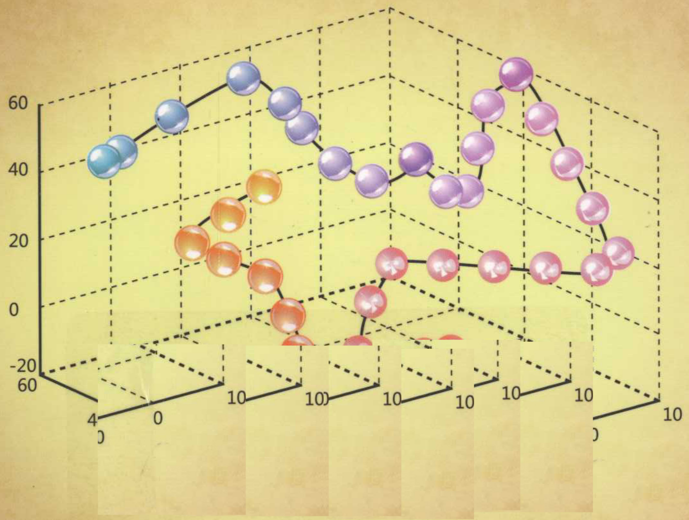
TURING 图灵数学·统计学丛书 44



随机过程

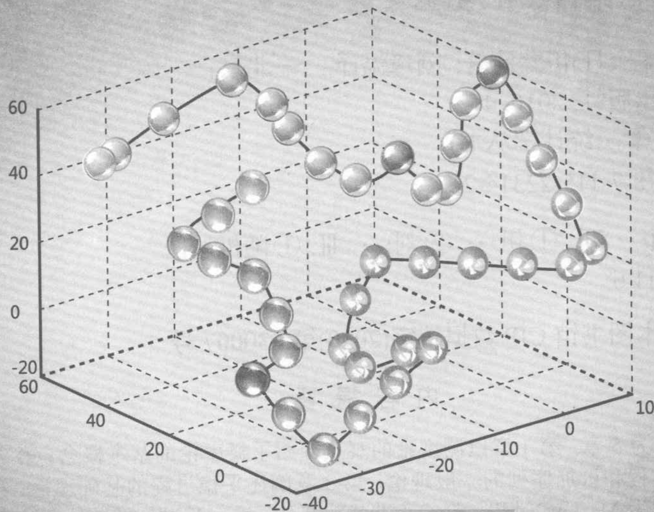
[日] 伊藤清 著

刘璋温 译

郑绍濂 审校

人民邮电出版社
POSTS & TELECOM PRESS

TURING 图灵数学·统计学丛书 44



随机过程

[日] 伊藤清 著
刘璋温 译
郑绍濂 审校

人民邮电出版社
北京

图书在版编目(CIP)数据

随机过程/(日)伊藤清著；刘璋温译. —北京：人民邮电出版社，2010.4
(图灵数学·统计学丛书)
ISBN 978-7-115-22314-2

Ⅰ.①随… Ⅱ.①伊… ②刘… Ⅲ.①随机过程 Ⅳ.①O211.6

中国版本图书馆 CIP 数据核字(2010)第 030007 号

内 容 提 要

本书分为 5 章. 第 1 章以测度论的观点介绍了概率论的基本概念；第 2 章叙述可加过程和可加序列的一般理论；第 3 章阐述平稳过程的基础理论；第 4 章和第 5 章为 Markoff 过程，前一章讲基础部分，后一章讨论扩散的一些现代理论和方法.

本书可供高等院校数学系、物理系等相关专业师生及工程师作参考.

图灵数学·统计学丛书

随机过程

◆ 著 [日] 伊藤 清
译 刘璋温
审校 郑绍濂
责任编辑 明永玲

◆ 人民邮电出版社出版发行 北京市崇文区夕照寺街 14 号
邮编 100061 电子函件 315@ptpress.com.cn
网址：http://www.ptpress.com.cn
北京隆昌伟业印刷有限公司印刷

◆ 开本：880×1230 1/32
印张：6.5
字数：208 千字 2010 年 4 月第 1 版
印数：1–3 000 册 2010 年 4 月北京第 1 次印刷

著作权合同登记号 图字：01-2008-3847 号

ISBN 978-7-115-22314-2

定价：25.00 元

读者服务热线：(010) 51095186 印装质量热线：(010) 67129223

反盗版热线：(010) 67171154

版权声明

译 者 序

本书是根据伊藤清所著的《随机过程 I》和《随机过程 II》翻译的. 伊藤的原著是日本岩波书店从 1957 年起连续出版的一套"岩波讲座——现代应用数学"基础篇中的两本.

随机过程是概率论的一个主要组成部分, 它的研究与现代科学技术有密切的关系. 本书以不大的篇幅用简练的笔法对随机过程论的几个主要方面和一些最新成就作了精练和严格的阐述. 具有初步概率论和泛函分析知识的读者, 通过阅读这本书, 可以较快地掌握现代随机过程的基本理论和发展方向.

本书的翻译工作起于 E. B. Dynkin 教授来我国讲学, 他通过译者的介绍了解到原书的内容后, 非常称赞, 并建议早日把它翻译出来. 在他的指导下,《随机过程 I》的俄译本已于 1960 年出版. 本译稿是在 1958 年 6 月底完成的, 其中第 4 章的前几节为余潜修先生所译. 随后译稿一直流传于北京大学、复旦大学和南开大学, 作为专门化的参考教材. 这次承复旦大学郑绍濂同志和南开大学王梓坤同志分别整理前三章和后两章, 并由郑绍濂同志统一校订. 校样排出后, 译者重新校阅了一遍, 根据原文又作了一些修正.

本书的出版与上列许多同志的帮助分不开, 译者在此对他们表示深切的感谢.

刘璋温

1961 年 4 月北京

前　　言

1957 年, 岩波书店曾将拙著《随机过程 I》和《随机过程 II》作为 “岩波讲座 —— 现代应用数学” 中的两个分册出版, 现又作为单行本再次发行. 本书概括介绍了随机过程的三部分重要内容, 即可加过程、平稳过程和 Markoff 过程, 还讲解了一维扩散过程. 特别是针对一维扩散过程, 本书第一次以随机分析的方式规范介绍了有关局部构造和边界点分类的内容, 而在当时这些才刚刚由 William Feller 发现. 为此, 我本人感到颇为自豪. 然而, 从那时起到现在本书再版, 已经过去半个世纪, 身为著者确实感慨颇深.

本书首版前, 拙著《概率论》(岩波书店, 1953) 中也概括讲解了有关可加过程、平稳过程和 Markoff 过程的内容. 本书首版后, 在以下两本英文拙著中还详细讲解了可加过程和 Markoff 过程:

Lectures on Stochastic Processes, Tata Institute of Fundamental Research, Bombay, 1960

Stochastic Processes, edited by Ole E. Barndorff-Nielsen and Ken-iti Sato, Springer, 2004 (原本作为 Aarhus 大学讲义于 1969 年出版)

在以上列出的 Tata 研究所的讲义录以及我与 H. P. McKean 合著的另外两本书

Diffusion Processes and their Sample Paths, Springer, 1965;

Classics in Mathematics, Springer, 1996

的合订本中还详细讲解了一维扩散过程的局部构造. 在该合订本的 4.6 节中, 用概率论的方法证明了一维扩散过程中齐次方程解的边界附近的具体性质. 为便于不太熟悉概率论的读者理解, 本书的 §60 和 §61 用分析的方法给出其证明.

在编写本书时, 随机过程的规范研究才刚刚兴起. 正如 “后记” 所述, 在当时相关文献信息极为不足. 在随后的半个世纪中, 随机过程研究得到极大发展, 国内外都出版了大量优秀的相关书籍. *Stochastic Processes*, Springer, 2004 一书的著者前言和编者序中都列举了一些相关书目.

本书首次出版后, 在 E. B. Dynkin 的指导下, A. D. Wentzell 将本书翻译成俄文, 并分别于 1960 年和 1963 年在莫斯科出版了第 I 册和第 II 册. 此外, 在 1959 年, 已故耶鲁大学教授角谷静夫注意到本书记述的一维扩散过程, 建议当时还是耶鲁大学研究生院学生的伊藤雄二把第 II 册翻译为英文. 最后, 油印了打字机原稿, 并分发给耶鲁大学校内外的数学研究人员. 历经半个世纪, 本书的两个分册都由伊藤雄二翻译为英文, 并作为美国数学学会数学翻译丛书中的 1 卷在*Essentials of Stochastic Processes*系列下出版. 此事令我极为欣喜. 伊藤雄二辛勤翻译期间, 福岛正俊和渡边信三仔细订正了首版中的错字等细小错误, 还替我适当修订了第 4 章和第 5 章的部分内容. 此次再版采用了与英译时相同的修订. 这里再次对福岛和渡边两位的努力以及岩波书店吉田宇一的帮助表示衷心感谢.

今年 6 月, 正在为本书新出版的单行本做准备时, 我终生的恩师弥永昌吉先生突然辞世. 而前述英译单行本的校对即将完成之时, 于 2004 年 8 月收悉角谷静夫先生的讣讯. 从 1942 年我开始写本领域论文以来, 两位先生就经常热情地给予我建议和鼓励, 到现在已历经 60 多年. 当然, 本书首版时我也得到了两位先生的极大帮助. 两位先生不能喜见装订精美的英文版和日文版新书, 甚为遗憾. 追忆两位先生生前的身影, 我衷心地为他们祈求冥福.

著 者

于 2006 年 12 月

目 录

第1章　基本概念

§1　测度论观点下的概率论 (1) 直观的背景

假设甲、乙二人掷硬币, 以先掷出正面者为胜. 现在试让甲先掷并考虑下面两个问题.

(i) 甲胜的概率是多少?

(ii) 直到决定胜负为止所掷的平均次数是多少?

首先, 观察这种比赛, 它可能出现哪些结果. 若以 O 表示掷出正面, U 表示掷出反面, 则可能产生的结果为

$$\left\{\begin{array}{l} \omega_1 = \mathrm{O} \\ \omega_2 = \mathrm{UO} \\ \omega_3 = \mathrm{UUO} \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \omega_n = \underbrace{\mathrm{UU}\cdots\mathrm{U}}_{(n-1)}\mathrm{O} \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \omega_\infty = \mathrm{UUU}\cdots \end{array}\right\} \tag{1.1}$$

ω_1 是甲先掷出正面而结束比赛的情形, ω_2 是甲掷出反面, 乙掷出正面而结束比赛的情形. ω_n 是直到第 $n-1$ 次为止二人都掷出反面, 在第 n 次才掷出正面而结束比赛的情形, 因此可由 n 的奇偶来决定是甲或是乙获胜. 最后, ω_∞ 是甲、乙双方总是掷出反面的情形, 这时比赛将无限继续下去. 由于 $\omega_1, \omega_2, \cdots, \omega_\infty$ 是这种比赛经过的各种样本, 所以叫做**样本点**(sample point), 其全体的集合 $\Omega = \{\omega_1, \omega_2, \cdots, \omega_\infty\}$ 就叫做**样本空间**(sample space).

现在来考虑各个样本点的概率. 第 1 次掷出正面或者反面, 这都是同等可能的, 所以 ω_1 的概率是 1/2, 而剩下的 $\{\omega_2, \omega_3, \cdots, \omega_\infty\}$全体的概率就应该是 1/2. 又由同样的理由, 后者的概率 1/2 应平均地分给 ω_2 和剩下

的 $\{\omega_3,\omega_4,\cdots,\omega_\infty\}$, 即各为 1/4. 据此可知, 分布于各个点 $\omega_1,\omega_2,\cdots,\omega_\infty$ 上的概率分别为 1/2, 1/4, 1/8, $\cdots$,0.

以 $P(\omega)$ 表示分布于 ω 上的概率. 于是就有

$$\begin{aligned}&P(\omega_1)=1/2,P(\omega_2)=1/4,\cdots,\\&P(\omega_n)=1/2^n,\cdots,P(\omega_\infty)=0.\end{aligned}\tag{1.2}$$

又若 E 为 Ω 的子集合, 则分布于 E 上的概率是分布于 E 的各点上的概率的总和, 即

$$P(E)=\sum_{\omega\in E}P(\omega).\tag{1.3}$$

这样一来, 就得到了一个集合函数 $P(E)$. 这个集合函数叫做**概率分布**(probability distribution).

其次, 考虑 (ii) 的问题. 直到决定胜负为止的次数可由各个样本点唯一决定. 比如 ω_1 时为 1, ω_2 时为 2, ω_n 时为 n. 这就是定义在样本空间上的函数, 以 $x(\omega)$ 来表示. 在样本空间上如此定义的函数叫做**随机变量**(random variable). 问题 (ii) 就是求随机变量 $x(\omega)$ 的平均值. 平均值有各种各样的定义方法, 而最普遍被采用的是以下的所谓**数学期望**(expectation):

$$E(x)=\sum_{\omega\in\Omega}x(\omega)P(\omega)=\sum_{n=1}^{\infty}nP(\omega_n)=\sum_{n=1}^{\infty}n\cdot\frac{1}{2^n}=2.\tag{1.4}$$

现在转到问题 (i) 上. 所谓甲胜就是上述的 $x(\omega)$ 为奇数的情形, 因此可用条件 "$x(\omega)=$ 奇数" 来表示. 这种可以由有关样本点 ω 的条件来表示的事情就叫做**事件**(event). 事件的概率规定为分布于满足该条件的样本点全体的集合 E [这叫做该事件的**外延**(extension)] 上的概率 $P(E)$. 因此

$$P(\text{甲胜})=P(E)=\sum_{n=0}^{\infty}P(\omega_{2n+1})=\sum_{n=0}^{\infty}\frac{1}{2^{2n+1}}=\frac{2}{3}.\tag{1.5}$$

现在把上例中的本质部分抽出来看一看. 首先, 作为基础的东西就是命名为样本空间的集合 Ω 和它上面的概率分布 P. 此外, 还有一个定义在 Ω 上而称为随机变量的函数 ①, 它的数学期望由 (1.4) 左端的等式来定

① 随机变量的严格的定义将在下面讲到. 此处的定义只适用于样本空间只包含可数个元素的情形. —— 校者注

义. 与 Ω 里的点有关的条件叫做事件, 而该事件的概率就可规定为 P 对其外延 E 的值 $P(E)$. 但在这种一般的场合下, 如何来给出概率分布就需要讨论一下. 由上面的例子可知, 假如 Ω 为一可数集合, 只要给予各个样本点的概率使得其总和为 1, 则就可由 (1.3) 给出它的概率分布. 但是, 在把直线上的点集、平面上的点集以及更一般的集合 (例如做 Brown 运动的粒子的各种各样的运动途径) 分别看做是 Ω 的场合下, 按照上述朴素的方法就不可能给出概率分布. 然而, 若注意到概率分布的想法类似于质量分布, 而后者的数学理论就是测度论, 由此, 就很自然地可以设想, 测度论也将适用于概率分布.

正是基于这种出发点, 我们把概率论构造成为测度论的形式, 因而使得长期以来以常识或者直观为基础, 而且缺乏严格的逻辑推理的概率论真正成为数学理论的一个分支, 并且已在许多的应用上获得了有价值的成果.

§2 概 率 分 布

令 X 为一个集合, $\boldsymbol{B}$ 为由 X 的子集所构成的 Borel 集合体. 对于 $\boldsymbol{B}$ 的元素 (集合) 定义了 Lebesgue 测度 $P(E)$, 而且满足

$$P(X)=1 \tag{2.1}$$

时, 就称 P 为 $X(\boldsymbol{B})$ 上的**概率测度**(probability measuro) 或者**概率分布**, 或简称为**分布**.

首先, 作为最简单的情况, 试考虑 X 为有限集合时的情形. 令 X 的元素为 $x_1, x_2, \cdots, x_n$. 这时通常可将 $\boldsymbol{B}$ 取为由 X 所有的子集组成的集合 2^X. 若只包含一个点的集合的测度 $P(x_1), P(x_2), \cdots, P(x_n)$ 已给定, 则由测度的可加性, 对任意的 $E \subseteq X$ 的概率可以定义为

$$P(E)=\sum_{x\in E} P(x). \tag{2.2}$$

因此, 在这样的场合下, 只需给定点函数 $P(x)$ 就可决定概率分布了. 显然, $P(x)$ 应满足条件

$$P(x) \geqslant 0, \quad \sum P(x)=1. \tag{2.3}$$

特别是当 $P(x)$**不依赖于**x, 而且 $P(x)=1/n$ 时, 就称 P 为**均匀分布**(uniform distribution).

其次, 对于 X 是可数无限集合的情形, 它与有限集合的情况大致相同. 但此时不再存在均匀分布了.

当 X 是实数集合 $\mathbb{R}^1$ 时, 问题就较困难了. 除了个别的特殊情况外, $\boldsymbol{B}$ 不可能取为由 $\mathbb{R}^1$ 的所有子集组成的集合 $2^{\mathbb{R}^1}$. $\boldsymbol{B}$ 的最自然的取法是取为包含所有开集的最小 **Borel 集合体**$\boldsymbol{B}^1$. 通常称 $\boldsymbol{B}^1$ 的元素为 **Borel 集合**. 令 P 为 $\mathbb{R}^1(\boldsymbol{B}^1)$ 上的概率分布. 对 $x\in\mathbb{R}^1$ 的任意邻域 U, 若它的 P 测度 $P(U)$ 为正时, 就称 x 为 P 的支撑点, 而这样的点的全体的集合就叫做 P 的**支集**(support). 特别若 $P(x)>0$, 则显然 x 是 P 的支撑点, 对于这样的 x, 我们另外给它一个名称, 叫做 P 的**不连续点**(discontinuity point). P 的不连续点的全体 D 至多是一可数集合. 当 $P(D)$=1 时, 称 P 为**纯粹不连续**(purely discontinuous); 而 $P(D)=0$ 时, 就称 P 为连续 (continuous). 作为比连续稍微强的条件, 有**绝对连续**(absolutely continuous) 的概念. 若 E 的通常的 Lebesgue 测度 $|E|$=0 时, 必有 $P(E)=0$, 则称 P 为绝对连续. 这时 P 具有密度, 而且可写为

$$P(E)=\int_E f(x)\mathrm{d}x. \tag{2.4}$$

此处 $f(x)$ 应满足的条件是

$$f(x)\geqslant 0, \int_{\mathbb{R}^1} f(x)\mathrm{d}x=1. \tag{2.5}$$

连续但不是绝对连续的概率分布叫做**奇异**(singular) 分布. 纯粹不连续分布、绝对连续分布和奇异分布是在 $\mathbb{R}^1(\boldsymbol{B}^1)$ 上的分布中重要的三种形状, 而任意的分布都可以由这三种形状的分布的凸组合 (convex combination) 来表示.(所谓 a 是 $a_1,a_2,\ldots,a_n$ 的凸组合, 就是指能够写成 $a=\sum c_ia_i(c_i\geqslant 0,\sum c_i=1)$ 的形状.) 这就是 **Lebesgue 分解定理**.

例 1 **δ分布**(delta distribution) $\delta(\cdot;a)$. 这是纯粹不连续分布, 也就是上述的 D 只含有一点 a 的场合. 特别当 $a=0$ 时就叫做**单位分布**(unitary distribution).

例 2 **二项分布**(binomial distribution) $B(\cdot;p,n), 0<p<1, n$ 是自然数. 这是纯粹不连续分布, 也就是由 $D=\{0,1,2,\cdots,n\}$,

$$P(k) = \binom{n}{k} p^k \cdot q^{n-k}, \quad q = 1-p, \quad k = 0,1,2,\cdots,n \tag{2.6}$$

给定的分布. 因为 $P(k)$ 等于 $(p+q)^n$ 展开式中的第 k 项, 所以有二项分布的名称.

例 3　Poisson 分布(Poisson distribution) $P(\cdot;\lambda), \lambda > 0$. 这是纯粹不连续分布, 也就是由 $D = \{0,1,2,\cdots\}$,

$$P(k) = \mathrm{e}^{-\lambda}\frac{\lambda^k}{k!}, \quad k = 0,1,2,\cdots \tag{2.7}$$

给定的分布.

例 4　正态分布(normal distribution) $N(\cdot;a,v), a$ 是实数, $v > 0$. 这是绝对连续分布, 密度由下式给定:

$$f(x) = \frac{1}{\sqrt{2\pi v}}\mathrm{e}^{-\frac{(x-a)^2}{2v}}. \tag{2.8}$$

例 5　Cauchy 分布(Cauchy distribution) $C(\cdot;a,c), a$ 是实数, $c > 0$. 这是绝对连续分布, 密度由下式给定:

$$f(x) = \frac{c}{\pi} \cdot \frac{1}{c^2 + (x-a)^2}. \tag{2.9}$$

当 X 为 m 维空间 $\mathbb{R}^m$ 时, 仍然可以照 $\mathbb{R}^1$ 的场合将结果推广. $\boldsymbol{B}$ 与一维的情况同样, 是包含所有开集的最小 Borel 集合体 $\boldsymbol{B}^m$. $\boldsymbol{B}^m$ 的元素叫做 Borel 集合. 与上述一样, 分布有三种形状, 并且 Lebesgue 分解定理成立.

例 6　δ 分布 $\delta(\cdot;\boldsymbol{a})$ 与一维的情况一样, 不过 $\boldsymbol{a}$ 是 $\mathbb{R}^m$ 的元素而已.

例 7　多项分布(multinomial distribution) $B(\cdot;\boldsymbol{p},n), n$ 是自然数, $\boldsymbol{p} = (p_1,p_2,\cdots,p_n), p_i \geqslant 0, \Sigma P_i = 1$. 这是纯粹不连续分布, 而 D 是格点 $\boldsymbol{k} = (k_1,k_2,\cdots,k_m), \Sigma k_i = n, k_i \geqslant 0$ 的全体.

$$P(\boldsymbol{k}) = \frac{n!}{k_1!k_2!\cdots k_m!}p_1^{k_1}p_2^{k_2}\cdots p_m^{k_m}. \tag{2.10}$$

例 8　正态分布 $N(\cdot;\boldsymbol{a},\boldsymbol{v})$, $\boldsymbol{a}$ 是 $\mathbb{R}^m$ 的元素, $\boldsymbol{V}$ 是对称正定 (狭义) 矩阵. 这是绝对连续的, 其密度为

$$f(\boldsymbol{x}) = (2\pi)^{-\frac{m}{2}}(\det \boldsymbol{V})^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}(\boldsymbol{V}(\boldsymbol{x}-\boldsymbol{a}),(\boldsymbol{x}-\boldsymbol{a}))\right\}. \tag{2.11}$$

此处 $\boldsymbol{V}(\boldsymbol{x}-\boldsymbol{a})$ 是把线性变换 $\boldsymbol{V}$ 作用于向量 $\boldsymbol{x}-\boldsymbol{a}$ 而得的, (,) 表示内积.

如上所述, 从一维的 $\mathbb{R}^1$ 的情形可平行地推广到 m 维的 $\mathbb{R}^m$ 的情形, 但是从 m 维推广到无限维时却有本质上的困难. 比如在无限维空间上, 由于不存在像在 $\mathbb{R}^m$ 的场合下那样普通的 Lebesgue 测度, 所以不可能定义像绝对连续那样的概念. 若令 A 为任意的集合, 那么 $\mathbb{R}^A$ 就是 A 中的元素 α 所对应的实数 ξ 排成的 $\prod\limits_{\alpha}\xi_\alpha$ 全体的集合. 若 A 是有限集合, 则 $\mathbb{R}^A$ 是有限维空间; 但若 A 为无限集合, 则 $\mathbb{R}^A$ 是无限维的. 使 $\prod\limits_{\alpha}\xi_\alpha\to\alpha_0$ 的坐标 ξ_{α_0} 的映象叫做**射影**(projection), 以 p_{α_0} 来表示. $\prod\limits_{\alpha}\xi_a\to\mathbb{R}^A$ 中的点 $(\xi_{\alpha_1},\xi_{\alpha_2},\cdots,\xi_{\alpha_n})$($\alpha_i$ 都不相同) 的映象也叫做射影, 以 $p_{\alpha_1\cdots\alpha_n}$ 来表示. 若令 $E^{(n)}$ 为 n 维的 Borel 集合, 由形状如 $p^{-1}_{\alpha_1\cdots\alpha_n}(E^{(n)})$ 所表示的 $\mathbb{R}^A$ 的子集就叫做 $\mathbb{R}^A$ 的 **Borel 柱集**(cylinder set). 若以 $\boldsymbol{B}^A$ 表示包含所有 Borel 柱集的最小 Borel 集合体, 则 $\boldsymbol{B}^A$ 的元素就叫做 $\mathbb{R}^A$ 的 **Borel 集合**. 今令 P 为 $\mathbb{R}^A(\boldsymbol{B}^A)$ 上的分布. 对于不相同的 $\alpha_1,\alpha_2,\cdots,\alpha_n\in A$, 定义 $\mathbb{R}^n(\boldsymbol{B}^n)$ 上的分布 $P_{\alpha_1\cdots\alpha_n}$, 使得

$$P_{\alpha_1\cdots\alpha_n}(E^{(n)})=P(p^{-1}_{\alpha_1\cdots\alpha_n}(E^{(n)})),E^{(n)}\in\boldsymbol{B}^n. \tag{2.12}$$

这就是分布 **P 的射影**. 考虑所有这类的 $P_{\alpha_1\cdots\alpha_n}$, 且令其全体为 $\mathfrak{P}$, 则 $\mathfrak{P}$ 满足下面所述的 Kolmogoroff **相容性条件**(consistency condition):

(K.1) 若令 $i(1),i(2),\cdots,i(n)$ 为 1,2, $\cdots,n$ 的排列, 则

$$P_{\alpha_{i(1)}\alpha_{i(2)}\cdots\alpha_{i(n)}}(E_{i(1)}\times E_{i(2)}\times\cdots\times E_{i(n)})=P_{\alpha_1\cdots\alpha_n}(E_1\times E_2\times\cdots\times E_n).$$

(K.2) $P_{\alpha_1\alpha_2\cdots\alpha_n}(E^{(n-1)}\times\mathbb{R}^1)=P_{\alpha_1\alpha_2\cdots\alpha_{n-1}}(E^{(n-1)})$.

反之, 对于满足这两个条件的分布系 $\mathfrak{P}$, 存在唯一的 $\mathbb{R}^A(\boldsymbol{B}^A)$ 上的分布 P 满足 (2.12). 这就叫做 Kolmogoroff 定理 ①.

现在, 假设已对 A 的各个元素 α 定义了 $\mathbb{R}^1(\boldsymbol{B}^1)$ 上的分布 P_α. 这时若定义

$$P_{\alpha_1\cdots\alpha_n}=P_{\alpha_1}\times P_{\alpha_2}\times\cdots\times P_{\alpha_n}(\text{直积测度}), \tag{2.13}$$

则 $\mathfrak{P}=\{P_{\alpha_1\cdots\alpha_n}\}$ 满足上述的 (K.1) 和 (K.2). 因此, 按照 Kolmogoroff 定理, 可以由 $\mathfrak{P}$ 定出 $\mathbb{R}^A(B^A)$ 上的分布 P. 称它为 $P_\alpha,\alpha\in A_1$ 的**直积概率**

① 参见 A. H. Колмогоров著的中译本《概率论基本概念》, 丁寿田译, 商务印书馆, 1953. —— 校者注

测度(direct product probability measure), 且以 $\prod\limits_{\alpha\in A} P_\alpha$ 来表示. 显然 P 是由

$$P(p_{\alpha_1}^{-1}(E_1)\cap p_{\alpha_2}^{-1}(E_2)\cap\cdots\cap p_{\alpha_n}^{-1}(E_n))=\prod_{i=1}^{n}P_{\alpha_i}(E_i) \tag{2.14}$$

所表达的 $\mathbb{R}^A(\boldsymbol{B}^A)$ 上的分布. 同理, 当 P_α 为高维 (有限或者无限) 的分布时 (维数可以由 α 而变), 也可以定义其直积概率测度.

§3 测度论观点下的概率论 (2) 逻辑的构成

固定一个集合 Ω, 这集合叫做样本空间. 在 Ω 上取一 Borel 集合体 $\boldsymbol{B}$, 及 $\Omega(\boldsymbol{B})$ 上的概率分布 P, 则 $\Omega, \boldsymbol{B}, P$ 有一个总的名称, 叫做 Ω 上的**概率空间**(probability space), 并记以 $\Omega(\boldsymbol{B}, P)$.

因为 $\Omega(\boldsymbol{B}, P)$ 是一种测度空间, 所以在其上可以建立 Lebesgue 积分理论. $\Omega(\boldsymbol{B}, P)$ 上的可测实函数叫做**随机变量**(random variable). 令 $x(\omega)$ 为随机变量, 则称

$$\Phi(E)=P\{\omega/x(\omega)\in E\}\equiv P(x^{-1}(E)),\quad E\in\boldsymbol{B}^1 \tag{3.1}$$

为 x 的**分布**(distribution). 这就是 $\mathbb{R}^1(\boldsymbol{B}^1)$ 上的概率分布. 称

$$E(x)=\int_\Omega x(\omega)P(\mathrm{d}\omega) \tag{3.2}$$

为 $x(\omega)$ 的**期望**(expectation) 或**均值**(mean). 这可以利用 x 的分布写成形式 ①

$$E(x)=\int_{\mathbb{R}^1}\xi\Phi(\mathrm{d}\xi). \tag{3.3}$$

因为可测函数不一定是可积的, 所以随机变量的均值不一定存在.

将若干随机变量并列起来就可以定义随机向量. 令 A 为有限或者无限集合, 且对于 A 的每个元素 α 都有一随机变量 $x_\alpha(\omega)$ 与之对应. 此时, 若定义

① 参见 Paul R. Halmos 著的中译本《测度论》, 王建华译, 科学出版社, 1958, §39, 定理 3. —— 校者注

$$\boldsymbol{x}(\omega)=\prod_{\alpha\in A}x_\alpha(\omega), \tag{3.4}$$

则 $\boldsymbol{x}(\omega)$ 是定义于 Ω 上而在 $\mathbb{R}^A$ 中取值的函数. 并且在

$$E^{(A)}\in \boldsymbol{B}^A\Rightarrow \boldsymbol{x}^{-1}(E^{(A)})\in \boldsymbol{B} \tag{3.5}$$

的意义下为可测. $\boldsymbol{x}(\omega)$ 就叫做随机向量. 若 A 的势为 m, 则叫做 m 维随机向量. 特别是把二维随机向量 $(x_1(\omega),x_2(\omega))$ 写成 $x_1(\omega)+ix_2(\omega)$, 就得**复随机变量**. 对于随机向量来说, 其分布也可以由 (3.1) 来定义, 再对各个分量进行积分, 就可以定义它的均值向量.

令 φ 为由 $\mathbb{R}^A$ 到 $\mathbb{R}^B$ 内的映象, 并且满足

$$E^{(B)}\in \boldsymbol{B}^B\Rightarrow \varphi^{-1}(E^{(B)})\in \boldsymbol{B}^A. \tag{3.6}$$

这时称 φ 是 **Borel 可测的**(Borel measurable) 或称 **$\boldsymbol{B}$–可测的**(B-measurable), 或者更简单地称**可测**(measurable). 由随机向量的可测映象所产生的象是可测的. 也就是说, 若把上述的可测映象 φ 作用于取值于 $\mathbb{R}^A$ 内的随机向量 $x(\omega)$ 上, 就得到一个在 $\mathbb{R}^B$ 内取值的随机向量 $\varphi(\boldsymbol{x}(\omega))$. 这时 $\varphi(\boldsymbol{x}(\omega))$ 的均值向量可由

$$E[\varphi(\boldsymbol{x}(\omega))]=\int_{\mathbb{R}^A}\varphi(\xi)\Phi\ (\mathrm{d}\xi),\quad \Phi\text{是}\boldsymbol{x}(\omega)\text{的分布} \tag{3.7}$$

给定. 这样的 $\varphi(\boldsymbol{x}(\omega))$ 称为**关于 $\boldsymbol{x}(\omega)$ 是可测的**. 若有若干随机向量 $\boldsymbol{x}_\alpha(\omega),\alpha\in A$, 也可以把它们排列起来定义更高维的随机向量

$$\boldsymbol{x}(\omega)=\prod_{\alpha}\boldsymbol{x}_\alpha(\omega). \tag{3.8}$$

若 $\boldsymbol{x}(\omega)$ 的分布是 $\boldsymbol{x}_\alpha(\omega)(\alpha\in A)$ 的分布的直积分布, 则称 $\boldsymbol{x}_\alpha(\omega)(\alpha\in A)$ 为**独立**(independent). 这也可以由下述条件来表达.

对任意不相同的 $\alpha_1,\alpha_2,\cdots,\alpha_n\in A$, 都有

$$P\{\omega/\boldsymbol{x}_{\alpha_i}(\omega)\in E_i,i=1,2,\cdots,n\}=\prod_{i=1}^{n}P\{\omega/\boldsymbol{x}_{\alpha_i}(\omega)\in E_i\}. \tag{3.9}$$

设 $A=\sum\limits_{\lambda}A_\lambda$(直和), $\lambda\in\Lambda$, 且令 $x_\alpha(\omega)(\alpha\in A)$ 为独立, 则

$$\boldsymbol{y}_\lambda(\omega)=\prod_{\alpha\in A_\lambda}\boldsymbol{x}_\alpha(\omega),\quad \lambda\in\Lambda \tag{3.10}$$