高等学校教学用书

随机过程与随机微分方程

SUIJIGUOCHENGYUSUIJIWEIFENFANGCHENG

武宝亭 李庆士 杨跃武 编著

$$dX(t)=(A(t)X(t)+a(t))dt+\sum_{i=1}^{m}(D_i(t)X(t)+d_i(t))dW_i(t)$$

Wiener

$$W(t)=(W_1(t),W_2(t),\cdots\cdots,W_n(t))$$

随机过程与随机微分方程

电子科技大学出版社

高等学校教学用书

随机过程与随机微分方程

武宝亭　李庆士　杨跃武　编著

电子科技大学出版社
· 1994 ·

(川)新登字 016

随机过程与随机微分方程

武宝亭　李庆士　杨跃武编著

*

电子科技大学出版社出版

(中国成都建设北路二段四号)

电子科技大学出版社激光照排中心照排

四川省平武县印刷厂胶印

四 川 省 新 华 书 店 经 销

*

开本　787×1092　1/16　印张 14.625　字数 348 千字

版次 1993 年 6 月第一版　印次 1994 年 2 月第二次印刷

印数 1001～2000 册

中国标准书号　ISBN7－81016－633－6/O・16

定价(压膜):8.80 元

304－314

335(1)

序

数学作为一门基础学科和一种科学语言，一直在工程技术和其他科学中起着重要的作用。随着科学技术向高级、精密及尖端发展，随着所研究的对象和系统愈来愈大型、复杂，现代的工程技术已经进入到“数理工程技术”的新时代。这个时代的特征，是数学方法已成为现代工程技术的一个有机的组成部分，并起着举足轻重、甚至是决定性的作用。“高技术本质上是一种数学技术”的观点，已为愈来愈多的人们所逐渐认识。另一方面，工程技术科学的飞速进步，特别是很多高技术、新技术的兴起，也对数学提出了挑战，使广大数学工作者面临着许多急待解决的新课题，从而有力地推动着数学学科本身的发展。

面对这样的形势和发展趋势，如何进一步加强广大工科学生的数学基础，提高他们的数学水平，使他们更多地掌握并学会运用在工程科学技术中行之有效的种种现代数学方法，已成为工科教育中的一个重要的问题。这不仅是关系到所培养的工科人才的质量，而且将会在今后相当长的一段时间中直接影响到我国工业现代化的进程。

现在的这一本书，作为工科硕士生的教材，应该说是在这方面的一个有益的尝试。它根据目前我国工科学生的实际数学基础及今后工作的需要，从比较低的起点出发，由浅入深，循序渐进，在不太大的篇幅中比较系统地介绍了随机过程及随机微分方程的一些基本内容，具有较好的可读性及较宽的适应面。它曾在工业大学为力学、地震工程、土木建筑、环境工程及电类等专业的硕士生试用过几次，并收到了良好的效果。它的正式出版，不仅是作者过去工作的一个可喜的结晶，也相信将会对促进工科数学的教学改革起到积极的作用，特为之序。

李大潜

1991年4月28日

于复旦大学

目　录

第一章　概率论提要及若干补充

§1　事件的概率

概率论与数理统计研究的对象是随机现象的统计规律性。

一个试验，如果可以在相同的条件下不断重复进行，每次试验的可能结果有许多个，每次试验前不能确定哪个结果会出现，但一切可能出现的结果是事先知道的。这样的试验称作随机试验。

在随机试验中，对一次试验可能出现也可能不出现的事件称为随机事件（简称事件）。特别地，每个可能出现的结果也是一个随机事件，称为基本事件。所有基本事件组成的集合，称该试验的样本空间，记为 $S=\{e\}$。一个事件是一些基本事件组成的集合。即，每一事件是 S 的一个子集。

以下设 A,B,C 都是同一随机试验的事件，并以 S 表示必然事件，Φ 表示不可能事件，则随机事件之间的关系和运算可与集合之间的关系和运算对应起来。

(1)如果 A 发生，必导致 B 发生，称 A 为 B 的子事件，记作 $A\subset B$，若 $A\subset B$ 且 $B\subset A$，称 A、B 相等，记作 $A=B$。

(2)称"A、B 中至少有一个出现"这个事件为事件 A,B 的和，记为 $A\cup B$。称"A 出现而 B 不出现"这一事件为 A,B 的差，记为 $A-B$。

(3)称"A,B 同时出现"为事件 A、B 的积事件，记为 $A\cap B$ 或 AB。

(4)若 A 与 B 不可能同时出现，即 $AB=\varnothing$，称 A 与 B 互斥或互不相容。若 A、B 互斥，和事件 $A\cup B$ 可记为 $A+B$。

(5)若 $AB=\varnothing$，$A\cup B=S$，则称 A 与 B 互逆，B 为 A 的逆事件，记为 $B=\overline{A}$。

事件的概率是在数量上表示事件出现的可能性大小的量。

定义　设 S 是随机试验的样本空间，如果对每一事件 $A\subset S$，都有一个实数 $P(A)$ 与之对应且满足：

(1) $0\leqslant P(A)\leqslant 1\qquad \forall\ A\subset S$

(2) $P(S)=1$

(3)对于两两互斥的可数多个事件 $A_1,A_2,\cdots$ 有 $P(A_1+A_2+\cdots)=P(A_1)+P(A_2)+\cdots$

则称 $P(A)$ 为事件 A 的概率

可证：$P(\varnothing)=0\qquad P(\overline{A})=1-P(A)$

$$P(A\cup B)=P(A)+P(B)-P(AB)$$

§2　条件概率与事件的独立性

定义1　设 A,B 为两个事件，$P(A)>0$，则事件 A 发生的条件下，事件 B 发生的条件概率

定义为 $P(B/A)=\frac{P(AB)}{P(A)}$

定义 2 若 $P(AB)=P(A)P(B)$则称事件 A 与 B 相互独立。

若 $P(AB)=P(A)P(B)$，$P(BC)=P(B)P(C)$

$P(AC)=P(A)P(C)$，则称 A,B,C 三事件两两相互独立，若还有$(ABC)=P(A)P(B)P(C)$则称三事件相互独立。

定理 1(全概率公式) 设 A 及 $B_1,B_2\cdots\cdots$为事件，$P(B_i)>0(i=1,2\cdots)$，$B_iB_j=\varnothing(i\neq j)$，$\bigcup B_i=S$

则 $P(A)=\sum P(B_i)P(A/B_i)$

定理 2(贝叶斯公式) 在定理 1 的条件下，若还有 $P(A)>0$，则：$P(B_k/A)=\frac{P(B_k)P(A/B_k)}{\sum P(B_i)P(A/B_i)}$

§3 随机变量与分布

定义 1 设 $S=\{e\}$为随机试验的样本空间，定义在 S 上的单值实函数 $X(e)$，$e\in S$ 称为随机变量。

随机变量取什么值，取决于试验结果。设 $\Omega\subset R^1$ 为实数集，记 $E=\{e|e\in S,X(e)\in\Omega\}$，则 E 表示一个随机事件。特别地$\{X<x_0\}$，$\{x_1<X<x_2\}$等等都是随机事件。随机变量的引入使随机事件与实数集的关系更密切，从而有可能用数学分析的方法去研究随机试验。

定义 2 若随机变量 X 所有可能取的值是有限个或可列个，则称之为离散型随机变量。

设离散型随机变量 X 所有可能取的值为 $a_1,a_2\cdots$，$P(X=a_i)=P_i(i=1,2,\cdots)$则下表

X	a_1	a_2	……	a_i	……
P	P_1	P_2	……	P_i	……

其中 $0\leqslant P_i\leqslant 1(i=1,2\cdots)$ $\sum_i P_i=1$ 称为 X 的分布密度。

定义 3 $F(x)=P\{X<x\}-\infty<x<+\infty$称为随机变量的分布函数。

显然，分布函数具有下列性质

(1)$0\leqslant F(x)\leqslant 1(-\infty<x<+\infty)$

(2)$F(x)$是单调不减函数。

(3)$F(-\infty)=0$，$F(+\infty)=1$

(4)$F(x-0)=F(x)$即 $F(x)$处处左连续。

(5)若 $F(x)$在 $x=x_0$ 点不连续，则其跃度等于 $P(X=x_0)$

定义 4 若随机变量 X 的分布函数为 $F(x)$，存在非负函数 $f(x)$，使 $F(x)=\int_{-\infty}^{x}f(x)dx$ 则称 X 为连续型随机变量，称 $f(x)$为 X 的分布密度。

显见，连续型随机变量的分布密度 $f(x)$有如下性质：

(1)$f(x)\geqslant 0$

(2)$\int_{-\infty}^{+\infty}f(x)dx=1$

(3) $P\{a \leqslant z < b\} = F(b) - F(a) = \int_a^b f(x)dx$

(4)设点 x 是 $f(x)$ 的连续点，则 $f(x)=F'(x)$

随机变量的概念还可推广到 n 维，以二维随机变量为例，离散型随机变量的分布密度成矩形状：

$X \backslash Y$	b_1	b_2	……	b_j	……
a_1	p_{11}	p_{12}	……	p_{1j}	……
a_2	p_{21}	p_{22}	……	p_{2j}	……
⋮	…	…	…	…	…
a_i	p_{i1}	p_{i2}	……	p_{ij}	……
⋮	…	…	…	…	…

其中 $0 \leqslant P_{ij} \leqslant 1 \quad \sum P_{ij} = 1$

定义 5　设 Z、Y 为两个随机变量，二元函数

$$F(x,y) = p\{X < x, Y < y\} \quad (-\infty < x < +\infty, -\infty < y < +\infty)$$

称为二维随机变量 (X,Y) 的分布函数，若存在非负的函数 $f(x,y)$ 使 $F(x,y) = \int_{-\infty}^{y}\int_{-\infty}^{x} f(x,y)dxdy$ 则称 (X,Y) 为连续型二维随机变量，$f(x,y)$ 称为密度函数。当 $f(x,y)$ 连续时，有 $f(x,y) = \frac{\partial^2}{\partial x \partial y}F(x,y)$

定义 6　设 (X,Y) 是二维随机变量，$F(X,Y)$ 为其分布函数，称 $F_1(x)=P\{X<x\}=P\{X<x, Y<+\infty\}=\lim\limits_{y\to+\infty} F(x,y)$ 为 (X,Y) 关于 X 的边缘分布函数，同样可得关于 Y 的边缘分布函数 $F_2(y)=P(Y<y)=P\{X<+\infty, Y<y\}=\lim\limits_{x\to+\infty} F(x,y)$

设 $f(x,y)$ 是二维连续型随机变量 (X,Y) 的密度函数则它的两个边缘分布也是连续型的，且其分布密度分别为：$f_1(x) = \int_{-\infty}^{+\infty} f(x,y)dy$

及　　$f_2(y) = \int_{-\infty}^{+\infty} f(x,y)dx$

定义 7　在事件 A 发生的条件下，随机变量 Y 的分布函数为 $F(y/A)=P\{Y<y/A\}=\frac{P\{Y<y, A\}}{P(A)}$

例如：当事件 A 为随机变量 $X<x$ 时，有：

$$F(y/X<x) = \frac{P\{X<x, Y<y\}}{P\{X<x\}} = \frac{F(x,y)}{F_1(x)}$$

当事件 A 为 $X=x$ 时，条件分布函数可记为 $F(y/x)$ 对于连续型随机变量 X 来讲 $P(X=x)=0$，上述定义就不适用了。若对任意正数 ε 有 $P\{x-\varepsilon \leqslant X < x+\varepsilon\} > 0$ 且极限 $\lim\limits_{\varepsilon\to0^+} P\{Y<y/x-\varepsilon \leqslant X < x+\varepsilon\}$ 存在，则可得：

$$F(y/x) = \lim_{\varepsilon\to0^+} P\{Y<y/x-\varepsilon \leqslant X < x+\varepsilon\}$$

$$= \lim_{\varepsilon\to0^+} \frac{P\{Y<y, x-\varepsilon \leqslant X < x+\varepsilon\}}{P\{x-\varepsilon \leqslant X < x+\varepsilon\}}$$

$$=\lim_{\varepsilon\to 0^+}\frac{F(y,x+\varepsilon)-F(y,x-\varepsilon)}{F_1(x+\varepsilon)-F_1(x-\varepsilon)}$$

$$=\frac{\frac{\partial}{\partial x}F(x,y)}{\frac{d}{dx}F_1(x)}$$

$$=\frac{\int_{-\infty}^{y}f(x,y)dy}{f_1(x)} \qquad (f_1(x)\neq 0)$$

两边对 y 求导得：

$$f(y/x)=\frac{f(x,y)}{f_1(x)}$$

称 $f(y/x)$ 为在条件 $X=x$ 下，Y 的条件分布密度。

类似地可得在条件 $Y=y$ 下，X 的条件分布密度 $f(x/y)=\dfrac{f(x,y)}{f_2(y)}$.

定义 8 若 X,Y 是两个连续型随机变量，其分布密度及条件分布密度符号见前所述，若

$$f(x/y)=f_1(x),f(y/x)=f_2(y)$$

对一切 x,y 都成立，则称 X 和 Y 是相互独立的。

显见，X 与 Y 是相互独立的充要条件是：

$f(x,y)=f_1(x)f_2(y)$，其中 $f(x,y)$ 是二维随机变量 (X,Y) 的分布密度。

对于多个随机变量的联合分布，边缘分布及条件分布，可类似地得出相应的结论，在此不再赘述。

§4 随机变量的数字特征

定义 1 设离散型随机变量 X 的分布密度为

X	a_1	a_2	……	a_i	……
P	p_1	p_2	……	p_i	……

若级数 $\sum_{k=1}^{\infty}a_kp_k$ 绝对收敛，称它的值为 X 的数学期望，记为 $EX=\sum_{k=1}^{\infty}a_kp_k$

定义 2 设连续型随机变量 X 的分布密度函数为 $f(x)$，若 $\int_{-\infty}^{+\infty}xf(x)dx$ 绝对收敛，称它为 X 的数学期望，记为 $EX=\int_{-\infty}^{+\infty}xf(x)dx$.

设 $g(x)$ 为一连续函数，则 $Y=g(X)$ 也是随机变量，若 X 为离散型定义 $EY=E(g(X))=\sum_{k=1}^{\infty}g(a_k)p_k$.

若 X 为连续型定义 $EY=E(g(X))=\int_{-\infty}^{+\infty}g(x)f(x)dx$，其中级数及积分都假定为绝对收敛的。

更可以推广到多个随机变量的情况，例如 X,Y 都是随机变量。联合分布密度为 $f(x,y)$，又设 $g(x,y)$ 为连续函数，则 $Z=g(X,Y)$ 也是随机变量，则定义 $EZ=E(g(X,Y)=$

$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y)f(x,y)dxdy$ 为 Z 的数学期望，若积分是绝对收敛的话。

定义 3 设 X 为随机变量，称 $E\{x^n\}$ 为 X 的 n 阶原点矩，记作 m_n。

称 $E\{(X-EX)^n\}$ 为 X 的 n 阶中心矩，记作 μ_n。

显然 $m_0=1, m_1=EX, m_2$ 亦称均方值。

又可得 $\mu_1=0, \mu_2$ 称为方差，常记作 $D(X)$ 或 $\sigma^2(X)$，而 $\sqrt{\mu_2}=\sigma(X)$ 又称为标准差。

数学期望有性质：

(1)若 K 为常数，则 $E\{KX\}=KEX$

(2)$E(X\pm Y)=EX\pm EY$

(3)当 X,Y 相互独立时，$E(XY)=EX\cdot EY$

方差有性质：

(1)常数的方差为零

(2)若 K 为常数，则 $\sigma^2(KX)=K^2\sigma^2(X)$

(3)若 X,Y 相互独立，则 $\sigma^2(X\pm Y)=\sigma^2(X)+\sigma^2(Y)$，另外可证 $\sigma^2=m_2-m_1^2=EX^2-(EX)^2$

定义 4 设 X,Y 为两个随机变量，称 $E(X^nY^k)$ 为 X 和 Y 的 $n+k$ 阶联合原点矩，记作 m_{nk}，称 $E\{(X-EX)^n(Y-EY)^k\}$ 为 X 和 Y 的 $n+k$ 阶联合中心矩，记作 μ_{nk}。

显然，$m_{n0}=E(X^n)$　$m_{0k}=E(Y^k)$ 分别为 X 的 n 阶，Y 的 k 阶矩。又 $\mu_{20}=\sigma^2(X)$　$\mu_{02}=\sigma^2(Y)$

$\mu_{11}=E\{(X-EX)(Y-EY)\}$ 称为 X,Y 的协方差（或二阶混合中心矩）常记作 $\mathrm{cov}(X,Y)=\mathrm{cov}(Y,X)$ 表示 X 与 Y 间的关联程度，归一化后，记作：

$$\rho(X,Y)=\frac{\mathrm{cov}(X,Y)}{\sigma(X)\sigma(Y)}=\frac{E\{(X-EX)(Y-EY)\}}{[E\{(X-EX)^2\}\cdot E\{(Y-EY)^2\}]^{1/2}}$$

并称为 X 与 Y 的相关系数。

相关系数有性质：

(1)$|\rho(X,Y)|\leqslant 1$

(2)若 X 与 Y 间以概率 1 存在线性关系，即 $P\{Y=aX+b\}=1$（a,b 为常数）则 $|\rho(X,Y)|=1$

定义 5 设 X,Y 为两个随机变量，若 $\mathrm{cov}(X,Y)=0$（或 $\rho(X,Y)=0$）称 X 与 Y 互不相关。

若 X 与 Y 相互独立，则 X 与 Y 互不相关，但反之不然。

若 X 与 Y 互不相关，则 $E(XY)=E(X)E(Y)$　$\sigma^2(X\pm Y)=\sigma^2(X)+\sigma^2(Y)$

作为本节结束，今证明一个重要的不等式

许瓦兹不等式 $[E(XY)]^2\leqslant E(X^2)E(Y^2)$

证明：设 ζ,η 为任意实数，则

$E\{(\zeta X+\eta Y)^2\}=E(X^2)\zeta^2+2E(XY)\zeta\eta+E(Y^2)\eta^2$ 上式对任意的 ζ,η 都非负，故由二次型性质知

$$[E(XY)]^2\leqslant E(X^2)\cdot E(Y^2)$$

相关系数 $|\rho|\leqslant 1$ 可作为其直接结果。

§5 几种常见的概率分布

一、二项分布

假定进行 n 次独立试验，每次试验中 A 出现的概率是 p，不出现的概率是 $q=1-p$，记事件 A 在 n 次试验中发生 m 次的概率为 $p_n(m)$，则

$$p_n(m) = \frac{n!}{m!(n-m)!}p^m q^{n-m}$$

设 X 为 n 次试验中 A 出现的次数，则 X 是离散型随机变量，可取 0 到 n 的任一整数，且

$$P(X=m) = C_n^m p^m q^{n-m} (m = 0,1,2,\cdots n)$$

则称 X 服从二项分布

$$EX = \sum_{m=0}^{n} C_n^m p^m q^{n-m} \cdot m = np$$

$$DX = E\{X^2\} - (EX)^2 = npq$$

二、泊松分布

在二项分布中，若试验次数 n 无限增大，p 无限减小，且 $np=\lambda$ 不变，则：

$$p_n(m) = p(X=m) = \frac{\lambda^m}{m!}e^{-\lambda}$$

此时称随机变量 X 服从泊松分布，即 X 可取一切非负整数，且 $p(X=m)=\frac{\lambda^m}{m!}e^{-\lambda}$

这也是离散型随机变量，$EX=\lambda$，　$DX=\lambda$

三、均匀分布

若连续型随机变量 X 的分布密度为

$$f(x) = \begin{cases} \dfrac{1}{b-a} & a < x < b \\ 0 & \text{其它} \end{cases}$$

则称 X 为区间 (a,b) 上的均匀分布

$$EX = \frac{a+b}{2} \qquad DX = \frac{(b-a)^2}{12}$$

四、正态分布

称分布密度为 $f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-a)^2}{2\sigma^2}}(-\infty<x<+\infty)$ 的连续型随机变量为服从正态分布的，简记为 $N(a,\sigma^2)$，也称该分布为高斯分布

$$EX = a \qquad DX = \sigma^2$$

中心极限定理　设 $X_1,X_2,\cdots$ 为独立的有同一分布的随机变量，$EX_i=a$　$DX_i=\sigma^2(i=1,2\cdots\cdots)$

则对任意 x，有

$$\lim_{n\to\infty} p\{\frac{1}{\sigma\sqrt{n}}\sum_{i=1}^{n}(X_i - a) < x\} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{x^2}{2}}dx$$

定理的证明见一般概率论教科书，今给出如下解释。

记 $Y_n = \sum_{i=1}^{n} X_i$ 则由于 X_i 相互独立，得

$$EY_n = na \quad DY_n = n\sigma^2$$

归一化后记 $\bar{Y}_n = \frac{Y_n - na}{\sqrt{n}\,\sigma} = \frac{Y_n - EY_n}{\sqrt{DY_n}}$

则 $E\bar{Y}_n = 0 \qquad D\bar{Y}_n = 1$

中心极限定理可记为：

$$\lim_{n\to\infty}\{(\bar{Y}_n < x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{x^2}{2}}dx$$

即当 $n\to\infty$ 时，$\bar{Y}_n$ 的分布趋于标准正态分布 $N(0,1)$ 从而 Y_n 的分布近似于 $N(na, n\sigma^2)$。

上述结论表明，任何物理过程，若为许多独立作用之和，则该过程趋于高斯过程。具体地说，若有很多独立的服从同一分布的随机变量，不论它们服从什么分布，只要具有有限方差，每个变量对总和的影响很小，则它们的和作为随机变量，其分布总是近似于正态分布的。这正是正态分布成为科技领域内最常见的重要分布的原因。

设 X 服从 $N(0,1)$，则其 n 阶矩为：

$$E(X^n) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} x^n e^{-\frac{x^2}{2}}dx$$

显然当 n 为奇数时，$E(X^n)=0$

当 n 为偶数时，由分部积分得：

$$E(X^n) = -\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} x^{n-1} de^{-\frac{x^2}{2}} = \frac{n-1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} x^{n-2}e^{-\frac{x^2}{2}}dx$$

$$= (n-1)E(X^{n-2})$$

由于 $E(X^0)=1$ 故

$$E(X^n) = \begin{cases} 0 & n \text{ 为奇数} \\ (n-1)(n-3)\cdots 5.3.1 = (n-1)!! & n \geqslant 2 \text{ 偶数} \end{cases}$$

若

X 服从 $N(a,\sigma^2)$，则 $Y = \frac{X-a}{\sigma}$ 服从 $N(0,1)$

$$E\{(X-a)^n\} = E(\sigma^n Y^n) = \sigma^n E(Y^n) = \begin{cases} 0 & n \text{ 为奇数} \\ (n-1)!!\sigma^n & n \geqslant 2 \text{ 偶数} \end{cases}$$

§6 特征函数

设 X、Y 为实随机变量，则 $Z = X + iY$ 为复随机变量，定义 $EZ = EX + iEY$，则有关矩，特别是数学期望与方差的一些性质对复随机变量仍正确。

设 X 为实随机变量，t 为实参数，则 e^{itX} 为复随机变量。

定义 函数 $\varphi(t) = E(e^{itX})$ 称为随机变量 X 的特征函数。

对连续型随机变量 X 而言，$\varphi(t)=\int_{-\infty}^{+\infty}f(x)e^{itx}dx$，其中 $f(x)$ 为其分布密度。对离散型随机变量 X 而言，$\varphi(t)=\sum_j p_j e^{ita_j}$，其中 $p(X=a_j)=p_j(j=1,2,\cdots)$

定理 1 特征函数在 $-\infty<t<+\infty$ 上连续，且

$$\varphi(0)=1 \qquad |\varphi(t)|\leqslant 1$$

需要说明的是，对任意随机变量，其特征函数都存在，且连续，即使对离散型随机变量而言也如此。

定理 2(唯一性定理) 分布函数由其特征函数唯一决定。

此定理述而不证。由本定理可知，随机变量的分布函数与特征函数之间有一一对应关系，两者同样可完整地描述一个随机变量。

定理 3 设 $Y=aZ+b(a,b$ 为常数)，则

$$\varphi_Y(t)=\varphi_X(at)e^{ibt}$$

其中 $\varphi_Y(t)$ 和 $\varphi_X(t)$ 分别表示 Y 和 X 的特征函数。

[证] $\varphi_Y(t)=E(e^{itY})=E[e^{it(aX+b)}]=e^{itb}E(e^{itaX})=e^{itb}\varphi_X(at)$

定理 4 相互独立的随机变量之和的特征函数，等于各随机变量特征函数之积。

[证] 以两个相互独立的随机变量为例。设 $Z=Z+Y$，由于 X 与 Y 相互独立。则 e^{itX} 与 e^{itY} 亦相互独立。故：

$$E(e^{itZ})=E(e^{it(X+Y)})=E(e^{itX}\cdot e^{itY})=E(e^{itX})E(e^{itY})$$

即 $\varphi_Z(t)=\varphi_X(t)\cdot\varphi_Y(t)$

定理 5 若随机变量 X 有 n 阶绝对矩，则其特征函数有 n 阶导数，且当 $k\leqslant n$ 时

$$m_k=E(X^k)=(-i)^k\varphi^{(k)}(0)$$

[证] 仅对连续型随机变量作证明，

此时：$\varphi(t)=\int_{-\infty}^{+\infty}e^{itx}f(x)dx$

两边对 t 求 k 次导数得

$$\varphi^{(k)}(t)=\int_{-\infty}^{+\infty}(ix)^k e^{itx}f(x)dx$$

积分号下求导的可能性由 k 阶绝对矩的存在性保证。令 $t=0$ 立得 $\varphi^{(k)}(0)=i^kE(X^k)$

$$即：E(X^k)=(-i)^k\varphi^{(k)}(0)$$

由以上讨论可见，对于求独立随机变量之和的分布，以及求随机变量的各阶矩，特征函数都是很有力的工具，而且适用于一切类型的随机变量。

例 1 设随机变量 X 服从二项分布，求其特征函数。

解：记 $X=\mu_1+\mu_2+\cdots+\mu_n$，其中 μ_j 是随机变量，当事件 A 在第 j 次试验中出现时，取值为 1，否则取值为 $0(j=1,2\cdots n)$

$$\because\quad P(\mu_j=1)=p \qquad p(\mu_j=0)=q=1-p$$

$$\therefore\quad \varphi_j(t)=E(e^{it\mu_j})=e^{it\cdot 0}\cdot q+e^{it\cdot 1}p=q+pe^{it}$$

而 μ_j 相互独立，由定理 4 知

$$\varphi_X(t)=\prod_{j=1}^{n}\varphi_j(t)=(q+pe^{it})^n$$

例 2　设 X 服从泊松分布，求其特征函数

解　$$\varphi(t) = E(e^{itX}) = \sum_{j=0}^{\infty} e^{itj} \cdot \frac{\lambda^j}{j!} e^{-\lambda} = e^{-\lambda} \sum_{j=0}^{\infty} \frac{(\lambda e^{it})^j}{j!}$$

$$= e^{\lambda(e^{it}-1)}$$

例 3　设 X 为区间 (a,b) 上的均匀分布，求其特征函数

解　$$\varphi(t) = E(e^{itX}) = \int_a^b \frac{e^{itx}}{b-a} dx = \frac{e^{itb} - e^{ita}}{it(b-a)}$$

例 4　设 X 服从正态分布 $N(a,\sigma^2)$ 求其特征函数

解　$$\varphi(t) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} e^{itx - \frac{(x-a)^2}{2\sigma^2}} dx$$

令　$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty - it\sigma}^{+\infty - it\sigma} e^{-\frac{y^2}{2}} dy \cdot e^{iat - \frac{\sigma^2 t^2}{2}} = e^{iat - \frac{\sigma^2 t^2}{2}}$$

用特征函数求数学期望及方差特别简单

$$\because \quad EX = m_1 \qquad DX = m_2 - m_1^2$$

$$\therefore \quad EX = -i\varphi'(0) \qquad DX = -\varphi''(0) + [\varphi'(0)]^2$$

如果令　$$\psi(t) = \ln\varphi(t) \qquad \text{则} \qquad \psi'(t) = \frac{\varphi'(t)}{\varphi(t)}$$

$$\psi''(t) = \frac{\varphi''(t)\varphi(t) - [\varphi'(t)]^2}{[\varphi(t)]^2}$$

令　$t = 0$　注意到 $\varphi(0) = 1$　得：$\psi'(0) = \varphi'(0)$

$$\psi''(0) = \varphi''(0) - [\varphi'(0)]^2$$

$$\therefore \quad EX = -i\psi'(0) \qquad DX = -\psi''(0)$$

读者不难对上述四个例子计算其数学期望与方差，将得到完全与前面一致的结果。

特征函数的概念可以推广到多维情形，仅以二维为例简述如下。

定义　函数 $\varphi(u,v) = E\{\exp(iuX + ivY)\}$ 称为随机变量 X 和 Y 的联合特征函数。对连续型二维随机变量 (X,Y) 而言，$\varphi(u,v) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y) e^{i(ux+vy)} dx dy$ 其中 $f(x,y)$ 为联合密度函数。

令 $u=0$ 或 $v=0$ 便得出边缘特征函数：

$$\varphi_X(u) = \varphi(u,0), \qquad \varphi_Y(v) = \varphi(0,v)$$

定理 6　若 X 与 Y 相互独立，则 $\varphi(u,v) = \varphi_X(u) \cdot \varphi_Y(v)$

定理 7　若 X 和 Y 有 $n+k$ 阶联合绝对矩，则其特征函数 $\varphi(u,v)$ 有 $(n+k)$ 阶导数，且：

$$m_{nk} = E(X^n Y^k) = (-i)^{n+k} \left. \frac{\partial^{n+k} \varphi(u,v)}{\partial u^n \partial v^k} \right|_{\substack{u=0 \\ v=0}}$$

§7　多元正态分布随机变量

由于正态分布的特殊重要性，故对 n 维正态分布单独作为一节叙述，而不再局限于二维的叙述。

n 维随机变量可用随机矢量表示之

$$\zeta=\begin{bmatrix}\zeta_1\\ \zeta_2\\ \vdots\\ \zeta_n\end{bmatrix},\qquad \mu=E\zeta=\begin{bmatrix}E\zeta_1\\ E\zeta_2\\ \vdots\\ E\zeta_n\end{bmatrix}=\begin{bmatrix}\mu_1\\ \mu_2\\ \vdots\\ \mu_n\end{bmatrix}$$

记 $b_{ik}=\mathrm{cov}(\zeta_i,\zeta_k)=E\{(\zeta_i-\mu_i)(\zeta_k-\mu_k)\}$ 为 ζ_i,ζ_k 之间的协方差，显然 $b_{ii}=D\zeta_i$ 是 ζ_i 的方差，矩阵 $B=\begin{bmatrix}b_{11} & b_{12} & \cdots & b_{1n}\\ b_{21} & b_{22} & \cdots & b_{2n}\\ \cdots\cdots & & \cdots\cdots & \\ b_{n1} & b_{n2} & \cdots & b_{nn}\end{bmatrix}$ 称为协方差矩阵，显然为对称阵。

对于任意实数组 $\lambda_1,\lambda_2,\cdots\cdots,\lambda_n$，由

$$E(\sum_{i=1}^{n}\lambda_i(\zeta_i-\mu_i))^2\geqslant 0$$

得 $\sum_{i,k=1}^{n}b_{ik}\lambda_i\lambda_k\geqslant 0$，故知协方差矩阵具有非负定性。

n 维正态分布随机变量的协方差矩阵正定且分布密度函数为

$$f(\mathbf{x}^T)=\varphi(x_1,x_2,\cdots,x_n)=\frac{1}{(2\pi)^{\frac{n}{2}}|B|^{\frac{1}{2}}}\exp[-\frac{1}{2}(\mathbf{x}-\mu)^TB^{-1}(\mathbf{x}-\mu)]$$

其中 $\mathbf{x}^T$ 表示矩阵 $\mathbf{x}$ 的转置，$|B|$ 为矩阵 B 的行列式。或写成：

$$f(x_1,x_2,\cdots,x_n)=\frac{1}{(2\pi)^{\frac{n}{2}}|B|^{\frac{1}{2}}}\exp[-\frac{1}{2|B|}\sum_{i,k=1}^{n}\bar{b}_{ik}(x_i-\mu_i)(x_j-\mu_j)]$$

其中 $\bar{b}_{ik}$ 为行列式 $|B|$ 中相应元素的代数余子式。

特征函数 $\varphi(\mathbf{t}^T)=\varphi(t_1,t_2,\cdots,t_n)=\exp(i\mu^T\mathbf{t}-\frac{1}{2}\mathbf{t}^TB\mathbf{t})=\exp[i\sum_{j=1}^{n}\mu_jt_j-\frac{1}{2}\sum_{j,k=1}^{n}b_{ij}t_jt_k]$

由上述可知，正态分布的随机矢量完全由其期望矢量及协方差矩阵确定。当 n 个随机变量 ζ_i 互不相关时，$b_{ik}=0(i\neq k)$，$B=\mathrm{diag}'[\sigma_1^2,\sigma_2^2,\cdots,\sigma_n^2]$，$|B|^{\frac{1}{2}}=\prod_{j=1}^{n}\sigma_j$，故概率密度为：

$$f(x_1,x_2,\cdots,x_n)=\prod_{j=1}^{n}\frac{1}{\sqrt{2\pi}\sigma_j}e^{-\frac{(x_j-\mu_j)^2}{2\sigma_j^2}}=\prod_{j=1}^{n}f_{\zeta_j}(x_j)$$

即 ζ_i 相互独立。

设 $\{\zeta_{k_1},\zeta_{k_2},\cdots,\zeta_{k_m}\}^T(m<n)$ 为 ζ 的任一个子向量，则在上式中令 $t_j=0(j\neq k_1,k_2,\cdots,k_m)$，便得边缘特征函数

$$\varphi(t_{k_1}\ \ t_{k_2}\ \ \cdots\ \ t_{k_m})=\exp[i\tilde{\mathbf{t}}^T\tilde{\mu}-\frac{1}{2}\tilde{\mathbf{t}}^T\tilde{B}\,\tilde{\mathbf{t}}]$$

其中 $\tilde{\mathbf{t}}=\{t_{k_1},t_{k_2},\cdots,t_{k_m}\}^T$ $\quad\tilde{\mu}=\{\mu_{k_1},\mu_{k_2},\cdots,\mu_{k_m}\}^T$，$\tilde{B}$ 为保留 B 的第 $k_1,k_2,\cdots,k_m$ 行与列所得的 $m\times m$ 阶阵。

由唯一性定理知：$\{\zeta_{k_1},\zeta_{k_2},\cdots,\zeta_{k_m}\}$ 也服从正态分布。

定理1 ζ 服从 n 维正态分布 $N(\mu,B)$ 的充要条件是它的任一线性组合 $\zeta=\sum_{k=1}^{n}a_k\zeta_k=\mathbf{a}^T\zeta$ 服从一维正态分布

$$N(\sum_{k=1}^{n}a_k\mu_k,\sum_{i,k=1}^{n}a_ia_kb_{ik})=N(\mathbf{a}^T\mu,\mathbf{a}^TB\mathbf{a})$$