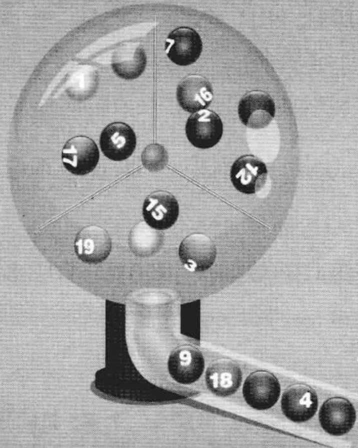
图灵数学·统计学丛书 52

伊藤清概率论

[日] 伊藤清 著
闫理坦 译

人民邮电出版社
POSTS & TELECOM PRESS

TURING 图灵数学·统计学丛书 52



伊藤清概率论

[日] 伊藤清 著

闫理坦 译

人民邮电出版社
北京

图书在版编目(CIP)数据

伊藤清概率论 / (日) 伊藤清著；闫理坦译. —北京：人民邮电出版社, 2011.4
(图灵数学·统计学丛书)
ISBN 978-7-115-24883-1

Ⅰ.①伊… Ⅱ.①伊… ②闫… Ⅲ.①概率论 Ⅳ.①O21

中国版本图书馆 CIP 数据核字(2011)第 030847 号

内 容 提 要

本书是世界级概率论大师伊藤清的名著。篇幅短小，内容丰富，既包括事件、概率、概率空间、均值、特征函数等基本概念，又包括大数定律、Poisson 小数定律、遍历定理以及随机过程的基本内容。

这是一本经典的概率论入门书，适合相关领域的本科生、研究生和教师作为参考书，是每一位概率学者的案头佳作。

图灵数学·统计学丛书
伊藤清概率论

- 著　　　　[日] 伊藤清
  译　　　　闫理坦
  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
  北京艺辉印刷有限公司印刷
- 开本：880×1230 1/32
  印张：5.875
  字数：117 千字　　2011 年 4 月第 1 版
  印数：1-3 000 册　　2011 年 4 月北京第 1 次印刷

著作权合同登记号　图字：01-2008-3602 号

ISBN 978-7-115-24883-1

定价：29.00 元

读者服务热线：(010)51095186　印装质量热线：(010) 67129223
反盗版热线：(010)67171154

新 版 序

这次有机会将我在 60 年前撰写的《概率论基础》一书以现代通俗易懂的表达方式再版，新版同样由岩波书店发行. 作为一个目睹了该领域巨大发展的人，又作为一个在此发展过程中略有小成的人，我不敢言前一版带来了多大的意义.

前一版反映了第二次世界大战后概率论的发展状况. 依照此书学习概率分析，之后又与我一同推动其发展的各位，提出了再版建议，我深感幸福. 特别是池田信行先生和渡边信三先生，二位担任了修正工作，也承蒙岩波书店吉田宇一先生的悉心照顾，在此表示由衷的谢意.

新版书的出版反映了近年来读者对概率论的广泛关注，若能对各位略有帮助，我将备感欣慰.

伊藤清

2004 年春于京都

初 版 序

相比数学的其他分支，概率论的发展极为缓慢，这是由于之前未能找到用数学方式明确表现概率的方法. 尽管如此，集合论、抽象空间论的发展给我们带来了十分鲜明易懂的表达方式，据此有“概率即是勒贝格测度”之说. 没有什么比这句话更能说明概率的数学本质了.

从这个角度讲，以往没有被明确定义的随机变量和事件等概念首次有了明确的表达. 随机变量指的是可测函数，而事件则表示可测集合.

这种定义在近二三十年才慢慢得到广泛认可，其精确的表达方式应归功于 A. Kolmogorov 先生.

本书旨在介绍概率论在新的解释中的基本知识，以及具体问题的解决方法.

执笔本书之际，弥永昌吉先生为我解惑并给予指导，北川敏男先生和畏友河田敬义先生等给予多处指点，吉田耕作先生审校时，他指出了多处本质性的错误，不胜感激. 另外，承蒙岩波书店的各位，尤其是布川角左卫门先生、黄寿永先生以及精兴社各位的关照，在此一并表示深厚的谢意.

伊藤清

1944 年秋于东京

目　录

第 1 章　概率论的基本概念

§1　概率空间的定义

概率空间是用数学观点阐述与研究随机现象的出发点. 在定义概率空间之前, 我们先回顾一些测度论中的有关概念.

抽象空间　集合的别称. 如全体实数的集合构成一个抽象空间, 又如赋予欧氏 (Euclid) 距离的 n 维向量空间也是抽象空间, 本书中将它们分别用 $\mathbb{R}$ 和 $\mathbb{R}^n$ 表示.

完全加法族　如果一个抽象空间的子集类满足下列三个条件, 则称它为该抽象空间的完全加法族(也称为 σ 代数).

1°　抽象空间本身是其中的一个元. 假设此空间为 Ω, 此子集类为 $\mathscr{F}$, 则

$$\Omega \in \mathscr{F}.$$

2°　属于 $\mathscr{F}$ 的可数无限个集合的并也属于 $\mathscr{F}$, 即

$$E_1, E_2, E_3, \cdots \in \mathscr{F} \Longrightarrow \bigcup_{k=1}^{\infty} E_k \in \mathscr{F}.$$

3°　属于 $\mathscr{F}$ 的集合的余集也属于 $\mathscr{F}$, 即若 $E \in \mathscr{F}$, 则

$\Omega - E \in \mathscr{F}$.

利用这三个条件，我们可以推出下列结论.

4° 空集 (今后用 $\varnothing$ 表示) 也属于 $\mathscr{F}$. 事实上，在 3° 中取 $E = \Omega$ 即可.

5° 如果 $E_1, E_2, E_3, \cdots \in \mathscr{F}$, 则 $\bigcap_{k=1}^{\infty} E_k \in \mathscr{F}$.

由恒等式 $\bigcap_{k=1}^{\infty} E_k = \Omega - \bigcup_{k=1}^{\infty} (\Omega - E_k)$ 以及 2° 和 3° 可得这个结论.

6° 如果 $E_1, E_2 \in \mathscr{F}$, 则 $E_1 \cup E_2, E_1 \cap E_2, E_1 - E_2 \in \mathscr{F}$. 这是由于

$$E_1 \cup E_2 = E_1 \cup E_2 \cup \varnothing \cup \varnothing \cup \varnothing \cup \cdots,$$

$$E_1 \cap E_2 = E_1 \cap E_2 \cap \Omega \cap \Omega \cap \Omega \cap \cdots,$$

$$E_1 - E_2 = E_1 \cap (\Omega - E_2).$$

一个抽象空间 Ω 的完全加法族不是唯一的. 其中 $\{\Omega, \varnothing\}$ 是最小的一个，Ω 的所有子集的全体是最大的完全加法族. 如果 Ω 有多个完全加法族，那么它们的交仍然是 Ω 的完全加法族.

对于 Ω 的一个子集族，包含这个子集族的 Ω 的最小完全加法族存在并且唯一，称其为**由这个子集族生成的完全加法族**. 例如，由 $\mathbb{R}$ 的全体区间构成的族所生成的完全加法族为 Borel 集合族. 再如，端点为有理数的全体区间构成的族也生成同一个 Borel 集合族. $\mathbb{R}$ 上的完全加法族有很多种，但是 Borel 集合族是最有用的一个.

将空间 Ω 与其子集构成的一个完全加法族 $\mathscr{F}$ 结合来考虑时，所产生的序偶 $(\Omega, \mathscr{F})$ 称为可测空间. 然而，当 $\Omega = \mathbb{R}$ 时，通

常取 $\mathscr{F}$ 为 $\mathbb{R}$ 的 Borel 集合族 $\mathscr{B}$，并且在多数场合下将 $(\mathbb{R},\mathscr{B})$ 简单地写成 $\mathbb{R}$. 对于空间 $\mathbb{R}^n$，同样将 n 维可测空间 $(\mathbb{R}^n,\mathscr{B}^n)$ 也简单地写成 $\mathbb{R}^n$. 此外，当 Ω 为可分距离空间 D 时，由 D 与 D 的邻域族生成的完全加法族构成的可测空间，也简单地记成 D.

测度 假设 $\mathscr{F}$ 为抽象空间 Ω 的一个完全加法族，m 为 Ω 上的集函数并且满足：

1° m 的定义域为 $\mathscr{F}$;

2° 对于所有的 $E \in \mathscr{F}$，$m(E) \geqslant 0$;

3° m 是完全可加的，即对任意两两不相交的集合序列 $\{E_1, E_2, \cdots\} \subset \mathscr{F}$，均成立

$$m\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} m(E_i).$$

这时，称 m 为 $(\Omega,\mathscr{F})$ 上的**测度**. 当 $\Omega = \mathbb{R}^n$ 而 $\mathscr{F}$ 为 Borel 集合族 $\mathscr{B}^n$ 时，称 m 为 $\mathbb{R}^n$ 上的测度[①]. 序偶 $(\Omega,\mathscr{F},m)$ 称为测度空间.

现在，我们来给出概率空间的定义.

定义 1.1 给定抽象空间 Ω 与其上的一个完全加法族 $\mathscr{F}$. 可测空间 $(\Omega,\mathscr{F})$ 上满足

$$P(\Omega) = 1$$

① 关于测度，请参看高木贞治教授的著作《解析概论》的第 9 章. (本书中文版即将由人民邮电出版社出版.)

的测度 P，称为 $(\Omega, \mathscr{F})$ 上的**概率测度**. 对于 $E \in \mathscr{F}$，称 $P(E)$ 为 E 的**概率**或 E 的 ***P*-测度**.

将 Ω, $\mathscr{F}$, P 一起考虑时，所产生的序偶 $(\Omega, \mathscr{F}, P)$ 称为**概率空间**.

§2 概率空间的实际意义

针对想理解后面出现的定理含义的读者，这里有必要对前一节定义的抽象概率空间在实际随机现象研究中的应用加以说明，仅对推理感兴趣的读者另当别论.

随机现象的研究基本上分为以下三个步骤.

第一步：试验 如投掷一枚均匀的骰子，观察出现的点数；又如从有 6 个黑球 4 个白球的盒子中任取一个球，等等.

第二步：设定标记 为将试验的结果在脑海中清晰地描绘出来，有必要事先确定结果的精度. 以掷骰子的试验为例，要确定的是像德川时代[①]的赌徒们所想的那样只是分偶数与奇数，还是像双六游戏[②]那样关注骰子的点数本身. 为此我们引入标记，也就是说某一个空间 Ω，它可以是 $\mathbb{R}$，$\mathbb{R}^n$ 甚至是更一般的

① 德川时代自公元 1603 年德川家康受任征夷大将军在江户设幕府开始，至 1867 年第 15 代将军庆喜将政治大权奉还朝廷 (即大政奉还) 为止，约 265 年，为继镰仓、室町幕府之后最强盛也是最后的武家政治组织.

—— 译者注

② 一种掷骰子决定赌博的游戏. —— 译者注

抽象空间. 当确定抽象空间后，将试验的结果与 Ω 中的点对应并将其点在脑海中描绘出来，这就是标记. 在前面赌博的场合 $\Omega=\{偶, 奇\}$，在双六游戏的场合 $\Omega=\{1,2,3,4,5,6\}$.

对标记应该注意的是:

1° 与空间 Ω 的任何点都不对应的现象不会发生;

2° 与空间 Ω 的两个点或更多点对应的现象也不会发生;

3° 空间 Ω 中存在与现象不对应的点也无妨. 例如在双六游戏的场合，$\Omega=\mathbb{R}$ 也可以，其中 $1,2,3,4,5,6$ 以外的点与现象不对应.

这里讲的 Ω 相当于概率空间 $(\Omega,\mathscr{F},P)$ 中的 Ω.

第三步: 引入概率 若如上面那样考虑，要确定试验结果发生的可能性的程度就必须引入实际的 Ω 上的概率测度 P. 基本方法有以"**等可能性**"为基础的，有以**频率**为出发点的，等等，但是无论哪种，P 均满足前一节所叙述的有关概率测度的条件. 但由于我们不清楚是否满足完全可加性，因此可以先验证**有限可加性**.

如果 E_1 与 E_2 不相交，则

$$P(E_1\cup E_2)=P(E_1)+P(E_2),$$

这称为**全概率原理**.

据此，对于两两互不相交的集合 $\{E_1,E_2,\cdots,E_n\}$, 我们可以推出

$$P\left(E_1\cup E_2\cup\cdots\cup E_n\right)=P(E_1)+P(E_2)+\cdots+P(E_n).$$

这个等式称为**有限可加性**.

以此类推，仅依靠形式的推理是不能导出完全可加性的. 将概率的完全可加性作为基础来假设，是数学上的理想化模式. 你渐渐地便能理解这种理想化不是与实际相悖的，反而是与其一致的.

综合以上三个步骤的分析便获得概率空间 $(\Omega, \mathscr{F}, P)$.

§3 概率测度的简单性质

假设 $(\Omega, \mathscr{F}, P)$ 为概率空间. 对于 Ω 的任意子集 E，定义

$$\bar{P}(E) = \inf\{P(E');\ E' \supset E,\ E' \in \mathscr{F}\},$$

$$\bar{P}(E) = \sup\{P(E');\ E' \subset E,\ E' \in \mathscr{F}\}.$$

如果 $\bar{P}(E) = \underline{P}(E)$，则称集合 E 为 P-**可测的**. P-可测集合的全体 $\mathscr{F}'$ 构成一个完全加法族，并且$\mathscr{F}' \supset \mathscr{F}$. 对于$P$-可测集合 E，将 $\bar{P}(E)$ 或 $\underline{P}(E)$ 再记成 $P(E)$，这也是 $(\Omega, \mathscr{F}')$ 上的概率测度，并且它是 $P(E)$ 的扩张. 由于类似的扩张总是可能的，所以我们约定，今后谈到 $(\Omega, \mathscr{F})$ 上的概率测度 P，总是考虑成如上介绍的扩张.

定义 3.1 点 ω 构成的单点集合，记成 $\{\omega\}$，如果其概率 $P(\{\omega\})$ 是正的，则称 ω 为 P 的不连续点.

推论 不连续点的全体至多是可数的.

定义 3.2 对于不连续点的全体 C，如果 $P(C)=1$，则称 P 为**纯不连续的概率测度**.

例 假设 $\Omega=\mathbb{R}$，$\mathscr{F}$ 是 $\mathbb{R}$ 的 Borel 集合族. 定义 $\mathscr{F}$ 上的集函数 P 如下：

$$p_k=\mathrm{e}^{-\eta}\frac{\eta^k}{k!},\quad k=0,1,2,\cdots;$$

$$P(E)=\sum_{k\in E}p_k,\quad E\in\mathscr{F}.$$

则 P 为纯不连续的概率测度. 这时，Ω 的任意子集均是 P-可测的. 关联的 P 称为**Poisson 分布**.

定义 3.3 没有不连续点的概率测度 P 称为**连续的**.

例 假设 Ω 是实数轴 $\mathbb{R}$ 上的区间 $[a,b]$，$\mathscr{F}$ 是属于该区间的 Borel 集合的族. 对于 $E\in\mathscr{F}$，定义集函数 P 如下：

$$P(E)=\frac{m(E)}{b-a},$$

其中 m 是 Lebesgue 测度，则 P 为 $(\Omega,\mathscr{F})$ 上的连续概率测度，关联的 P 称为 $[a,b]$ 上**均匀分布**.

$(\Omega,\mathscr{F})$ 在完全一般的场合下也可以得出上面的定义. 我们也常遇到 $(\Omega,\mathscr{F})$ 较为特殊且其上事先给定了测度 m 的情况. 例如 Ω 是 n 维欧氏空间 $\mathbb{R}^n$, $\mathscr{F}$ 为其上的 Borel 集合族，m 为其上一个给定的 Lebesgue 测度. 在这种场合下，将 $(\Omega,\mathscr{F})$ 上的概率测度 P 与 m 结合起来考虑时，得到下面的定义.

定义 3.4 如果 $m(N)=0$ 蕴含 $P(N)=0$，则称 P**关于 m 是绝对连续的**，或称**绝对连续 (m)**，并在不至于产生误解的条件下简单地称为**绝对连续**.

根据测度论中众所周知的定理可以获得如下定理.

定理 3.1 (Radon-Nikodym 定理) 如果 Ω 可以表示成可数个 m- 测度有限的集合的并，则关于 m 绝对连续的概率测度 P 可以表示成关于测度 m 的积分. 即对任意集合 $E\in\mathscr{F}$，均有

(1) $P(E)=\displaystyle\int_E f(\omega)m(\mathrm{d}\omega)$,

其中 $f(\omega)$ 为定义在 Ω 上关于 m 可积的可测函数，并且满足

(2) $\displaystyle\int_\Omega f(\omega)m(\mathrm{d}\omega)=1$,

(3) 关于测度 m 几乎处处有 $f(\omega)\geqslant 0$.

反之，由满足上面条件 (2) 与 (3) 的关于 m 可积的函数，按照 (1) 式定义的集函数 $P(E)$ 构成一个关于 m 绝对连续的概率测度 P.

此外，满足 (1) 的可测函数 $f(\omega)$ 在 Ω 上关于测度 m 几乎处处唯一，即如果关于 m 可积的可测函数 $f_1(\omega)$ 和 $f_2(\omega)$ 均满足 (1)，则

$$m(E\{\omega\,;\,f_1(\omega)\neq f_2(\omega)\})=0.$$

定义 3.5 上面定理中的 $f(\omega)$ 称为绝对连续的概率测度 P(关于 m) 的**概率密度**.

例 1 (均匀分布) 区间 $[a,b]$上的均匀分布是绝对连续的，其概率密度 $f(x)=\dfrac{1}{b-a}$，$x\in[a,b]$.

例 2 (Gauss 分布) 当 $\Omega=\mathbb{R}$ 时，容易验证 Ω 上的函数

$$f(\omega)=\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{1}{2\sigma^2}(\omega-\mu)^2}\qquad(\sigma>0,\mu\in\mathbb{R})$$

满足上面的条件 (2) 和 (3). 称以 $f(\omega)$ 为概率密度的概率测度为 $\mathbb{R}$ 上的 Gauss 分布.

例 3($\mathbb{R}^n$ 上的 Gauss 分布) 定义 $\mathbb{R}^n$ 上的实值函数 $f(\xi_1,\xi_2,\cdots,\xi_n)$ 如下：

$$f(\xi_1,\xi_2,\cdots,\xi_n)=\frac{\sqrt{\Delta}}{\pi^{n/2}}\exp\Big\{-\sum_{i,j}a_{ij}(\xi_i-m_i)(\xi_j-m_j)\Big\},$$

其中 a_{ij},m_i 以及 $j=1,2,\cdots,n$ 均为实数，Δ 为行列式 $|a_{ij}|$ 并且 $\sum\limits_{i,j}a_{ij}(\xi_i-m_i)(\xi_j-m_j)$ 是对称正定二次型.

则 $f(\xi_1,\xi_2,\cdots,\xi_n)>0$ 并且在 $\mathbb{R}^n$ 上可测. 下面证明

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}f(\xi_1,\xi_2,\cdots,\xi_n)\mathrm{d}\xi_1\mathrm{d}\xi_2\cdots\mathrm{d}\xi_n=1.$$

假设矩阵 $(a_{ij})_{n\times n}$ 的特征根为 $\lambda_1,\lambda_2,\cdots,\lambda_n$，则 $\lambda_i>0$，$i=1,2,\cdots,n$. 向量 $(\xi_1,\xi_2,\cdots,\xi_n)$ 到向量 $(\xi_1',\xi_2',\cdots,\xi_n')$ 的适当正交变换蕴含

$$\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}(\xi_i-m_i)(\xi_j-m_j)=\sum_{i=1}^{n}\lambda_i(\xi_i')^2.$$

另外，正交变换的 Jacobi 行列式的绝对值等于 1，即

$$\mathrm{d}\xi_1\mathrm{d}\xi_2\cdots\mathrm{d}\xi_n=\mathrm{d}\xi_1'\mathrm{d}\xi_2'\cdots\mathrm{d}\xi_n'.$$

因此，

$$\begin{aligned}&\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}f(\xi_1,\xi_2,\cdots,\xi_n)\mathrm{d}\xi_1\mathrm{d}\xi_2\cdots\mathrm{d}\xi_n\\
=&\frac{\sqrt{\Delta}}{\pi^{n/2}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\exp\left\{-\sum_{i=1}^{n}\lambda_i(\xi_i')^2\right\}\mathrm{d}\xi_1'\mathrm{d}\xi_2'\cdots\mathrm{d}\xi_n'\\
=&\frac{\sqrt{\Delta}}{\pi^{n/2}}\prod_{i=1}^{n}\int_{-\infty}^{\infty}\mathrm{e}^{-\lambda_i(\xi_i')^2}\mathrm{d}\xi_i'\\
=&\frac{\sqrt{\Delta}}{\pi^{n/2}}\prod_{i=1}^{n}\frac{\sqrt{\pi}}{\sqrt{\lambda_i}}=\frac{\sqrt{\Delta}}{\sqrt{\lambda_1\lambda_2\cdots\lambda_n}}.\end{aligned}$$

注意 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 为矩阵 $(a_{ij})_{n\times n}$ 的特征根的事实蕴含

$$\Delta=\lambda_1\lambda_2\cdots\lambda_n.$$

从而上面最后一个代数式等于 1.

由此可见，$f(\xi_1,\xi_2,\cdots,\xi_n)$ 满足概率密度的条件. 由此定义的 $\mathbb{R}^n$ 上的概率测度称为 n 维正态分布或 $\mathbb{R}^n$ 上的**Gauss 分布**.

作为连续概率测度, 与绝对连续概率测度相对应的是奇异概率测度.

定义 3.6　假设 m 是连续测度, 如果存在 N 使得 $m(N)=0$，并且 $P(N)=1$，则称 P 为关于 m 的**奇异概率测度**.

例　将区间 $I=[0,1]$ 三等分， 取中间的部分为 I_1. 如果 $x\in I_1$，定义 $g(x)=\dfrac{1}{2}$.