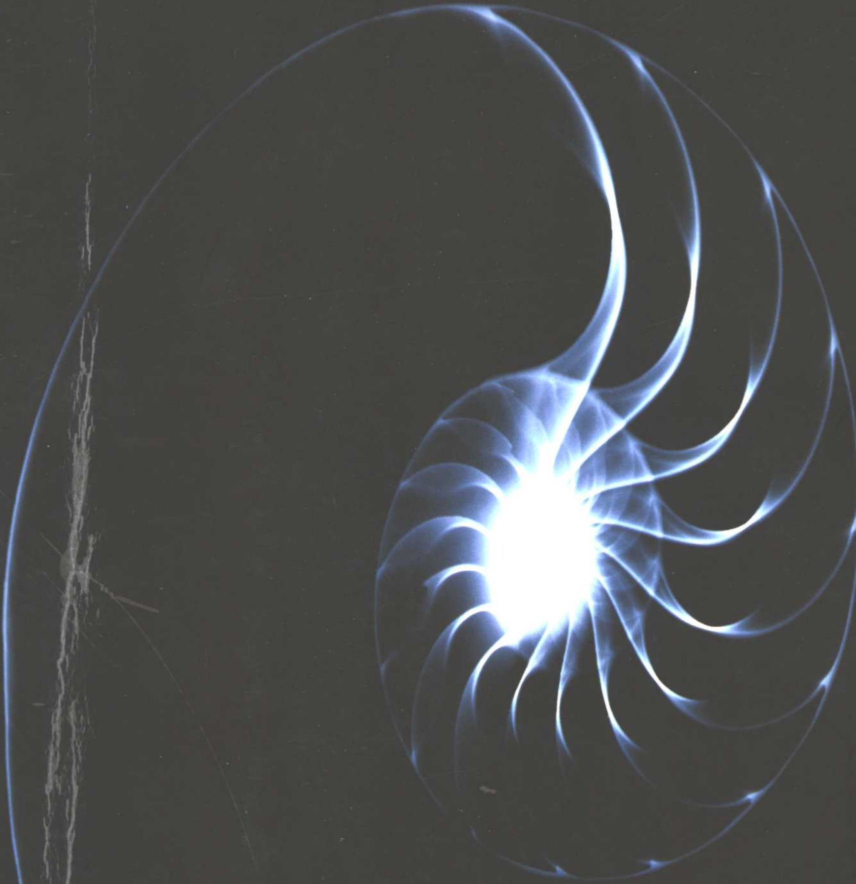

The Stochastic Process of Nonlinear Systems

非线性系统的随机过程

彭仕政/蔡绍洪/唐延林/著

非线性系统的随机过程

彭仕政　蔡绍洪　唐延林　著

贵州人民出版社

图书在版编目(CIP)数据

非线性系统的随机过程/彭仕政著. 一贵阳:贵州人民出版社,2000.7
ISBN 7-221-05110-0

Ⅰ. 非... Ⅱ. 彭... Ⅲ. 非线性系统(自动化)-随机过程 Ⅳ. TP271

中国版本图书馆 CIP 数据核字(2000)第 22480 号

非线性系统的随机过程

著 作 人:彭仕政 蔡绍洪 唐延林
责任编辑:杨民生
特约编辑:赵树民
封面设计:邹 刚
技术设计:施德端
出版发行:贵州人民出版社
(贵阳市中华北路 289 号)
印 刷:贵州兴隆印务有限责任公司
开 本:850×1168 1/32
字 数:149 千字
印 张:6
版 次:2001 年 7 月第 1 版第 1 次印刷
印 数:1—1000 册

ISBN 7-221-05110-0/G·1970 **定价**:28.00 元

前言

20世纪60年代以来,非线性科学已成为跨学科的研究前沿。包括自然科学和社会科学在内的各门学科,都存在着各自的非线性现象,过去研究的线性现象,不过是非线性现象有条件的特例。非线性科学的迅速崛起,被认为是继量子理论和相对论后20世纪物理学的第三次革命。

确定性的线性系统其运动结果是确定的,确定性的非线性系统却可能出现不确定的随机结果,这就表现出了非线性动力系统的内在随机性。事实上,任何系统的存在都不是孤立的,系统之间始终是相互依存、相互影响、相互制约的,这就是非线性。这里的依存、影响和制约的程度稍有变化,亦即涨落,系统发生和运动的结果也就会呈现出差异,这也就表现出系统运动的随机性。事物的相互作用正是非线性的反映。用一个幅度特性分布服从高斯概率密度函数的随机信号输入线性系统时,系统会有满足高斯分布的随机响应输出,但对于非线性系统,则这种高斯激励将不一定有对应的高斯输出。对于远离临界点的稳定平衡态系统,小的信号涨落不能改变系统的状态,输出响应仅围绕平衡点作相应的小涨落,与线性系统的表现类似。但是在临界点附近,信号涨落或噪声将使系统状态发生大的变化,响应发生大幅值的内随机响应,这种激励对能演化成稳定周期激励的非稳定平衡态的稳定性是一种检测。不稳定平衡态在临界点附近,微涨落就可能激发其内部的巨涨落,从而使得系统失去原有的有序模态而演化为无序的随机运动。不稳定平衡态在远离临界点处,涨落到对系统原有的有序周期运动的影响大大减弱,小涨落此时实际已难改变系统原有运动状态和模式。对任意系统,如既有稳定平衡态又有非稳定平衡态的系统,涨落对

含有周期信号的非线性系统的激励不仅能激发系统响应的无序运动、随机运动，而且还能增大系统的有序周期运动。随机共振过程就是序参量涨落的激励，增加了系统响应有序运动的例证。越靠近临界点，随机共振的模态特征就越明显，反之，远离临界点，系统对之响应的随机共振过程特征就越不明显。

非线性系统的随机运动过程实际上是非常普遍的现象，为使更多的人对此有一定了解和关注，我们试做了一定的工作，因而有了本书。本书纂写过程中参考、借鉴了众多专家学者的研究思想、观点、方法和著述，有关文献资料，一并附于书末，由于水平有限，书中不足乃至错误之处在所难免，请批评、指正。

作　者

2001 年 3 月

目　录

第一章　随机变量

随机过程有各种类型，处理不同过程所需采用的办法也有所不同，因而有必要将各种类型随机过程所涉及的数学描述作一介绍.

研究随机现象，常把可以预想的试验结果用一样本点表示，记为 S. 试验中所有样本点构成样本空间 S. 构成 S 的事件称为 S 的子集，两个子集不相交称其互不相容. 互不相容事件不是同一试验的结果.

设 $P(A)$表示事件 A 的概率，则：

(1)整个样本空间的概率为 1. $P\{S\}=1$　(1.1)

(2)对任何事件 A，$P\{A\}\geqslant 1$ 表示事件的概率是非负的.　(1.2)

(3)互不相容事件之一发生的概率为：

$$P\{A_1 \text{或} A_2 \text{或} \cdots\}=\sum_{i=1}^{\infty}P\{A_i\} \tag{1.3}$$

一、　随机变量

试验中变量的取值不可预言，但该变量取不同值的概率是确定的，则这样的变量称为随机变量. 定义样本空间 S 上的随机变量为单值函数$X(s)$. 将 S 中的一个事件的概率解释为随机变量 $X(s)$的值在某一区间或某一实数域内的概率. 随机变量常可分离散随机变量和连续随机变量，前者在有限个数中取值，对任一可能的取值 x_i，随机变量以惟一概率取 x_i，表为：

$$P\{X(s)=x_i\}=P_i,\quad i=0,1,\cdots \tag{1.1.1}$$

而称概率 P_i 的序列$\{P_i\}$为随机变量 $X(s)$的概率分布，称累积概率

$$P\{X(s)\leqslant x\}=\sum_{x_i\leqslant x}P_i=F(x)\quad -\infty<x<\infty \tag{1.1.2}$$

为随机变量 $X(s)$ 的分布函数.

通常,若 $X(s)$ 只能取一个值,称其为退化的随机变量,此时 $X(s)$ 为常数;若 $\{p_i\}$ 满足

$$\sum_{i=1}^{n}\{p_i\}=1 \tag{1.1.3}$$

称 $X(s)$ 为正规随机变量,反之,若

$$\sum_{i=1}^{n}\{p_i\}<1 \tag{1.1.4}$$

称 $X(s)$ 为非正规随机变量,而 $1-\sum_{i=1}^{n}p_i$ 是 $X(s)$ 不取有限值的概率.

若存在一个非负的函数 f,使对任意 $a\leqslant b$ 有

$$P\{a<X\leqslant b\}=\int_a^b f(x)dx \tag{1.1.5}$$

则称 X 为连续随机变量,$f(x)$ 称为 X 的概率密度函数. X 的分布函数则为:

$$F(x)=P\{X\leqslant x\}=\int_{-\infty}^{x}f(t)dt \tag{1.1.6}$$

和

$$P\{a<X\leqslant b\}=F(b)-F(a) \tag{1.1.7}$$

连续随机变量可在数轴的某一区间内连续取值,因而 X 取某一值的概率无意义,亦即 $P\{X=x\}=0$,这时可将 X 取值在 $x\sim x+dx$ 区间内的概率为:

$$dP|_{X=x}=f(x)dx \tag{1.1.8}$$

若密度函数满足.

$\int_{-\infty}^{\infty}f(x)dx=1$,则称此连续随机变量是正规的,如果上述积分值小于 1,则为非正规的.

二、 多变量随机变量

离散随机变量和连续随机变量都可以推广到包含多个随机变量的多变量情形.

由 n 个随机变量 $X_1, X_2, \cdots, X_n$ 构成一组 n 维随机变量 $\mathbf{X}=\begin{pmatrix} X_1 \\ \cdot \\ \cdot \\ \cdot \\ X_n \end{pmatrix}$，即将这一组随机变量看成定义在样本空间上的一个随机向量.

将 $f_n(x_1, x_2 \cdots, x_n)$ 为 X_i 取值 $x_i (i=1,2\cdots,n)$ 的联合概率分布. 考虑到今后重点讨论连续随机变量，我们重点地按概率的正定性和归一性得到.

$$f_n(x_1, x_2 \cdots x_n) \geqslant 0 \tag{1.2.1}$$

$$\int \cdots \int f(x_1, x_2 \cdots x_n) dx_1 dx_2 \cdots dx_n = 1 \tag{1.2.2}$$

由 n 维分布可求出低维($r<n$)的约化分布函数

$$f_r(x_1, x_2 \cdots x_r) = \int \cdots \int f_n(x_1, x_2, \cdots, x_n) dx_{r+1} \cdots dx_n \tag{1.2.3}$$

(1.2.3)式表示 X_i 取值 x_i，而 X_j 可取任何值的概率分布密度(其中 $i=1,2\cdots,r; j=r+1,\cdots,n$). 由约化分布又可定义条件概率

$$\mathrm{P}(x_1, \cdots, x_s | x_{s+1} \cdots\cdots, x_n) = \frac{f_n(x_1, x_2, \cdots, x_n)}{f_{n-s}(x_{s+1}, \cdots, x_n)}$$

$$= \frac{f_n(x_1, \cdots, x_n)}{\int \cdots \int f_n(x_1, x_2 \cdots, x_n) dx_1 \cdots dx_s} \tag{1.2.4}$$

表示在 $X_{s+1}, \cdots, X_n$ 分别取值为 $x_{s+1}, \cdots, x_n$ 的条件下 $X_1, \cdots, X_s$ 分别取值为 $x_1, \cdots, x_s$ 的条件概率分布函数. 其中

$$\mathrm{P}(x_1 | x_2) = \frac{f_2(x_1, x_2)}{\int f_2(x_1, x_2) dx_1} \tag{1.2.5}$$

称为 x_1 对 x_2 的条件分布.

对正规离散随机变量 X、Y，可定义其联合概率分布 $\{p_{ij}\}$ 为：

$$p_{ij}=P\{X=x_i,Y=y_i\} \qquad i,j=0,1\cdots\cdots \tag{1.2.6}$$

且

$$\sum_i\sum_j p_{ij}=1 \tag{1.2.7}$$

令

$$p_i=\sum_j p_{ij} \tag{1.2.8}$$

$$q_j=\sum_i p_{ij} \tag{1.2.9}$$

则

$$p_i=\sum_j p_{ij}=\sum_j P\{X=x_i,Y=y_j\}=P\{X=x_i\} \tag{1.2.10}$$

$$q_j=\sum_i p_{ij}=P\{Y=y_j\} \tag{1.2.11}$$

$\{p_i\}$、$\{q_j\}$ 分别称为 X 和 Y 的边缘分布，联合分布 $\{p_{ij}\}$ 决定 $\{p_i\}$ 和 $\{q_j\}$. 非正规随机变量则不然，这点应务必认识.

对 $X=x_i$，可定义 Y 的条件概率分布：对 $p_i>0$，

$$P\{Y=y_j|X=x_i\}=\frac{P\{X=x_i,Y=y_j\}}{P\{X=x_i\}}=\frac{p_{ij}}{p_i} \tag{1.2.12}$$

对 $Y=y_j$，

$$P\{X=x_i|Y=y_j\}=\frac{P\{X=x_i,Y=y_j\}}{P\{Y=y_j\}}=\frac{p_{ij}}{p_j} \tag{1.2.13}$$

若对所有 i 和 j，

$$p_{ij}=p_i p_j \tag{1.2.14}$$

则 X、Y 随机独立，或称 X 和 Y 是独立分布的.

若 X、Y 为正规连续，则可有

$$P\{a<X\leqslant b,c<Y\leqslant d\}=\int_c^d\int_a^b \xi(x,y)dxdy \tag{1.2.15}$$

及

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\xi(x,y)dxdy=1 \tag{1.2.16}$$

其中之 $\xi(x,y)$ 称为连续随机变量的联合概率密度函数.

如果
$$\left.\begin{aligned}P\{a < X \leqslant b\} &= \int_a^b f(x)dx \\ P\{c < Y \leqslant d\} &= \int_c^d g(y)dy\end{aligned}\right\} \tag{1.2.17}$$

则 $f(x) = \int_{-\infty}^{\infty} \xi(x,y)dy$ 和 $g(y) = \int_{-\infty}^{\infty} \xi(x,y)dx$ (1.2.18)

称为边缘密度函数.相应地,条件密度函数定义为:

$$f(x|y)=\frac{\xi(x,y)}{g(y)},g(y|x)=\frac{\xi(x,y)}{f(x)} \tag{1.2.19}$$

其中 $f(x)>0,g(y)>0$ 且满足

$$f(x) = \int_{-\infty}^{\infty} f(x|y)g(y)dy,g(y) = \int_{-\infty}^{\infty} g(y|x)f(x)dx \tag{1.2.20}$$

若
$$\xi(x,y)=f(x)g(y) \tag{1.2.21}$$

对所有 x 和 y 成立,X、Y 即随机独立,其相关函数 $\varphi(X)$、$\psi(Y)$ 也为独立随机变量.

三、 数学期望

随机变量的数学期望定义为该随机变量的所有可能值的加权平均,其权重为相应概率.离散变量 X 的数学期望为:

$$\langle X\rangle=\sum_i x_i p_i \tag{1.3.1}$$

连续变量 X 的期望则定义为:

$$\langle X\rangle = \int_{-\infty}^{\infty} xf(x)dx \tag{1.3.2}$$

若上两式之结果绝对收敛,则$\langle X\rangle$存在.

一般,若函数 $\varphi(X)$ 的数学期望存在,则

$$\langle\varphi(X)\rangle=\sum_i \varphi(x_i)p_i \tag{1.3.3}$$

$$\langle\varphi(X)\rangle = \int_{-\infty}^{\infty} \varphi(x)f(x)dx \tag{1.3.4}$$

为离散变量或连续随机变量的数学期望表式.它可以推广到多变量分布.它是一个线性算子.同时,随机变量线性组合的数学期望等于各随机变量期望的线性组合.即

$$\langle C_0+C_1X_1+\cdots+C_nX_n\rangle=C_0+C_1\langle X_1\rangle+\cdots+C_n\langle X_n\rangle \qquad (1.3.5)$$

(其中 $C_i, i=0,1,\cdots,n$ 为常数).

显然,常数的期望就是该常数本身.

按式(1.3.3)和(1.3.4),当 φ 为线性函数时 $\langle\varphi(X_1,\cdots,X_n)\rangle=\varphi(\langle X_1\rangle,\cdots,\langle X_n\rangle)$ 成立,但并不对任意函数 φ 成立,通常:

若一个非退化随机变量 X 仅仅取正值,则 X 的倒数的期望,大于期望的倒数,即

$$\left\langle\frac{1}{X}\right\rangle>\frac{1}{\langle X\rangle} \qquad (1.3.6)$$

称为 Jensen 不等式.若 X 仅取负值,相反的不等式成立.

随机变量的条件期望:

给定 y 时,X 的条件期望是指此时 X 在条件分布下的期望,离散和连续情形分别如下:

$$\langle X|y_j\rangle=\sum_i x_i\frac{p_{ij}}{p_j} \qquad p_j>0 \qquad (1.3.7)$$

$$\langle X|y\rangle=\int_{-\infty}^{\infty}x\frac{\xi(x,y)}{g(y)}dx, \qquad g(y)>0 \qquad (1.3.8)$$

上两式中,$p_j>0$ 和 $g(y)>0$ 是必要的.若将 x 或 x_i 当一随机变量 X,则 $\langle Y|X\rangle$ 也是随机变量,若 X 取 x_i,则 $\langle Y|X\rangle$ 取值 $\langle Y|x_i\rangle$;若定义函数的条件期望,则可将作为条件的随机变量视为常数,例如,令 $\varphi(X、Y)=X+Y$,$\varphi(X、Y)=XY$,则:

(1) $\langle\varphi(X、Y)|X\rangle=\langle(X+Y)|X\rangle=X+\langle Y|X\rangle$

(2) $\langle\varphi(X、Y)|X\rangle=X\langle Y|X\rangle$

对任意两个随机变量 X 和 Y,有

$$\langle\langle Y|X\rangle\rangle=\langle Y\rangle \qquad (1.3.9)$$

和

$$\langle XY\rangle=\langle X\langle Y|X\rangle\rangle \qquad (1.3.10)$$

若 $X_1,\cdots,X_n$ 是独立随机变量，则

$$\langle X_1\cdots X_n\rangle \doteq \langle X_1\rangle\langle X_2\rangle\cdots\langle X_n\rangle \tag{1.3.11}$$

以上，(1.3.3)式到(1.3.11)式，我们今后均不加证明地加以使用.

四、 矩、方差和协方差

(一)矩、方差和协方差的定义

与随机变量有关的期望有：

(1) 矩 $\langle X^r\rangle, r=1,2,\cdots$ (1.4.1)

(2) 阶乘矩 $\langle X(X-1)\cdots(X-r+1)\rangle, r=1,2,\cdots$ (1.4.2)

(3) 中心矩 $\langle X-\langle X\rangle\rangle^r, r=1,2,\cdots$ (1.4.3)

如果能确定一个随机变量 X 的分布函数 $f(x_i)$，则已可得到关于它的一切可能的信息，但实际上这往往十分困难，通常 $f(x_i)$ 很难确定，但常能得到关于 X 的矩的信息. 事实上，如上(1.4.1)式有：

$$\langle X^r\rangle = \sum_i x_i^r f(x_i) \tag{1.4.4}$$

其一阶矩 $\langle X\rangle$ 称为 X 的平均值；二阶中心矩，记为：

$$\begin{aligned}\sigma_x{}^2 &= \langle X-\langle X\rangle\rangle^2 \\ &= \langle X^2\rangle - \langle X\rangle^2\end{aligned} \tag{1.4.5}$$

称为 X 的方差，而

$$\sigma_x = (\langle X^2\rangle - \langle X\rangle^2)^{1/2} \tag{1.4.6}$$

称为 X 的标准差.

对离散随机变量 X 而言，它的矩提供了关于分布函数 $f(x_i)$ 的范围和形状的信息. 最重要的矩是较低级的几个，因为其中包含了有关分布函数整体行为的信息. 一级矩给出了 X 的平均值；标准差给出了分布函数 $f(x_i)$的宽度，标准偏差小意味着 $f(x_i)$在一级矩附近有尖锐的峰且能十分肯定 X 有接近于$\langle X\rangle$的值.

方差和标准差常用来度量分布的分散性，一个退化随机变量

或一个常数的方差应为零：

$$\sigma_c{}^2=0 \tag{1.4.7}$$

一个随机变量的方差不依赖于测量原点而有赖于测量尺度，所以

$$\sigma_{a+bx}^2=b^2\sigma_x{}^2 \tag{1.4.8}$$

对于具有联合分布的两个随机变量 X 和 Y，我们发现

$$\sigma_{xy}=\langle X-\langle X\rangle\rangle\langle Y-\langle Y\rangle\rangle \tag{1.4.9}$$

$$=\langle XY\rangle-\langle X\rangle\langle Y\rangle \tag{1.4.10}$$

记 $\sigma_{xy}=\mathrm{Cov}(X、Y)$表示 X、Y 的协方差. 如果 X 和 Y 为随机独立，$\sigma_{xy}=0$，但并不意味着随机变量的分布是独立的.

对于连续随机变量，我们可令 X 在区间 $a\leqslant X\leqslant b$ 定义为一个具有连续值的集合. 假定存在一个分段连续的函数 $f(x)$，使 X 的值落在 $a\leqslant X\leqslant b$ 中的概率 $P(a\leqslant X\leqslant b)$等于在 $x=a$ 和 $x=b$ 之间的曲线 $f(x)$下面的面积(图 1—4—1)：

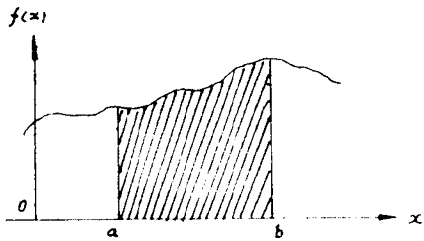

图 1—4—1　阴影部分的面积为发现随机变量 X 处在 $a\leqslant X\leqslant b$ 内的概率

$$P(a\leqslant X\leqslant b)=\int_a^b f_x(x)dx \tag{1.4.11}$$

此处 $f_x(x)$是 X 的概率密度，并满足条件

$$f_x(x)\geqslant 0 \tag{1.4.12}$$

和

$$\int_a^b f_x(x)dx=1 \tag{1.4.13}$$

连续随机变量的第 n 级矩定义为：

$$\langle X^n\rangle = \int x^n f_x(x)dx \tag{1.4.14}$$

$\langle X\rangle$仍为 X 的平均值或期望. 方差和标准差的定义与(1.4.5)和(1.4.6)式相同. 若所有的矩$\langle X^n\rangle$都知道，则概率密度完全被确定. 为了清楚地看出这一点，引入特征函数 $\varphi_X(k)$. 定义：

$$\varphi_X(k) = \langle e^{ikx}\rangle = \int e^{ikx} f_x(x)dx = \sum_{n=0}^{\infty} \frac{(ik)^n\langle X^n\rangle}{n!} \tag{1.4.15}$$

上式中的级数展开，须当$\langle X^n\rangle$足够小使得级数收敛时才有意义. $f_x(x)$是特征函数的 Fourier 变换.

$$f_x(x) = \frac{1}{2\pi}\int e^{-ikx}\varphi_X(k)dk \tag{1.4.16}$$

所以，只要所有矩知道，则通过(1.4.15)和(1.4.16)即可将 $f_x(x)$ 完全确定. 若给出了特征函数，则又可通过微商求矩

$$\langle X^n\rangle = \frac{1}{i^n}\frac{d^n\varphi_X(k)}{dk^n}\Big|k=0 \tag{1.4.17}$$

显然，若已知 $f_x(x)$，用上法求矩较为简单.

另若，定义累积量展开为：

$$\varphi_X(k) = \exp\left[\sum_{n=1}^{\infty}\frac{(ik)^n}{n!}C_n(X)\right] \tag{1.4.18}$$

其中 $C_n(X)$为第 n 级累积量，将(1.4.15)和(1.4.18)按 k 的幂次展开并令同次项相等，则所得前四个累积量可表示如下：

$$C_1(X) = \langle X\rangle \tag{1.4.19a}$$

$$C_2(X) = \langle X^2\rangle - \langle X\rangle^2 \tag{1.4.19b}$$

$$C_3(X) = \langle X^3\rangle - 3\langle X\rangle\langle X^2\rangle + 2\langle X\rangle^3 \tag{1.4.19c}$$

$$C_4(X) = \langle X^4\rangle - 3\langle X^2\rangle^2 - 4\langle X\rangle\langle X^3\rangle + 12\langle X^2\rangle\langle X\rangle^2 - 6\langle X\rangle^4 \tag{1.4.19d}$$

若更高级累积量迅速趋于零，则常可于(1.4.18)展式中保留前面几个累积量而得到 $\varphi_X(k)$的较好近似.

对于与 $f_x(x)$ 相关的随机变量如 $Y=G(x)$ 的概率密度，常可表为：

$$f_Y(y)=\int\delta(y-G(x))f_x(x)dx \tag{1.4.20}$$

其中 $G(x)$ 是随机变量的已知函数而 $\delta(y-G(x))$ 是 δ 函数.

对(1.4.9)和(1.4.10)所描述的协方差，随机变量 X 和 Y 的联合概率密度 $f(x,y)$ 须满足

$$f(x,y)\geqslant 0$$

和
$$\iint f(x,y)dxdy=1 \quad (\text{归一}) \tag{1.4.21}$$

(二)随机变量线性函数的方差

定义记号 $V_{ar}(X)$ 表示方差，则随机变量线性函数的方差为

$$\begin{aligned}
&V_{ar}(a+bX+cY)\\
&=\langle a+bX+cY-\langle a+bX+cY\rangle\rangle^2\\
&=\langle b[X+\langle X\rangle]+c[Y-\langle Y\rangle]\rangle^2\\
&=b^2\langle X-\langle X\rangle\rangle^2+c^2\langle Y-\langle Y\rangle\rangle^2+2bc\langle X-\langle X\rangle\rangle\langle Y-\langle Y\rangle\rangle\\
&=b^2{\sigma_X}^2+c^2{\sigma_Y}^2+2bc\sigma_{XY}
\end{aligned}$$

则线性函数 $C_0+C_1X_1+\cdots+C_nX_n$ 的方差为：

$$\begin{aligned}
&V_{ar}(C_0+C_1X_1+\cdots+C_nX_n)\\
&=\sum_{i=1}^{n}C_i^2\sigma_{X_i}^2+\sum_{\substack{i=1\\ i\neq j}}^{n}\sum_{j=1}^{n}C_iC_j\sigma_{X_i,X_j}
\end{aligned} \tag{1.4.22}$$

特别地，若 $X_1,\cdots,X_n$ 为独立随机变量，则

$$V_{ar}(C_1X_1+\cdots+C_nX_n)=C_1^2\sigma_{X_1}^2+\cdots+C_n^2\sigma_{X_n}^2 \tag{1.4.23}$$

若 $X_1,\cdots,X_n$ 互相独立，有相同的方差 σ_X^2，则随机变量 X 的平均值

$\langle X\rangle=\dfrac{1}{n}(X_1+\cdots+X_n)$ 的方差为：

$$V_{ar}\langle X\rangle=\frac{1}{n}\sigma_X^2 \tag{1.4.24}$$

(三)随机变量两个线性函数间的协方差