个仅取非负整数值的随机变量. 记

$$m(t)=\sum_{n=0}^{\infty}e^{tn}p_n \tag{6.1}$$

其中 $p_n=\mathrm{P}\,[N=n]$, $n=0,1,\cdots$. 又设 $X_1,X_2,\cdots$ 是独立同分布的随机变量序列 (以下简记独立同分布为 iid), 并以 F 与 $M(t)$ 分别表示它们的公共 cdf 与 mgf. 再假定诸 X_i 和 N 也是相互独立的, 并记

$$S=X_1+\cdots+X_N \tag{6.2}$$

当 $N=0$ 时, 约定 $S=0$. 以下称 S 为随机和, 并称 N 为求和次数, 而称诸 X_i 为 S 的加项.

【例 6.1】 对于在一给定时间区间内的保单组合来说，以 N 表示 **索赔次数**，再以 X_i 表示 **第 i 次的索赔额**，则 S 等于 **索赔总额**.

为了求出 S 的 *cdf*，可如下运用累次法则：

$$
\begin{aligned}
F_S(s) &= \mathrm{P}[S \le s] = \mathrm{E}\big[\mathrm{P}[S \le s|N]\big] \\
&= \sum_{n=0}^{\infty} \mathrm{P}[S \le s|N=n]p_n = \sum_{n=0}^{\infty} F^{*n}(s)p_n \qquad (6.3)
\end{aligned}
$$

类似地，可得到 mgf 的表示式

$$
\begin{aligned}
M_S(t) &= \mathrm{E}[e^{tS}] = \mathrm{E}\big[\mathrm{E}[e^{tS}|N]\big] \\
&= \mathrm{E}[M(t)^N] = m(\log M(t)) \qquad (6.4)
\end{aligned}
$$

其中 m 是由 (6.1) 式定义的 N 的 mgf.

这样，

$$M_S'(t) = m'\big(\log M(t)\big) \cdot \frac{M'(t)}{M(t)} \qquad (6.5)$$

特别地，

$$M_S'(0) = m'(0)M'(0)$$

这表明

$$\mathrm{E}[S] = \mathrm{E}[N]\mathrm{E}[X] \qquad (6.6)$$

若采用例 6.1 中的语言，上式表明，索赔总额的期望值恰等于索赔次数的期望值与个体索赔额的期望值之乘积 (这一结果并不令人感到惊奇). 将 (6.5) 式再微分一次，并置 $t=0$，可得

$$M_S''(0) = m''(0)M'(0)^2 + m'(0)\{M''(0) - M'(0)^2\} \qquad (6.7)$$

即有

$$\mathrm{E}[S^2] = \mathrm{E}[N^2](\mathrm{E}[X])^2 + \mathrm{E}[N]\ \mathrm{Var}\ [X] \qquad (6.8)$$

最后，在上式两端再分别减去 (6.6) 式两端的平方，便得

$$\mathrm{Var}\,[S] = \mathrm{Var}\,[N](\mathrm{E}\,[X])^2 + \mathrm{E}\,[N]\;\mathrm{Var}\,[X] \tag{6.9}$$

再采用例 6.1 中的语言，由上式即知，索赔总额的方差可分解成两个分量：第一个分量反映了索赔次数是随机的，第二个分量则反映个体索赔额是随机的.

7. 复合 Poisson 分布

本节讨论随机和的特殊情形，即求和次数 N 服从 Poisson 分布的情形. 假设 Poisson 参数为 λ. 这样，由方程 (6.3) 知，

$$F_S(s) = \sum_{n=0}^{\infty} F^{*n}(s)\frac{\lambda^n}{n!}e^{-\lambda} \tag{7.1}$$

这便是著名的 **复合 Poisson 分布**.

将

$$m(t) = \exp\{\lambda(e^t - 1)\}$$

代入 (6.4) 式，即可求出它的 mgf:

$$M_S(t) = \exp\{\lambda[M(t) - 1]\} \tag{7.2}$$

作为应用，考虑两个复合 Poisson 分布. 假设 $F^{(i)}$ 是一复合 Poisson 分布，它的 mgf 为 $M^{(i)}(t), i = 1, 2$. 再假定 λ_i 为 Poisson 参数，F_i 是以 $M_i(t)$ 为 mgf 的诸加项的分布. 显然，$F^{(1)} * F^{(2)}$ 的 mgf 为

$$\begin{aligned} M^{(1)}M^{(2)} &= \exp\{\lambda_1[M_1 - 1]\}\exp\{\lambda_2[M_2 - 1]\} \\ &= \exp\{\lambda[M - 1]\} \end{aligned} \tag{7.3}$$

其中

$$\lambda = \lambda_1 + \lambda_2, \quad M = \frac{\lambda_1 M_1 + \lambda_2 M_2}{\lambda}$$

这表明，复合 Poisson 分布的卷积仍为复合 Poisson 分布. 新的 Poisson 参数是原先的 Poisson 参数的和，新的加项的分布则是原有的加项分布的加权平均，其权数与各自的 Poisson 参数成正比.

本节的剩下部分将讨论如何通过标准的方法,将一个复合 Poisson 分布表示成一系列的卷积. 假设给定一复合 Poisson 分布，其 Poisson 参数为 $\lambda > 0$. (为避开技术上细节) 假定其加项分布 F 是离散的，不连续点为 $x_1, x_2, \cdots$. 令

$$q_i = \mathrm{P}[X = x_i] = F(x_i) - F(x_i - 0), \quad \forall i \geq 1 \tag{7.4}$$

再以 N_i 表示取值为 x_i 的诸加项的随机次数，亦即在 (6.2) 式中，$X_j = x_i$ $(j = 1, \cdots, N)$ 的项数. 显然，有

$$S = x_1 N_1 + x_2 N_2 + \cdots \tag{7.5}$$

不过，下述两事实则未必是明显的； (a) N_i 服从以 λq_i 为参数的 Poisson 分布； (b) 诸随机变量 $N_1, N_2, \cdots$ 是相互独立的.

为证明 (a), 首先注意到， (给定 N 后) N_i 服从以 N 与 q_i 为参数的二项分布. 这样，对于 $n = 0, 1, 2, \cdots$, 我们有

$$\begin{aligned}
\mathrm{P}[N_i = n] &= \mathrm{E}\left[\mathrm{P}[N_i = n | N]\right] \\
&= \sum_{m=n}^{\infty} \binom{m}{n} q_i^n (1 - q_i)^{m-n} \frac{\lambda^m}{m!} e^{-\lambda} \\
&= \frac{e^{-\lambda}}{n!} (\lambda q_i)^n \sum_{m=n}^{\infty} \frac{1}{(m-n)!} (1 - q_i)^{m-n} \lambda^{m-n} \\
&= \frac{e^{-\lambda}}{n!} (\lambda q_i)^n e^{\lambda(1-q_i)} = e^{-\lambda q_i} \frac{(\lambda q_i)^n}{n!}
\end{aligned} \tag{7.6}$$

为证明 (b), 首先观察到 $x_i N_i$ 的 mgf 等于

$$M_i(t) = \exp\{\lambda q_i (e^{t x_i} - 1)\} \tag{7.7}$$