随机过程

方兆本　缪柏其　编著

中国科学技术大学出版社

随 机 过 程

方兆本 缪柏其 编著

中国科学技术大学出版社
1993·合肥

(皖)新登字08号

随 机 过 程

方兆本 缪柏其 编著

中国科学技术大学出版社出版

(安徽省合肥市金寨路96号,邮政编码:230026)

金寨县印刷厂印刷

安徽省新华书店发行

开本:850×1168/32　印张:5.625　字数:143千

1993年1月第1版　1993年1月第1次印刷

印数:1—4000册

ISBN7-312-00379-6/O・122　定价:3.50元

内 容 提 要

本书介绍在应用中经常遇到的几种基本随机过程，如泊松过程，更新过程，马尔柯夫过程，平稳过程，布朗运动，以及线性随机微分方程，材料丰富．每章结合大量有实际背景的例子来解释基本概念，并配有一定量的习题．

本书可作为理、工科大学生和研究生的教学用书或教学参考书，也可作为工程技术人员应用随机过程的入门参考书．

前　　言

本书是作者与中国科学技术大学数学系同事们多年从事本课程教学的积累。它是为大学理工科本科生、研究生概率统计公共课所编写的教材、是陈希孺教授所著的《概率论与数理统计》的续篇。本书可以作非数学系应用数学辅修专业的随机过程课的教本，主要讨论随机过程的基本理论及其应用。每章正文之后有配套的习题供读者练习。随机过程是一门应用性很强的学科、各个领域中的科技工作者都能从中发现有启发性的模型。单纯照搬模型比较容易，难的是在实际问题中简化条件提炼出恰当的随机模型。学习这门课程应在这方面多做努力，这也是本书编写的宗旨。

学习本书要有微积分和初等概率论的基础，兼顾内容阐述的需要和学生的实际接受能力书中用到了矩母函数、生成函数、复变函数、微分方程求解及矩阵代数等数学工具。对某些工具不熟悉的读者可以跳过证明推理直接阅读有关结论。这并不影响对本课程的基本了解。

使用本教材的教员自然可以根据课时限制及各系各科的不同需求而有所侧重。比如，删去2.3节，3.3节中定理3.3和3.4的证明，3.4、3.6节、4.2节的定理4.2、4.4节、以及5.3节将仍是一份应用随机过程的 *ABC* 教材。书末曾备有习题答案或提示，但考虑到附在书后出版对教与学无益故予以删除。作者感谢胡太忠同志帮助演算了书中的全部习题。

作　　者

1991年9月于合肥

目　　录

第一章 引　论

1.1 引　言

1.1.1 基本概念和例子

随机过程是对一连串随机事件间动态关系的定量描述。它是在自然科学、工程科学、社会科学各领域研究随机现象的有力工具。其应用包罗万象：气象预报、天文观测、通讯工程、原子物理、宇航遥控、生物医学、管理科学、运筹决策、计算机科学、经济分析、人口理论、可靠性与质量控制等许许多多领域都已离不开用随机过程的理论来建立各种数学模型。

一般，把一族随机变量定义为随机过程。英文叫 stochastic process。"stochastic"一词源于希腊语"$\sigma\tau o\chi\alpha\sigma\tau\iota\kappa o\xi$"意思是"猜"。但这门科学不是乱猜。在研究随机过程时人们透过表面的偶然性找出必然的内在规律并以概率的形式来描述这些规律。从偶然中悟出必然正是这一学科的魅力所在。

随机过程的早期历史属于物理领域。人们可以追溯到 Gibbs，Boltzman，Poincaré 等人在统计力学中的研究以及后来 Einstein，Wiener，Lévy 等人对布朗运动的开创性工作。而 Erlang 等则在电话流中研究了 Poisson 过程。而整个学科的理论基础则是由 Kolmogorov 和 Doob 奠定的。"stochastic"这一用词也在这时流行。生灭过程是 Feller 首先引进的。Cramer 和 Lévy 研究了平稳过程。Xinchin，Palm 发展了排队论中的过程理论。Doob 则研究 Markov 过程和鞅。这些都是早期研究

的重要里程碑。目前，这一学科仍在理论和应用两方面以空前的深度和广度在迅速发展着。下面对随机过程作正式定义：

定义 1.1 随机过程就是一族随机变量$\{X(t), t\in T\}$，其中t是参数它属于某个指标集T，T称为是参数集。

一般，t代表时间.当$T=\{0,1,2,\cdots\}$时也称随机过程为随机序列。对$X(t)$可以这样看：随机变量是定义在空间Ω上的，所以是随t与$\omega\in\Omega$而变化的，于是可以记为$X(t, \omega)$。当固定一次随机试验，即取定$\omega_0\in\Omega$时，$X(t,\omega_0)$就是一条样本路径它是t的函数。它可能是连续的，也可能是有间断点和跳跃的。这是我们通常所观察到的过程.另一方面固定了时间$t=t_0$，$X(t_0,\omega)$就是一个随机变量，其取值随着随机试验的结果而变化。变化有一定的规律、叫做概率分布。随机过程$X(t)$取的值称作是过程所处的状态。状态的全体称为状态空间。根据T及状态空间的不同我们可以对过程进行分类。依照状态空间可分为连续状态和离散状态；依参数集T，当T为有限集或可列集则称之为离散参数过程否则称为是连续参数过程。当T是高维向量则$X(t)$称作是随机场。

例 1.1 英国植物学家 Brown 注意到漂浮在液面上的微小粒子不断进行无规则的运动。这种运动叫做 Brown 运动。它是分子大量随机碰撞的结果。若记$(X(t),Y(t))$为粒子在平面坐标上的位置，则它是平面上的 Brown 运动。在统计物理中对它有深入的研究。

例 1.2 一醉汉在路上行走，以概率p前进一步，概率$1-p$后退一步。以$X(t)$记他在街上的位置则$X(t)$就是直线上的随机游动。

例 1.3 神经细胞在细胞膜的位势达到某一临界值C时就要兴奋。刺激和抑制两种脉冲以一定的速率(比如 Poisson 过程)抵达细胞。前者使位势升高；后者使位势降低。升降的幅度服从相同的分布$H(x)$。神经细胞在兴奋过后位势恢复到 0，过程再度重

复。记 T_i 为两次兴奋的间隔时间，并记 $X(t)$ 为时刻 t 时细胞膜的位势则过程的一次实现如图 1.1 所示：

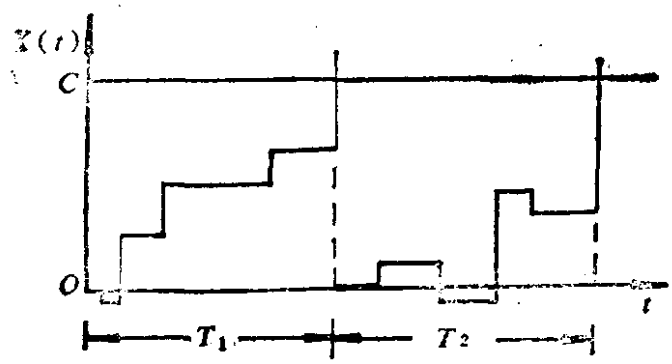

图 1.1 神经细胞膜的位势

例 1.4 到达总机交换台的呼叫次数为 Poisson 过程。每次呼叫是相互独立的，而间隔时间服从指数分布。交换台在同一时刻只能接通 K 个呼叫。人们常要了解在某一时刻的排队长度以及呼叫的平均等待时间。这是一种排队模型。

例 1.5 流行病学的研究中有如下模型：在时刻 0 时易感人群大小为 $X(0)$，$Y(0)$ 是已受传染的人数。假定易感人群被传染的概率为 p，则经过一段传染周期后（记之为单位时间）$X(0)$ 中有 $X(1)$ 没有染上病而 $Y(1)$ 却受到传染。传染过程一直漫延到再没有人会染上这种流行病时停止。于是 $X(t)=X(t+1)+Y(t+1)$，且当 $j\leqslant i$ 时有

$$P(X(t+1)=j|X(t)=i)=\binom{i}{i-j}p^{i-j}(1-p)^{j}.$$

$\{X(t),\ t=0,1,2,\cdots\}$ 就是以上式为状态转移概率的 Markov 过程。

例 1.6 设 $X(t)$ 为信号流，它满足方程

$$X(t+1)=aX(t)+\varepsilon(t),$$

其中 a 为实参数，$\varepsilon(t)$ 代表误差。真正的信号 $X(t)$ 并不可能观察到。人们所能观察到的是

$$Y(t)=X(t)+Z(t),$$

$Z(t)$为噪声。从观察值$Y(t)$出发，检测出$X(t)$是通讯工程中的重要课题。这是过程的预测与滤波的问题。

例 1.7 水库库容调度。记$Y(t)$为$(t, t+1)$年间的水库蓄水量，它是随机的。M为每年年底固定的清库泄洪量，$X(t)$为泄洪之后的库容量。则$X(t+1)=\min\{X(t)+Y(t), K\}-\min\{X(t)+Y(t), M\}$其中$K$为大坝的设计库容。过程$X(t)$为一Markov过程。把水库库容改为商店仓库的库存模型类似。

例 1.8 记$X(t)$为时刻t的商品价格。若$X(t)$适合线性模型

$$X(t)+\alpha_1 X(t-1)+\alpha_2 X(t-2)+\cdots\alpha_p X(t-p)$$
$$=Z(t)+\beta_1 Z(t-1)+\cdots+\beta_q Z(t-q).$$

其中α_k，β_j为实参数，$Z(t)$为独立同分布的不可观察的随机变量。则$X(t)$服从所谓ARMA模型——混合自回归滑动平均模型。这是在经济预测中十分有用的时间序列模型。

1.1.2 有限维分布和数字特征

知道了分布函数$F_X(x)=P(X\leqslant x)$就能了解随机变量X。类似的，对随机过程$\{X(t), t\in T\}$定义$F_t(x)=P\{X(t)\leqslant x\}$为过程的一维分布，$X(t)$的期望$E[X(t)]$为过程的均值函数记作$\mu_X(t)$。而$\mathrm{Var}[X(t)]$则被定义为过程的方差函数。不仅如此，我们还需要了解随机变量$X(t_1)$与$X(t_2)$的联合分布$P\{X(t_1)\leqslant x_1, X(t_2)\leqslant x_2\}$，这也就是过程在$t_1$，$t_2$两不同时刻值的联合二维分布记作$F_{t_1, t_2}(x_1, x_2)$。$\mathrm{Cov}(X(t_1), X(t_2))$则称为是过程的协方差函数，记作$r_X(t_1, t_2)$。于是有

$$r_X(t_1, t_2)=E\{(X(t_1)-\mu_X(t_1))(X(t_2)-\mu_X(t_2))\}. \quad (1.1)$$

协方差函数显然有对称性，即对任何s, t有$r_X(t, s)=r_X(s, t)$。当$t_1=t_2=t$时。$r_X(t, t)=\mathrm{Var}X(t)$。容易证明协方差函数是非负定的，也就是说对任何$t_1, t_2, \cdots, t_n\in T$及任意实数$b_1, \cdots, b_n$我们恒有

$$\sum_{i=1}^{n}\sum_{j=1}^{n} b_i b_j r_X(t_i, t_j) \geqslant 0。\tag{1.2}$$

因为协方差运算有线性性质，由

$$\begin{aligned}\operatorname{Var}\left(\sum_{j=1}^{n} b_j X(t_j)\right) &= \operatorname{Cov}\left(\sum_{i=1}^{n} b_i X(t_i), \sum_{j=1}^{n} b_j X(t_j)\right)\\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} b_i b_j r_X(t_i, t_j)\end{aligned}$$

知(1.2)式成立故它是非负定的。进而，我们定义随机过程的有限维分布族它是 $F_{t_1, t_2, \cdots t_n}(x_1, \cdots, x_n), t_1, \cdots, t_n \in T$ 全体。其中

$$F_{t_1, \ldots, t_n}(x_1, \cdots, x_n) = P\{X(t_1) \leqslant x_1, \cdots, X(t_n) \leqslant x_n\}。\tag{1.3}$$

知道了随机过程的有限维分布族就知道了过程$\{X(t), t \in T\}$中任意 n 个随机变量的联合分布。也就完全了解了这些变量之间的相互依赖关系。有限维分布也有对称性，它与变量 $X(t_1), \cdots, X(t_n)$ 的排列次序无关。对$\{1, \cdots, n\}$的任一置换$(i_1, \cdots, i_n)$有

$$F_{t_{i_1}, \ldots, t_{i_n}}(x_{i_1}, \cdots, x_{i_n}) = F_{t_1, \ldots, t_n}(x_1, \cdots, x_n)。$$

有限维分布还有相容性，意思是当某些 $x \to \infty$时高维分布的边缘分布与相应的低维分布是一致的。也即对 $m<n$ 有

$$\begin{aligned}&F_{t_1, \cdots, t_m, t_{m+1}, \cdots, t_n}(x_1, \cdots, x_m, \infty, \cdots, \infty)\\ &= F_{t_1, \cdots, t_m}(x_1, \cdots, x_m)。\end{aligned}$$

有趣的是若一族给定的分布函数有上述对称性和相容性则保证了存在一个随机过程$\{X(t), t \in T\}$使它的有限维分布族正好就是给定的分布函数族。这是 1931 年由 Kolmogorov 证明的相容性基本定理。

例 1.9 设 X_n 为第 n 次独立地扔一六面骰子的结果。则$\{X_n, n \geqslant 1\}$为一随机过程，参数集 T 为$\{1, 2, 3, \cdots\}$。而状态空间为$\{1, 2, 3, 4, 5, 6\}$。序列 3，1，4，6，3，2，5，5，1，…就是$\{X_n, n \geqslant 1\}$的一次可能的实现。均值函数 $EX_n = EX_1 \equiv 3.5$，除方差函数为 17.5 外协方差函数恒等于 0。任何有限维分布 $F_{n_1, \ldots, n_k}(x_1, \cdots, x_k) = F(x_1) \cdots F(x_k)$其中 $F(x)$为 X_1 的分布函数。

当然这种独立情形是太过于简单了。对于大多数随机过程，

$X(t)$之间是相依的。研究他们之间的相依关系是我们的主要课题。根据相依关系的不同，人们可以研究随机过程的不同类型。

1.1.3 平稳过程和独立增量过程

如果两个随机变量X_1，X_2的分布函数$F_{X_1}(x)$与$F_{X_2}(x)$对任何x都是相等的则称它们是等分布的记作$X_1 \overset{d}{=} X_2$。类似的如果一个随机向量$\boldsymbol{X}=(X_1,\cdots,X_n)$与另一随机向量$\boldsymbol{Y}=(Y_1,\cdots,Y_n)$有相同的联合分布则也称他们是等分布的记作$\boldsymbol{X}\overset{d}{=}\boldsymbol{Y}$。有一类重要的随机过程，它处于某种概率平衡状态其主要性质只与变量$X(t)$之间的时间间隔有关而与我们考查的起始点无关。这类过程叫做平稳过程。具体而言我们有

定义 1.2 如果随机过程$X(t)$对任意的$t_1,\cdots,t_n\in T$和任何h有

$$(X(t_1+h),\cdots,X(t_n+h))\overset{d}{=}(X(t_1),\cdots,X(t_n)), \quad (1.4)$$

则称作是严格平稳的。

条件 (1.4) 很强也不易验证。所以退而求其次有所谓宽平稳或二阶平稳过程。引入

定义 1.3 如果随机过程$X(t)$的所有二阶矩存在并有$EX(t)=m$及协方差函数$r_X(t, s)$只与时间差$t-s$有关则称作是宽平稳的或二阶矩平稳的。

对于宽平稳过程由于对$-\infty<s,t<\infty, r_X(s,t)=r_X(0,t-s)$所以可以记之为$R_X(t-s)$。显然对所有$t$ $R_X(-t)=R_X(t)$，即为偶函数。所以$R_X(t)$的图形是关于坐标轴对称的。其在 0 点的值$R_X(0)$就是过程的方差函数，即$R_X(0)=\mathrm{Var}X(t)$。

尽管$X(t)$之间常常不是相互独立的，但人们可以假定过程的增量之间是相互独立的。这就是

定义 1.4 如果对任何$t_1,\cdots,t_n\in T$，$t_1<t_2<\cdots<t_n$随机变量$X(t_2)-X(t_1),X(t_3)-X(t_2),\cdots,X(t_n)-X(t_{n-1})$是相互独立的

则$X(t)$称为是独立增量过程。如果进一步有对任何t_1，t_2
$X(t_1+h)-X(t_1)\overset{d}{=}X(t_2+h)-X(t_2)$则过程称为是有平稳独立增量的过程。

可以证明平稳独立增量过程的均值函数一定是t的线性函数。我们以后要介绍的Poisson过程和Brown运动都是这类过程。这两类过程是过程理论中的两块最重要的基石。

例1.10 设Z_i，$i=0,1,2,\cdots$是一串独立同分布的随机变量定义$X_n=\sum_{i=0}^{n}Z_i$则过程$\{X_n, n\geqslant 0\}$就是独立增量过程。一般称X_n为独立和。

重要的相依类型还有Markov链及Markov过程(这构成了第三章的主题)、更新过程和鞅等。更新过程是运筹学、排队论等管理学科中的重要工具；鞅是近代概率论及随机过程理论中的重要概念。受篇幅及课时限制都无法一一展开讨论。读者应注意到过程的分类不是绝对的，一个具体的过程可以同时属于上述多种类型。比如第二章要讨论的Poisson过程既有独立增量又有平稳增量。它既是连续时间的Markov链又是一类特殊的更新过程。若过程的参数为λ则Poisson过程$X(t)$减去均值函数即$X(t)-\lambda t$还是一个鞅。

1.2 条件期望、矩母函数

1.2.1 条件期望

要研究随机过程离不开讨论一族随机变量相互之间的关系。这时常要用到条件概率和条件期望这些基本的概念和工具。本节将对有关概念和工具作一简要回顾。学过概率论的读者都一定知道什么是事件的条件概率，即给定事件B发生时事件A发生的条件概率$P(A|B)$等于$P(A\cap B)/P(B)$。但究竟给了随机变量

Y 的取值后另一随机变量 X 的条件期望是什么却未必能说清楚。先看离散型随机变量 X 和 Y：一般对所有使 $P\{Y=y\}>0$ 的 y 定义给定 $Y=y$ 时 X 取 x 的条件概率为

$$P\{X=x|Y=y\}=\frac{P\{X=x, Y=y\}}{P\{Y=y\}}. \tag{1.5}$$

而给定 $Y=y$，X 的条件分布函数则定义为

$$F(x|y)=P\{X\leqslant x|Y=y\}. \tag{1.6}$$

相应的，给定 $Y=y$，X 的条件期望定义为

$$E(X|Y=y)=\sum_x xP\{X=x|Y=y\}. \tag{1.7}$$

对一般的连续型随机变量 $P(Y=y)$ 往往为 0，比式(1.5)没有意义我们怎么来定义条件期望呢？常用的办法是如果对任何包含 y 的小区间总有 $P(Y\in\Delta y)=0$，则定义 $P(X\in A|Y=y)=0$。否则，若 $P(Y\in\Delta y)>0$ 就可以定义

$$P(X\in A|Y=y)=\lim_{\Delta y\downarrow 0}P(X\in A|Y\in\Delta y).$$

这里 $\Delta y\downarrow 0$ 的意思是使包含 y 的小区间的长度缩小为 0。除了个别例外的 y 值这一极限总是存在。相应地，定义给定 $Y=y$ 时，X 的条件分布函数为

$$P(X\leqslant x|Y=y)=\lim_{\Delta y\downarrow 0}P(X\leqslant x|Y\in\Delta y), \tag{1.8}$$

记作 $F(x|y)$。进一步，如果存在一非负函数（记为）$f(x|y)$ 使得对任何集合 A 恒有

$$P(X\in A|Y=y)=\int_A f(x|y)\mathrm{d}x, \tag{1.9}$$

且 $\int f(x|y)\mathrm{d}x=1$ 则 $f(x|y)$ 称为是给定 $Y=y$ 时 X 的条件密度。不难看出条件分布函数是关于条件密度对变量 x 从 $-\infty$ 到 x 的积分，即

$$F(x|y)=\int_{-\infty}^{x}f(x|y)\mathrm{d}x. \tag{1.10}$$

如果把 X 与 Y 的联合密度 $f(x, y)$ 看作质量为 1 的平板的密度

(面密度)则条件密度 $f(x|y)$ 就是固定 $Y=y$ 时的线密度。它们之间的关系是

$$f(x,y)=f(x|y)f(y), \tag{1.11}$$

其中 $f(y)$ 为随机变量 Y 的边缘密度。此时，给定 $Y=y$，X 的条件期望为

$$E(X|Y=y)=\int xf(x|y)\mathrm{d}x。 \tag{1.12}$$

(1.7)式与(1.12)式本质上是一回事，在更深一点的课程中人们把它们统一记为

$$E(X|Y=y)=\int x\mathrm{d}F(x|y)。 \tag{1.13}$$

在行文不发生混淆时，也常记为 $E(X|Y)$。严格地讲(1.13)表示一个数值(与 y 有关)而 $E(X|Y)$ 则为随机变量。

例 1.11 扔一硬币出现正面的概率为 p，独立地作投币试验。记 S 为 n 次试验中出现正面的总次数，并设首次出现正面是在第 T 次试验。问题是求给定 n 次试验中仅出现了一次正面时随机变量 T 的条件概率分布，也即 $P(T=k|S=1)$。

一种想法认为给定了正面出现一次它会以相同的机会在 n 次试验中的任一次出现所以 T 的条件概率分布应该是 $1,2,\cdots,n$ 上的均匀分布。但有人也许会争辩：因为给定了只出现一次正面它很可能会较早出现。让我们来实际算一算：事件 $\{T=k,S=1\}$ 只包含一个序列，它仅在第 k 次试验时结果为正面。所以 $P\{T=k,S=1\}=p(1-p)^{n-1}$，而 $P(S=1)=\binom{n}{1}p(1-p)^{n-1}$。因此有 $P(T=k|S=1)=P(T=k,S=1)/P(S=1)=1/n$。可见前一看法是正确的。

将上例稍作推广，如果粒子依照参数为 λ 的 Poisson 分布进入计数器，即若记 S 为单位时间区间[0,1]上的粒子总数则有

$$P(S=k)=\frac{\lambda^k e^{-\lambda}}{k!},\ k=0,1,2\cdots$$

问题是给定在[0，1]时间段上仅有一个粒子到达计数器时该粒子到达时间 T 将服从什么分布？读者很自然会想到是[0，1]上的均匀分布。我们可以把[0，1]区间分成长度为 Δt 的 n 个小区间这时 $\Delta t=\frac{1}{n}$。近似地，好比作了 n 次独立试验粒子以概率 $\lambda\Delta t$ 落入每一小区间，而不在其中的概率为 $1-\lambda\Delta t$。由例 1.11 知

$$P((k-1)\Delta t\leqslant T<k\Delta t\,|\,S=1)=\frac{1}{n}=\Delta t.$$

取极限即知给定 $S=1$，T 是[0，1]上的均匀分布。本章习题 10 提供另一解法，不妨一试。

条件期望有些重要的性质我们总结为下面的

命题 1.1 *a*）若 X 与 Y 独立，则 $E(X|Y=y)=EX$。*b*）条件期望有所谓平滑性：

$$EX=\int E(X|Y=y)dF_Y(y)=E[E(X|Y)].\qquad(1.14)$$

c）对随机变量 X，Y 的函数 $\phi(X,Y)$ 恒有

$$E[\phi(X,Y)|Y=y]=E[\phi(X,y)|Y=y].\qquad(1.15)$$

证明：注意到当 X 与 Y 独立时，给定 $Y=y$ 的 X 的条件分布与无条件分布是一样的。*a*）由此不难得证。对于 *b*）仅证明离散随机变量的情形：

$$EX=\sum_k E(X|Y=y_k)P(Y=y_k).\qquad(1.16)$$

由定义 $E(X|Y=y_k)=\sum_j x_jP(X=x_j|Y=y_k)$，再由条件概率的定义(1.16)式右边为 $\sum_k\sum_j x_jP(X=x_j,\ Y=y_k)$ 先对 k 求和即得 $\sum_j x_jP(X=x_j)$ 由定义正是 EX。*c*）的证明留给读者作为练习。

命题 1.1 *b*）告诉我们计算 X 的期望可以分两步走，先计算给定 $Y=y$ 时 X 的条件期望。再对这条件期望作加权平均就是总的平均或期望。这是一个很有用的思想。

条件期望是概率统计中极其重要的概念，在均方误差最小准