应用随机过程

柳金甫　李学伟　编著

YINGYONG SUIJI GUOCHENG

中国铁道出版社

应用随机过程

柳金甫　李学伟　编著

中国铁道出版社

2000年·北京

（京）新登字 063 号

内 容 简 介

这是一本随机过程应用方面的教科书，是作者在多年教学实践与应用研究的基础上，整理出的一种面向工科、经济管理等专业的基础知识、基本理论和应用方法，包括随机过程的基本概念、随机分析、POISSON 过程、更新过程、MARKOV 链、BROWN 运动和平衡过程的谱分析及其应用等。

本书适用于电子工程、通信、经济、管理科学与工程等专业的研究生教学，也可供有关工程技术人员参考。

图书在版编目（CIP）数据

应用随机过程/柳金甫，李学伟编著．—北京：中国铁道出版社，2000.7
ISBN 7-113-03673-2

Ⅰ．应…　Ⅱ．①柳…　②李…　Ⅲ．随机过程-应用　Ⅳ．0211.6

中国版本图书馆 CIP 数据核字（2000）第 13971 号

书　　名：应用随机过程
作　　者：柳金甫　李学伟
出版发行：中国铁道出版社（100054，北京市宣武区右安门西街 8 号）
策划编辑：殷小燕
责任编辑：殷小燕
封面设计：李艳阳
印　　刷：北京市彩桥印刷厂
开　　本：787×1092　1/16　印张：13.5　字数：335 千
版　　本：2000 年 7 月第 1 版　　2000 年 7 月第 1 次印刷
印　　数：1～2500 册
书　　号：ISBN 7-113-03673-2/O・74
定　　价：25.80 元

序　言

本书是为学习工程、物理及应用数学的学生编写的一本随机过程导论性的书，包括了应用随机过程的主要内容。给出了随机过程理论所需的有关概率论的基本概念，并用较大篇幅论述了科技工作者非常关心的随机过程的理论、方法及应用。使读者既能了解该理论的基本内容，又能学到解决问题的思路与技巧。由于目前国内应用概率研究正在发展阶段，本书的出版是必要的。

本书是作者近年来在应用随机过程、排队论等课程的教学实践和部分研究工作的基础上写成的，大部分材料在北方交通大学的有关专业研究生、本科生及进修教师的教学中多次使用过。可作为理工科大学有关专业本科生、研究生的应用随机过程、随机模型等课程的教科书或参考书。对有兴趣的实际工作者，也能从本书中获得所需要的知识。

一般地说，初等概率论研究随机现象的静态性质，而随机过程论是研究随机现象的动态特性，研究随机现象发展过程的数量关系。初等概率论是随机过程的基础，而随机过程则是初等概率论的自然延伸。目前，随机过程论已发展成为内容十分丰富、应用极为广泛的一门数学分支，成为广大自然科学工作者、工程技术人员和社会、经济学家乐于应用的数学工具之一，并显示出其越来越重要的作用。

虽然本书不可能全面详尽地介绍随机过程论的一切方面，但作者在编写时尽可能对随机过程的基本概念和一些应用作了较为详细的论述。希望本书的出版对读者有较大的帮助。

王梓坤

前　　言

本书的初稿是作者给北方交通大学工科研究生(如信息工程学院、土建学院等)以及应用数学专业的本科生开的随机过程课的讲义。在多年的教学实践中,不断对教材充实改进,同时,广泛征求了各专业研究生导师及研究生们的意见,突出了在相关专业的应用特色,最后成为现在这本书。

在此,作者深深地感谢中科院院士王梓坤教授、陈希孺教授对本书编写的关怀和指导,感谢我的本校同仁高自友教授、常彦勋教授以及汪成咏、吴法恩、刘晓、郑神州、王兵团、付俐、王玲凤、渠刚莱副教授,他们关心本书的出版,并提出不少修改意见。感谢我的研究生桂文豪同学对其中部分书稿的审校。

在本书出版过程中,铁道出版社的殷小燕编辑严谨的工作作风和高投入的工作精神得以使本书顺利出版,特表感谢!

由于作者水平所限,书中不当之处敬请读者指正。

作　者

一九九九年十月于北方交通大学

目　　录

第 0 章　基础知识

§1　随机变量及其分布函数、密度函数

仅讨论随机事件及其概率，只能孤立地研究随机试验的一个或几个事件，对于随机现象的数学分析显然是不够的。为了更深入地研究随机现象，需要把随机试验的结果数量化，就是说用一个变量来描述随机试验的结果。我们首先引入随机变量(random variable)，要讨论随机变量，就必须对其给以一定的限制或规定。

设 Ω 是联系于某一随机试验的**基本事件空间**(或称**样本空间**)，Ω 中的元素(或者说"点")就是描述这试验的基本事件，即试验的可能结果。空间 Ω 的某些子集称**事件**。我们用 $\mathscr{F}$ 表示事件的全体，即 $\mathscr{F}$ 是一族满足一定条件的 ω 点集① $(\omega\in\Omega)$，我们称二元组 $(\Omega,\mathscr{F})$ 为**可测空间**。在 $\mathscr{F}$ 上定义一个集函数 $P(\cdot)$ 以度量 $\mathscr{F}$ 中事件发生的可能性大小，它满足：

(1)非负性：$0\leqslant P(A)\leqslant 1$，对任意 $A\in\mathscr{F}$；

(2)规范性：$P(\Omega)=1$；

(3)可列可加性：若 $A_1,A_2,\cdots,A_n,\cdots\in\mathscr{F}$，且 $A_i\cap A_j=\phi,i\neq j,i、j=1,2\cdots$ 则

$$P(\bigcup_{n=1}^{\infty}A_n)=\sum_{n=1}^{\infty}P(A_n)。$$

我们称 $P(A)$ 为事件 A 的概率，称三元组 $(\Omega,\mathscr{F},P)$ 为**概率空间**。

1.1 定义

给定一个随机试验 E，它的结果 ω 是样本空间 Ω 的元素，$(\Omega,\mathscr{F},P)$ 是相应的概率空间。如果对于每一 ω 有一个实数 $X(\omega)$ 与之对应，就得到定义在 Ω 上的实值点函数 $X(\omega)$，若对于任意实数 x，集 $\{\omega:X(\omega)\leqslant x\}$ 是 $\mathscr{F}$ 中的事件，即

$$(\omega:X(\omega)\leqslant x)\in\mathscr{F}$$

则称 $X(\omega)$ 为随机变量。

从定义看出，随机变量 $X(\omega)$ 总是联系着一个概率空间，意即关系式 $X\triangleq X(\omega)$ 将概率空间 Ω 中的每一元素 ω 映射到实轴 $R=(-\infty,\infty)$ 上的点 x，建立起样本空间 Ω 与实数(或复数)空间或其一部分的对应关系。习惯上，不必每次都写出概率空间 $(\Omega,\mathscr{F},P)$，并将 $X(\omega)$ 关于对 ω 的依赖性省略，简记为 X，$(\omega:X(\omega)\leqslant x)$ 简记为 $(X\leqslant x)$。另一方面，由于要求 $(X\leqslant x)\in\mathscr{F}$，因此 $P(X\leqslant x)$ 总是有意义的。

对随机变量概念的引入是概率论的重大进展，它使概率研究的对象由事件扩大为随机变量。

---

① 确切地说，$\mathscr{F}$ 是 Ω 的某些子集构成的 $\mathscr{F}$ 上的 σ 代数或 σ 域，即 $\mathscr{F}$ 中元满足

(1) $\Omega\in\mathscr{F}$；

(2)若 $A\in\mathscr{F}$，则 $\overline{A}=\Omega-A\in\mathscr{F}$；

(3)若 $A_1,A_2,\cdots,A_n,\cdots\in\mathscr{F}$，则 $\bigcup_{n=1}^{\infty}A_n\in\mathscr{F}$。

随机变量按其取值情形可分为两大类:离散型随机变量和连续型随机变量。

1.2 定义

随机变量 X 称离散的,如果存在一有限或可列无穷的数列$\{x_i\}$,且当 $i\neq j$ 时有 $x_i\neq x_j$,使得

$$P(X=x_i)=p_i>0$$

且
$$\sum_i p_i=1$$

此时 X 的分布函数定义为

$$F(X)=P(X\leqslant x)=\sum_{x_i\leqslant x}p_i \qquad x\in R^{①} \tag{0.1}$$

1.3 定义

随机变量 X 称为**连续型**的,如果存在一定义在 R 上的非负可积函数$f(x)$,使对任意 $x\in R$,其分布函数可表示为

$$F(x)=\int_{-\infty}^{x}f(y)\mathrm{d}y \tag{0.2}$$

且
$$\int_{-\infty}^{\infty}f(x)\mathrm{d}x=1$$

由积分性质容易推知,对任意 $-\infty\leqslant a<b<\infty$ 有

$$P(a<x\leqslant b)=\int_{b}^{a}f(x)\mathrm{d}x$$

注意对任意 $a\in R$ 有 $P(x=a)=\lim\limits_{n\to\infty}P(a-\frac{1}{n}<x\leqslant a)=\lim\limits_{n\to\infty}\int_{a-\frac{1}{n}}^{a}f(x)\mathrm{d}x=0$,即连续型随机变量取任一固定值的概率等于零,我们把$f(x)$称做连续型随机变量 X 的分布密度函数或频率函数。

§2 随机变量的数学期望(或均值)和方差

随机变量 x 的**数学期望**或**均值**记作 EX,它(如果存在的话)由下式定义:

$$EX=\int_{-\infty}^{\infty}x\mathrm{d}F(x)=\begin{cases}\sum\limits_i x_if(x_i)=\sum\limits_i x_iP_i & \text{离散型时,}\\ \int_{-\infty}^{\infty}xf(x)\mathrm{d}x & \text{连续型时。}\end{cases}$$

我们说随机变量 X 的均值存在,如果上面定义中的无穷级数或无穷积分是绝对收敛的。

随机变量 X 的**方差**定义为

$$\mathrm{Var}X=E(X-EX)^2=EX^2-(EX)^2$$

式中 X 的**标准差**定义为其方差的算术平方根$\sqrt{\mathrm{Var}X}$。

EX 表 X 取值的集中地点,为随机变量 X 的位置特征,它是概率论中一个极其重要的概念。实际上,在概率论各类数字特征中都是由数学期望定义的。

---

① 若用 δ 函数把$f(x)\triangleq\sum_i p_i\delta_{x_i}(x)$,称其为随机变量 x 的密度函数或质量函数,此时分布函数可表为 $F(x)=\int_{-\infty}^{x}f(y)\mathrm{d}y$,这里 $\delta_{x_i}(x)=1$ 若 $x=x_i$;$\delta_{x_i}(x)=0$,若 $x\neq x_i$,因此类似于连续情形也把质量函数称做密度函数。

方差 $\mathrm{Var}X$ 是用来表征随机变量 X 偏离其均值 EX 的分散程度的。$\mathrm{Var}X$ 越大，其取值的分散程度越大，反之，$\mathrm{Var}X$ 取值越小，其取值就越集中，特别，当 $\mathrm{Var}X=0$ 时，则有 X 几乎处处取某个常数(即 EX)。

一般地，我们把 X^k 的数学期望 $v_k=EX^k$ 称做 X 的 **k 阶原点矩**，把 $(X-EX)^k$ 的数学期望 $\mu_k=E(X-EX)^k$ 称为 X 的 **k 阶中心矩**，这里 k 是任意正整数。应当指出的是，一个随机变量的矩不一定存在。但如果对某一正整数 r，它的 r 阶矩存在，则对任意正整数 $k\leqslant r$，它的 k 阶矩也一定存在。显然，数学期望是 X 的一阶原点矩，$\mathrm{Var}X$ 则是 X 的二阶中心矩。

2.1 命题 设 X 是取非负整数值的离散型随机变量，则

$$EX=\sum_{k=1}^{\infty}P(X\geqslant k) \tag{0.3}$$

事实上，由定义

$$\begin{aligned}EX=\sum_{k=1}^{\infty}kP(X=k)=P(X=1)&\\ &+P(X=2)+P(X=2)\\ &+P(X=3)+P(X=3)+P(X=3)\\ &+\cdots\cdots\\ &+P(X=k)+P(X=k)+\cdots+P(X=k)\\ &+\cdots\cdots \qquad\qquad\qquad (\text{共 } k \text{ 项})\end{aligned}$$

最后对每一纵向列先求和即得。

2.2 命题 设 X 是取非负值的随机变量，则

$$EX=\int_0^{\infty}(1-F(x))\mathrm{d}x \tag{0.4}$$

这里 EX 或者有限，或者等于 ∞。

证 由分部积分公式，(0.4)式右端等于

$$\lim_{x\to\infty}x(1-F(x))+\int_0^{\infty}x\mathrm{d}F(x)$$

上式中第二项即为 EX，当 $EX<\infty$ 时，只需证第一项等于零。事实上

$$\begin{aligned}\lim_{x\to\infty}x(1-F(x))&=\lim_{x\to\infty}x\int_x^{\infty}\mathrm{d}F(y)\\ &\leqslant\lim_{x\to\infty}\int_x^{\infty}y\mathrm{d}F(y)=0\end{aligned}$$

若 $EX=\infty$，则由 $\lim\limits_{x\to\infty}x(1-F(x))\geqslant 0$，故(0.4)两端相等。

如果除去对 x 的非负性要求，我们可以利用类似的方法证明若 EX 存在，则有

$$EX=\int_0^{\infty}(1-F(x))\mathrm{d}x-\int_{-\infty}^{0}F(x)\mathrm{d}x。$$

§3 随机向量及其概率分布

当我们要同时考虑两个或更多个随机变量时，就需研究这些随机变量的联合分布，这里主要对二维随机向量 (X,Y) 展开讨论，对于变量数目大于 2，即 $(X_1,X_2,\cdots X_n)$ $(n>2)$ 情形，讨论是类似的。

设 X 和 Y 是任意两个随机变量，X 和 Y 的(二维)**联合概率分布**(简称联合分布)定义为

$$F(x,y)=P(X\leqslant x,Y\leqslant y),\ -\infty<x,y<\infty$$

X 和 Y 的边沿分布 $F_X(x)$，$F_Y(y)$ 可以从它们的联合分布求出：

$$F_X(x)\triangleq P(X\leqslant x)=\lim_{y\to\infty}P(X\leqslant x,Y\leqslant y)\triangleq F(x,\infty)$$

$$F_Y(y)\triangleq P(Y\leqslant y)=\lim_{x\to\infty}P(X\leqslant x,Y\leqslant y)\triangleq F(\infty,y)$$

当 (X,Y) 为离散型时，它们的联合密度函数(或称联合质量函数)是

$$f(x,y)=P(X=x,Y=y)$$

X、Y 的边沿密度(质量)函数依次是

$$f_X(x)=\sum_y f(x,y)$$

$$f_Y(y)=\sum_x f(x,y)$$

(X,Y) 为连续型随机向量，如果存在一个定义在 $R\times R$ 上的非负可积函数 $f(x,y)$，使对一切 $x,y\in R$ 有

$$F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}f(u,v)\mathrm{d}u\mathrm{d}v$$

并且称 $f(x,y)$ 为 (X,Y) 的联合密度函数，X 和 Y 的边沿密度可由 (X,Y) 的联合密度求出：

$$f_X(x)=\int_{-\infty}^{\infty}f(x,y)\mathrm{d}y$$

$$f_Y(y)=\int_{-\infty}^{\infty}f(x,y)\mathrm{d}x$$

进而有

$$F_X(x)=\int_{-\infty}^{x}f_X(u)\mathrm{d}u=\int_{-\infty}^{x}\int_{-\infty}^{\infty}f(u,y)\mathrm{d}u\mathrm{d}y$$

$$F_Y(y)=\int_{-\infty}^{y}f_Y(v)\mathrm{d}v=\int_{-\infty}^{y}\int_{-\infty}^{\infty}f(x,v)\mathrm{d}x\mathrm{d}v$$

随机变量 X、Y 称为**相互独立**的，如果对任意实数 x、y 有

$$F(x,y)=F_X(x)F_Y(y)$$

此时，无论 (X,Y) 为离散型的或连续型的，其联合密度函数均可表为

$$f(x,y)=f_X(x)f_Y(y),\qquad \forall\ x,y\in R$$

(在 (X,Y) 为连续型时，只要对几乎所有的 x,y(Lebesgue 测度)上式成立即可)

两个随机变量 X,Y 的**协方差** $\mathrm{Cov}(X,Y)$ 定义为

$$\begin{aligned}\mathrm{Cov}(X,Y)&=E\{(X-EX)(Y-EY)\}\\&=E(XY)-EX\cdot EY\end{aligned}$$

X、Y 的**相关系数** ρ_{XY} 定义为

$$\rho_{XY}=\frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}X\cdot\mathrm{Var}Y}}$$

当 X、Y 相互独立时，易证 $EXY=EXEY$，从而有 $\mathrm{Cov}(X,Y)=0$。需要指出的是，由 $\mathrm{Cov}(X,Y)=0$ 一般不能推出 X、Y 相互独立。

容易验证，运算“E”满足线性：对任意两随机变量 X、Y 都有

$$E(aX+bY)=aEX+bEY$$

这里 a、b 为常数，但运算“Var”一般并不具有这样的线性性质，

事实上
$$\mathrm{Var}(X+Y)=\mathrm{Var}X+\mathrm{Var}Y+2\mathrm{Cor}(X,Y)$$
$$\mathrm{Var}(aX)=a^2\mathrm{Var}X$$

可以把二维分布的概念直接推广到任意有限多维随机向量的情况。此时 n 维随机向量 $(X_1,X_2,\cdots,X_n)$ 的联合分布函数为
$$F(X_1,\cdots,X_n)\triangleq P(X_1\leqslant x_1,\cdots,X_n\leqslant x_n)$$
它表示事件 $X_1\leqslant x_1,\cdots,X_n\leqslant x_n$ 同时出现的概率,类似于二维的情形,它具有下列性质:

(1) $F(x_1,\cdots,x_{i-1},-\infty,x_{i+1},\cdots,x_n)=0$
$$i=1,2,\cdots,n;$$
(2) $F(\infty,\infty,\cdots,\infty)=1$;

(3)若 $F(x_1,\cdots x_n)$ 的偏导数
$$f(x_1,\cdots,x_n)=\frac{\partial F(x_1,\cdots,x_n)}{\partial x_1\cdots\partial x_n}$$
存在,则称 $f(x_1,\cdots,x_n)$ 为 n 维联合密度函数,从而有
$$F(x_1,\cdots,x_n)=\int_{-\infty}^{x_1}\cdots\int_{-\infty}^{x_n}f(x_1,\cdots,x_n)\mathrm{d}x_1\cdots\mathrm{d}x_n$$
对于可测集 $G_n\subset R^n$,随机点落在 G_n 的概率为
$$P((x_1,\cdots,x_n)\in G_n)=\int\underset{G_n}{\cdots}\int f(x_1,\cdots,x_n)\mathrm{d}X_1\cdots\mathrm{d}x_n。$$

在理论或实际中,常常需要把一个随机变量或随机向量通过某种变换变为一个新的随机变量或随机向量。这时一个重要的问题是要求出新的变量或向量的分布函数或分布密度。考察线性变换 $Y=aX+b$ 的简单情形。设 X 的分布函数为 $F_x(x)$,$a\neq0$,则 Y 的分布函数是
$$F_Y(y)=P(aX+b\leqslant y)=P(aX\leqslant y-b)$$
$$=\begin{cases}P(X\leqslant\frac{y-b}{a})=F_X(\frac{y-b}{a}) & \text{当 } a>0\\ P(X\geqslant\frac{y-b}{a})=1-F_X(\frac{y-b}{a}-0) & \text{当 } a<0\end{cases}$$
若 $F_X(x)$ 具有密度函数 $f_X(x)$,则 Y 的密度函数 $f_Y(y)$ 存在,而且有
$$f_Y(y)=f_X(\frac{y-b}{a})\cdot\frac{1}{|a|}$$
较为一般的结果有

3.1 定理 设 n 维随机向量 $X=(X_1,\cdots,X_n)$ 有密度函数 $f_X(x_1,\cdots,x_n)$,又设 n 元函数 $y_i=u_i(x_1,\cdots,x_n)$,$i=1,\cdots,n$ 满足:

(1)存在唯一的反函数 $x_i=v_i(y_1,\cdots,y_n)$,$i=1,\cdots,n$

(2) $u_i(x_1,\cdots,x_n)$ 和 $v_i(y_1,\cdots,y_n)$ 都连续且有一阶连续偏导数,若以 J 表示变换 $y_i=u_i(x_1,\cdots,x_n)$ 的雅可比行列式:
$$J=\begin{vmatrix}\frac{\partial x_1}{\partial y_1}\cdots\frac{\partial x_1}{\partial y_n}\\ \vdots \qquad \vdots\\ \frac{\partial x_n}{\partial y_1}\cdots\frac{\partial x_n}{\partial y_n}\end{vmatrix}$$
则分量由 $Y_i=u_i(X_1,\cdots,X_n)$,$i=1,\cdots,n$ 给定的 n 维随机向量 $Y=(Y_1,\cdots,Y_n)$ 有密度函数

$$f_Y(y_1,\cdots,y_n)=|J|\cdot f_X(x_1(y_1,\cdots,y_n),\cdots,x_n(y_1,\cdots,y_n)) \tag{0.5}$$

作为这一定理的应用,再次考虑线性变换 $Y=ax+b$,这时 $n=1$ 和 $X=\dfrac{Y-b}{a}$,于是有 $|J|=\left|\dfrac{\partial x}{\partial y}\right|=\dfrac{1}{|a|}$ 和 $f_X\left(\dfrac{y-b}{a}\right)$,故由(0.5)式又一次得到 $f_Y(y)=f_X\left(\dfrac{y-b}{a}\right)\dfrac{1}{|a|}$。

§4 矩母函数和概率生成函数

X 的矩母函数定义为

$$\Psi(t)=Ee^{tX}=\int_{-\infty}^{\infty}\mathrm{e}^{tx}\mathrm{d}F(x)$$

对 Ψ 逐次求导并计算在 $t=0$ 点的值能得到 X 的各阶矩,即

$$\Psi'(t)=E(X\mathrm{e}^{tX})$$
$$\Psi''(t)=E(X^2\mathrm{e}^{tX})$$
$$\cdots\cdots\cdots\cdots$$
$$\Psi^{(n)}(t)=E(X^n\mathrm{e}^{tX})$$

计算在 $t=0$ 点的值得

$$\Psi^{(n)}(0)=EX^n \qquad n\geqslant 1$$

这里应注意我们已假定求导与积分运算可交换是合理的。通常遇到的情形都是这样的。

当矩母函数存在时,它唯一地决定概率分布。这是十分重要的,因此这使我们能够用矩母函数刻划随机变量的概率分布。见表 0.1。

4.1 例 设 X 与 Y 是独立的正态随机变量,各自的均值为 μ_1 与 μ_2,方差为 σ_1^2 与 σ_2^2,它们和的矩母函数由下式给出

$$\begin{aligned}\Psi_{X+Y}(t)&=E\mathrm{e}^{t(X+Y)}\\&=E\mathrm{e}^{tX}E\mathrm{e}^{tY}\qquad(\text{由独立性})\\&=\Psi_X(t)\Psi_Y(t)\\&=\exp\{(\mu_1+\mu_2)t+(\sigma_1^2+\sigma_2^2)t^2/2\}\end{aligned}$$

其中最后一个等式来自表 0.2,于是 $X+Y$ 的矩母函数是均值为 $\mu_1+\mu_2$,方差为 $\sigma_1^2+\sigma_2^2$ 的正态随机变量的矩母函数。由唯一性,这就是 $X+Y$ 的概率分布。

表 0.1

离散概率分布	概率质量函数 $p(x)$	矩母函数 $\phi(t)$	均值	方差
二项分布参数 $n,p,0\leqslant P\leqslant 1$	$\binom{n}{x}p^x(1-p)^{n-x}$ $x=0,1,\cdots,n$	$[pe^t+(1-p)]^n$	np	$np(1-p)$
泊松分布 参数 $\lambda>0$	$\mathrm{e}^{-\lambda}\dfrac{\lambda^x}{x!}$ $x=0,1,2,\cdots$	$\exp\{\lambda(\mathrm{e}^t-1)\}$	λ	λ
几何分布 参数 $0\leqslant p\leqslant 1$	$p(1-p)^{x-1}$ $x=1,2,\cdots$	$\dfrac{p\mathrm{e}^t}{1-(1-p)\mathrm{e}^t}$	$\dfrac{1}{p}$	$\dfrac{1-p}{p^2}$
负二项分布 参数 r,p	$\binom{x-1}{r-1}p^r(1-p)^{x-r}$ $x=r,r+1,\cdots$	$\left[\dfrac{p\mathrm{e}^t}{1-(1-p)\mathrm{e}^t}\right]^r$	$\dfrac{r}{p}$	$\dfrac{r(1-p)}{p^2}$

表 0.2

连续概率分布	概率密度函数 $f(x)$	矩母函数 $\phi(t)$	均值	方差
(a,b)上的均匀分布	$\frac{1}{b-a}, a<x<b$	$\frac{e^{tb}-e^{ta}}{t(b-a)}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
指数分布 参数 $\lambda>0$	$\lambda e^{-\lambda x}, x\geqslant 0$	$\frac{\lambda}{\lambda-t}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Γ-分布 参数$(n,\lambda)\lambda>0$	$\frac{\lambda e^{-\lambda x}(\lambda x)^{n-1}}{(n-1)!}$ $x\geqslant 0$	$\left(\frac{\lambda}{\lambda-t}\right)^n$	$\frac{n}{\lambda}$	$\frac{n}{\lambda^2}$
正态分布 参数(μ,σ^2)	$\frac{1}{\sqrt{2\pi}\sigma}e^{-(x-\mu^2)/2\sigma^2}$ $-\infty<x<\infty$	$\exp\{\mu t+\frac{\sigma^2t^2}{2}\}$	μ	σ^2
B-分布 参数 a,b, $a>0,b>0$	$cx^{a-1}(1-x)^{b-1}$ $0<x<1$ $c=\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}$		$\frac{a}{a+b}$	$\frac{ab}{(a+b)^2/(a+b+1)}$

同矩母函数一样，概率生成函数是研究离散型随机变量的有效工具。设 X 是一具有概率分布

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

的离散型随机变量，则 X 的概率生成函数定义为

$$G(s)=\sum_{k=0}^{\infty}p_ks^k$$

这里 $p_k\geqslant 0$, $\sum_k p_k=1$。X 是取非负整数值随机变量，上式等价于

$$G(s)=Es^X$$

有时也称 $G(s)$为 X 的概率母函数。

易见，每一概率分布$\{p_k\}$对应唯一的概率母函数 $G(s)$；反之，每一概率母函数 $G(s)$确定唯一的概率分布$\{p_k\}$。事实上，每一 p_k 皆可由 $G(s)$通过

$$p_k=\frac{1}{k!}\left(\frac{d^kG(s)}{ds^k}\right)_{s=0} \qquad k=0,1,2,\cdots$$

给出。若对 $G(s)$相继求导，则可对$|s|<1$有

$$G'(s)=\sum_{k=1}^{\infty}kp_ks^{k-1}$$

$$G''(s)=\sum_{k=2}^{\infty}k(k-1)p_ks^{k-2}$$

故当 X 是非负整数值随机变量时有

$$EX=\sum_{k=1}^{\infty}kp_k=G'(1)$$

$$E(X(X-1))=\sum_{k=2}^{\infty}k(k-1)p_k=G''(1)$$

从而得

$$EX^2 = E(X(X-1)) + EX = G''(1) + G'(1)$$
$$\mathrm{Var}X = EX^2 - (EX)^2 = G''(1) + G'(1) - (G'(1))^2$$

一般地，X 的 k 阶阶乘矩由下式给出：

$$E[X(X-1)\cdots(X-k+1)] = \left(\frac{\mathrm{d}^kG(s)}{\mathrm{d}s^k}\right)_{s=1}$$

下面给出母函数的一些重要性质。设$\{a_k\}$、$\{b_k\}$是两实数序列，令

$$c_k = a_0b_k + a_1b_{k-1} + \cdots + a_kb_0$$

则数列$\{c_k\}$称做$\{a_k\}$和$\{b_k\}$的**卷积**，并记作

$$\{c_k\} = \{a_k\} * \{b_k\}$$

现设 X 和 Y 是两相互独立的非负整数值随机变量，分别具有概率分布$\{p_i\}$、$\{q_j\}$。于是它们的联合分布由 $P(X=i, Y=j) = p_iq_j$ 给出。而 $X = X + Y$ 的概率分布则由

$$r_k \triangleq P(z=k) = \sum_{i=0}^{k} p(X=i, Y=k-1)$$
$$= \sum_{i=0}^{k} p_iq_{k-i}$$

给出。即$\{r_k\}$是$\{p_k\}$和$\{q_k\}$的卷积。这一结果可以推广到任意有限多个独立随机变量的情形。

4.2 定理　n 个相互独立的非负整数值随机变量 $X_1, X_2, \cdots, X_n$ 之和 $Z \triangleq X_1 + X_2 + \cdots + X_n$ 的概率母函数 $G_z(s)$等于 $X_1, X_2, \cdots, X_n$ 的概率母函数 $G_1(s), \cdots, G_n(s)$之积：

$$G_z(s) = G_1(s)G_2(s)\cdots G_n(s)$$

证明：

$$G_z(s) = Es^z = E(s^{X1+\cdots+Xn}) = E(s^{X1}\cdots s^{Xn})$$
$$= Es^{X1}\cdots Es^{Xn} = G_1(s)\cdots G_n(s)$$

特别，若 $X_1, \cdots, X_n$ 独立同分布，具有同一母函数 $G(s)$时，有

$$G_z(s) = (G(s))^n$$

4.3 定理　设 X 是取非负整数值的随机变量，其概率母函数是 $P(s)$，又设 m、n 是两个非负整数，且 $m \neq 0$，则随机变量 $Y = mX + n$ 的概率母函数

$$Q(s) = s^nP(s^m)$$

证明　由 Y 的定义及有关数学期望的性质有

$$Q(s) = ES^Y = ES^{mX+n} = s^nES^{mX}$$
$$= s^nP(s^m)$$

§5　Lapace 变换和 Laplace – Stieltjes 变换

Laplace 变换是一种积分变换，可视为上节概率母函数的一种推广。这种变换适于用来研究一般的非负随机变量。

设 $f(t)$是任意定义在 $R_+ = [0, \infty)$上的函数，我们称由

$$\tilde{f}(s) = \int_0^\infty \mathrm{e}^{-st}f(t)\mathrm{d}t \tag{0.6}$$

定义的函数$\tilde{f}(t)$为$f(t)$的**Laplace 变换**（简记为 L 变换），$\tilde{f}(s)$有时也记作$\mathscr{L}\{f(t)\}$。当我们使用$\tilde{f}(s)$时，强调它是参数 s 的函数，而使用$\mathscr{L}\{f(t)\}$时则强调它是$f(t)$的 L 变换。显然，我们只能对那些使得(0.6)式积分存在的 s 值讨论。需要指出的是，在 Laplace 变换的一般理论中，参数 s 是取值于复数域的，但在本书中的应用通常只须考虑非负实值参数 s。

L 变换的某些重要性质

(1)线性性质

$$\mathscr{L}\left\{\sum_{i=1}^{n} c_i f_i(t)\right\} = \sum_{i=1}^{n} c_i \tilde{f}_i(s)$$

其中 n 是任意正整数，c_i 是任意常数，如 $\mathscr{L}\{\cosh at\} = \mathscr{L}\left\{\frac{e^{at}+e^{-at}}{2}\right\} = \frac{s}{s^2-a^2}$

(2)平移性质

$$\mathscr{L}\{e^{-at}f(t)\} = \tilde{f}(s+a)$$

如由 $\mathscr{L}\{t\} = \frac{1}{s^2}$ 有 $\mathscr{L}\{e^{-at}t\} = \frac{1}{(s+a)^2}$ 。若

$$g(t) = \begin{cases} f(t-a) & t > a, \\ 0 & t \leqslant a, \end{cases}$$

则 $\mathscr{L}\{g(t)\} = e^{-as}\tilde{f}(s)$

(3)尺度性质

$$\mathscr{L}\{f(at)\} = \frac{1}{a}\tilde{f}\left(\frac{s}{a}\right)$$

如

$$\mathscr{L}\{\sin at\} = \frac{1}{a}\left[\left(\frac{s}{a}\right)^2 + 1\right]^{-1} = \frac{a}{s^2+a^2}$$

(4)导数的 L 变换

$$\mathscr{L}\{f'(t)\} = s\tilde{f}(s) - f(0)$$

(5)积分的 L 变换

$$\mathscr{L}\left\{\int_0^t f(x)\,dx\right\} = \frac{1}{s}\tilde{f}(s)$$

(6)极限性质

$$\lim_{t\to\infty} f(t) = \lim_{s\to 0} s\tilde{f}(s)$$
$$\lim_{t\to 0} f(t) = \lim_{s\to 0} s\tilde{f}(s)$$

若上面的极限存在。

设 X 是一非负随机变量，分布函数为 $F(x)$。若 X 有密度函数$f(x)$，则我们能够借助密度函数的 L 变换研究随机变量的概率分布，其 L 变换为

$$\mathscr{L}\{f(x)\} = \int_0^\infty e^{-sX} f(x)\,dx$$

但是，如果 X 的密度函数不存在，则就无法使用 L 变换这一有力工具。为此我们可以对分布函数定义类似的变换

$$\mathscr{L}-\mathscr{S}\ \{F(x)\} \triangleq Ee^{-sX} = \int_0^\infty e^{-sX}\,dF(x) \tag{0.7}$$

式中的积分是 Stieltjes 积分，当 $s \geqslant 0$ 时这积分一定存在。我们把由上式定义的变换称为分布函数 $F(x)$（或随机变量 X）的 **Laplace - Stieltjes 变换**（简记为 L - S 变换）。