纯粹数学与应用数学专著·典藏版

第27号

统计渐近论基础

勒 康 罗昭容 著

科学出版社

纯粹数学与应用数学专著 第27号

统计渐近论基础

勒 康 罗昭容 著

科 学 出 版 社

1 9 9 3

（京）新登字 092 号

内 容 简 介

本书对在大样本情况下遇到的统计问题作系统的介绍．书中用简单但严格的数理方法讲解渐近论中的主要定理并提供一套估计方法以满足应用的需要．本书除总结近 50 年来渐近论发展的成果外，并在每章末附该章内容的发展简史，并叙述各定理的由来．

本书可供大学有关专业的大学生、研究生阅读．也可供研究人员参考．

图书在版编目(CIP)数据

纯粹数学与应用数学专著丛书：典藏版/杨乐主编. —北京：科学出版社，2018.1

ISBN 978-7-03-055754-4

Ⅰ. ①纯… Ⅱ. ①杨… Ⅲ. ①数学 Ⅳ. ①O1

中国版本图书馆 CIP 数据核字(2017) 第 298639 号

责任编辑：毕 颖／责任校对：李静科
责任印制：张 伟／封面设计：陈 敬

科学出版社 出版
北京东黄城根北街 16 号
邮政编码：100717
http://www.sciencep.com

北京东华虎彩印刷有限公司印刷
科学出版社发行 各地新华书店经销
*
1993 年 12 月第 一 版 开本：720 × 1000 1/16
2018 年 1 月 印 刷 印张：13 1/2
字数：172 000

《纯粹数学与应用数学专著》丛书

主　编　杨　乐

副主编　（以姓氏笔画为序）

王　元　王梓坤　石钟慈　严士健

张恭庆　胡和生　潘承洞

序

1968 年夏季，本书作者之一(勒康)应蒙特律大学 (University of Montréal) 邀请作为期 1 个月的学术演讲，讲题为统计决策渐近理论。承 Catherine Doleáns, Jean Haezendonck 及 Roch Roy 三位博士记录整理，此法文讲义随后被编入蒙特律大学出版社的高等数学讲座丛书发行. 20 年后，我们觉得这本讲义也许有翻译成中文的价值，后经 Shanti Gupta 教授之建议，我们决定将原来的法文版更新重写，分别出英文和中文版，这就是本书的来历.

中、英文版仍然保持原来讲义的大纲，但删除了过时的材料，并且增添了一些近 20 年来获得的新结果. 原版注重于 LAN 条件的讨论，而新版增添了 Jeganathan 获得的结果和对 LAMN 条件的讨论，并包括 Hájek-Le Cam 渐近极小极大定理和卷积定理，以及这些定理的含义. 本书并未对渐近理论作全面的讨论，因此书中有些定理只作叙述而不给证明.

本书的目的在于用简单的方式来讲解几个基本的概念及工具. 读者可从参考文献中查到这方面的一般资料. 我们希望此书能为读者提供一套考虑统计渐近问题的方法. 这套方法比用传统的极大似然方法来处理渐近问题要完整连贯些.

谨此感谢蒙特律大学出版社准许引用部分原版材料，并感谢 Springer-Verlag 发行此书。

我们同时感谢方开泰教授费心与科学出版社洽商，出版此书.

英文版在 Berkeley 由 Chris Bush 女士神速打字完成，谨此感谢 Springer-Verlag 的 Ruediger Gebauer 先生及 Susan Gordon 女士校对书稿.

我们非常感谢许建伦博士在准备中文版时所提供的一切协助,赵林城教授在阅读中文稿后所提供的宝贵意见,以及蒋继明先生校对中文稿。

勒 康(Lucien Le Cam)

羅昭容(Grace Lo Yang)

1989 年 10 月

目　　录

第一章 引 言

本书的目的是叙述一些在考虑统计渐近问题时非常有用的概念及工具，其主要思想是用一些比较容易处理或比较熟悉的测度集 $\mathscr{F}=\{Q_\theta:\theta\in\Theta\}$ 来逼近我们原有的概率测度集 $\mathscr{E}=\{P_\theta:\theta\in\Theta\}$.

举例来说，假定我们观测到大量的独立同分布的随机变量 $X_1, X_2,\cdots,X_n$. 假设它们在实轴上有哥西分布，其密度为

$$f(x,\theta)=\frac{1}{\pi}\frac{1}{1+(x-\theta)^2},\quad -\infty<x<\infty.$$

记 $P_{\theta,n}$ 为 $X_1,\cdots,X_n$ 的联合分布，令 Z_n 为另一随机变量，它在实轴上服从正态分布 $G_{\theta,n}$，其均值为 θ，方差为 $\frac{2}{n}$.

本书要阐述的理论是说，当 n 充分大时，$\mathscr{E}_n=\{P_{\theta,n}:\theta\in\mathscr{R}\}$ 和 $\mathscr{F}_n=\{G_{\theta,n}:\theta\in\mathscr{R}\}$ 这两个概率集对处理大部分统计问题而言，区别甚小.

另外一个例子是，设 $Y_1,Y_2,\cdots,Y_n$ 为 n 个独立同分布的随机变量，每个变量的密度为 $[1-|x-\theta|]^+$，其联合分布记为 $Q_{\theta,n}$. 其次令 $H_{\theta,n}$ 为正态分布，其均值为 θ，方差为 $\frac{1}{n\log n}$，则当 n 值充分大时，$\{Q_{\theta,n}:\theta\in\mathscr{R}\}$ 和 $\{H_{\theta,n}:\theta\in\mathscr{R}\}$ 这两个概率集会相当靠近.

第二章将介绍几个距离的定义，其目的是用严谨的数学定义来解释什么是两个概率集“相当靠近”. 至于利用距离的思想，可追溯到 Wald [1943] 的文章. 我们要用的这些距离定义亦与 Blackwell [1951] 等人所讨论的“实验比较”有关. 所谓实验，本书将依照 Blackwell 的定义，将它定义为 $\mathscr{X}$ 集上一个 σ- 域

$\mathscr{A}$ 上的任何一个概率集 $\mathscr{E}=\{P_\theta:\theta\in\Theta\}$，指标集$\Theta$通常被称为“参数空间”．有一种方便的看法是把每个$\theta$看成是一种理论，它对实验者要进行的实验，提供一个随机模型 P_θ.

有一点要注意的是在比较哥西实验 $\{P_{\theta,n}:\theta\in\mathscr{R}\}$ 及高斯实验 $\{G_{\theta,n}:\theta\in\mathscr{R}\}$ 的例子中，这两个实验有相同的参数空间Θ（此处 $\Theta=\mathscr{R}$），但有完全不同的样本空间 $\mathscr{X}$．哥西实验的观测值是在n维空间 $\mathscr{X}=\mathscr{R}^n$ 里，而高斯实验的观测值则在一维空间里．在第二章中我们所定义的实验距离，适用于任何两个有相同参数空间的实验 $\mathscr{E}=\{P_\theta:\theta\in\Theta\}$ 和 $\mathscr{F}=\{Q_\theta:\theta\in\Theta\}$．在同一章内，我们用 Blackwell 典型表达法求出一个实验的标准表达式，这个结果是在实验指标集Θ为有限的条件下求出的．运用这个标准表达式，我们可以证明当Θ为一固定有限集时，实验的收敛（在我们的距离定义下）与似然比分布的收敛等价．

第三章将讨论一些在研究似然比收敛时遇到的技巧问题．这些技巧问题通常可用“同居条件”予以彻底简化．同时，我们还将介绍 Hellinger 变换及 Hellinger 距离，它们对研究独立变量的实验特别有用．

极限定理是第四章的主题，所讨论的是一些在独立变量实验情况下获得的极限定理．承蒙 Hájek 及 Šidak 二位的恭维，本章包括了他们所称的“Le Cam 三引理”．

第五章阐明 LAN 条件．LAN 是 local asymptotic normality 之缩写．它的真正意义是用高斯移位实验（Gaussian shift experiment）对原来的实验作局部渐近逼近，此高斯实验的参数集为一个k维空间的线性指标集．本章将详述在一个参数值θ周围，因实验满足 LAN 条件而产生的一些结果．除此之外，本章还提供了一个构造估计量的方法．此法步骤如下：第一，任取一个“好”的初步估计量 θ_n^*，并选取一个适当的向量集 $\{u_{n,i}:i=0,1,\cdots,k\}$，其中 $u_{n,0}=0$，$\{u_{n,i}:\ i=1,\cdots,k\}$ 是参数空间 $\mathscr{R}^k$ 中的一组基；第二，算出对数似然比在所有 $\theta_n^*+u_{n,i}+u_{n,j}$ 点之值，这里 $i,j=0,1,\cdots,k$；第三，求出经过这些对数似然比

值的二次式拟合；第四，在 $\mathscr{R}^k$ 空间里算出使这二次式达到极大之点 T_n，T_n 即是我们用来估计 θ 的估计量.

在 LAN 条件下，依上述方法构造的估计量有渐近极小极大性，并有渐近充分性. 此估计量也满足 Hájek 卷积定理，我们将用 van der Vaart 的方法来证明此卷积定理. 本章最后一节是讨论 LAMN 条件，LAMN 是 locally asymptotically mixed normal 的缩写. 这里我们主要是引用 Jeganathan 的文章，其他的例子及资料请参看 Basawa 和 Prakasa Rao [1980]，Basawa 和 Scott [1983]，Prakasa Rao [1987]，Greenwood 和 Shiryayev [1985] 等书.

第六章讨论独立变量的情况. 我们叙述 LAN 条件的形式，特别是在"标准独立同分布"情况下的形式. 统计界一般所熟悉的是建立在 Cramér 条件下的极大似然理论，根据第五章结果所建立的理论虽与极大似然理论有几分相似，但是它与第二章所介绍的概念比较连贯一致些. 该理论所用的条件比 Cramér 用的要弱些. 第六章将详细介绍该理论的一个充分条件，即均方可微性. 最后，我们举例说明，如何应用这套理论到其他各种情况.

第七章讨论 Bayes 程序 (Bayes-procedure) 和 Bernstein-von Mises 定理. 本章将叙述此定理的一种形式，证明方法将强调其中关键步骤.

本书每章最后一节为附录，介绍该章内容的发展简史. 书末附有参考文献及中英文名词对照表. 因受篇幅所限，本书无法提供一套完整的参考资料，有兴趣的读者可从其他书籍提供的文献中找到补充资料. 与本书有关的书籍可参阅 Basawa 和 Prakasa Rao [1980]，Basawa 和 Scott [1983]，Greenwood 和 Shiryayev [1985]，Ibragimov 和 Has'minskii [1981]，Le Cam [1986]，Pfanzagl 和 Wefelmeyer [1982]，Prakasa Rao [1987]，Serfling [1980]，Strasser [1985] 等书.

第二章　实验、亏值、距离

§2.1　风险函数之比较

令 $\mathscr{A}$ 为 $\mathscr{X}$ 集内的一个 σ- 域，依照 Blackwell 的定义，我们称 $\mathscr{A}$ 上的任何一概率测度集 $\mathscr{E}=\{P_\theta:\theta\in\Theta\}$ 为一个实验，称指标集 Θ 为参数空间. 要描述 Wald 所谓的统计决策问题，我们还需要引入一个决策集 Z 和一个损失函数 W，其中 W 是定义在 $\Theta\times Z$ 上取值于 $(-\infty,+\infty]$ 的函数.

假设一个统计工作者在样本空间 $\mathscr{X}$ 内观察到一个值 x，他要在未知 θ 的情况下，从决策集 Z 里选取一个决策 z. 他的取法是先选定一个在 Z 上的概率测度 ρ_x，然后依照概率 ρ_x 随机抽样得到一值. 如果所抽到的是 z，则他受到的损失为 $W_\theta(z)$. 因此，当观察值为 x 时，他的平均损失为 $\int W_\theta(z)\rho_x(dz)$. 若 x 是按照概率分布 P_θ 抽出，则他的总平均损失为双重积分 $\int\left[\int W_\theta(z)\rho_x(dz)\right]P_\theta(dx)$.

我们称函数 $\rho:x\rightsquigarrow\rho_x$ 为随机化决策程序或决策函数. 称双重积分 $\int\left[\int W_\theta(z)\rho_x(dz)\right]P_\theta(dx)$ 为决策 ρ 的风险，这是指以 W 为损失函数而 θ 为真值情况下的风险. 我们记它为 $R(\theta,\rho)$，有时为强调这些函数的关系，我们将 $R(\theta,\rho)$ 记为 $W_\theta\rho P_\theta$.

为使上述定义有意义，显然需要附加条件，即这些积分必须存在. 为使积分存在，我们假定：对每一个 θ 值，$\inf\limits_z W_\theta(z)>-\infty$，令 $\mathscr{B}$ 为 Z 集上使 W_θ 为 z 的可测函数的 σ- 域，ρ_x 为 $\mathscr{B}$ 上的概率测度，对每一个 $B\in\mathscr{B}$，函数 $\rho:x\rightsquigarrow\rho_x(B)$ 是 x 对 $\mathscr{A}$ 的可测函数. 最后这个“可测”假定不仅是为了数学处理的方便，

而且是为了反映实际情况．可以想象，在实际工作中，统计工作者或实验者之所以假设 σ- 域 $\mathscr{A}$，是因为他对 $\mathscr{A}$ 中的事件 A 感兴趣并且能判断 A 是否发生．我们要求 $x \rightsquigarrow \rho_x(B)$ 为 $\mathscr{A}$ 的可测函数，原则上就是要求决策程序 ρ 必须为能从已知 $\mathscr{A}$ 的信息而"执行"的决策程序．

Wald 的统计决策函数论就是基于上述结构，他的观点是运用风险函数来评判和比较各种决策程序之优劣．

现假设有一实验者，在斟酌两个不同的实验方法，记这两实验为 $\mathscr{E}=\{P_\theta: \theta\in\Theta\}$ 和 $\mathscr{F}=\{Q_\theta: \theta\in\Theta\}$，这里 P_θ 是在测度空间 $(\mathscr{X}, \mathscr{A}_1)$ 上的概率测度，Q_θ 是测度空间 $(\mathscr{Y}, \mathscr{A}_2)$ 上的概率测度．要注意此处采用的是同一个参数集 Θ．原因是实验者根据某种已知理论"θ"而构造随机模型 $P_\theta\in\mathscr{E}$，如果他要采取另一种方法做实验 $\mathscr{F}$，原则上他对基本理论 θ 的认识是不变的，因此这两个实验的随机模型指标集 Θ 应为同一个．举例来说，假设他要设计一个物理实验来估计碳 14(C^{14}) 的半衰期．他首先假定碳 14 原子的寿命具有指数分布，其密度为 $\theta e^{-\theta x}$，$x>0$，$\theta\in(0,\infty)$．然后来观察 n 个碳 14 原子．下一步进行的方法则有好几种，一种方法是观测在某一给定时间 (t) 内这 n 个碳原子有多少 (X) 发生分裂．按照上述说明，理论"θ"则为 X 提供一个概率分布 P_θ，另一种实验方法是事先给定某个 $m(m\leqslant n)$，然后来测量这 m 个原子分裂所需的总时间 Y，一如从前，理论 θ 为变量 Y 提供一个概率分布 Q_θ．

对于这两个实验 $\mathscr{E}=\{P_\theta:\theta\in\Theta\}$ 和 $\mathscr{F}=\{Q_\theta:\theta\in\Theta\}$，我们能作比较吗？它们的统计性质差异何在？这要看样本 n 的大小以及选出的 t, m 值．本书的目的在于讨论一般理论，它可以用来说明以下事实：实验 $\mathscr{E}$ 与 $\mathscr{F}$ 也许不能直接比较，但是，如果所取的 t 值使得 EX 比 m 值大许多，则对于在 $\mathscr{F}$ 上能获得的统计结果，我们在 $\mathscr{E}$ 上亦能获得同样好或差不多好的结果．如果 m 与 EX 皆很大，并且 $\dfrac{EX}{m}$ 趋于 1，则 $\mathscr{E}$ 与 $\mathscr{F}$ 之差异变得

很小.

读者可以取 $n=10^{13}$, $t=2$ 小时及 $m=10^6$ 的特殊例子来体会这个问题.

至于测定两实验间的距离,我们使用的方法是考察分别在 $\mathscr{E}$ 和 $\mathscr{F}$ 上能达到的两个风险函数之差.

为此我们来定义一种函数集 $\mathbb{R}(\mathscr{E}, W)$. 函数集 $\mathbb{R}(\mathscr{E}, W)$ 包含所有具有下列性质的 θ 函数 $r(\theta)$: 在 $\mathscr{E}$ 实验上存在一个决策程序 ρ, 它的风险函数 $W_\theta\rho P_\theta$ 对所有的 θ 皆满足不等式 $W_\theta\rho P_\theta \leqslant r(\theta)$.

由于数学技巧的缘故, 我们不采用 $\mathbb{R}(\mathscr{E},W)$ 而用它的逐点闭包, 即在函数逐点收敛意义下的闭包 $\bar{\mathbb{R}}(\mathscr{E}, W)$. 如果我们假设:

(1) 所有的 P_θ 被同一个 σ- 有限测度 μ 控制.

(2) 只考虑使 Z 为紧集的决策问题, 并且, 损失函数 $z \rightsquigarrow W_\theta(z)$ 在 Z 上为下半连续. 则 $\mathbb{R}(\mathscr{E},W)$ 自然是一个闭集.至于实验 $\mathscr{E}$ 及决策空间 (Z,W) 的更一般定义, 读者可看 Le Cam [1986].

在这些假设下,我们给出下面两个定义:

定义 1 对任意一个损失函数 W, $0\leqslant W_\theta(z)\leqslant 1$,对每一个 $r_2\in\mathbb{R}(\mathscr{F},W)$和一个 $\varepsilon\in[0,1]$,考虑交集 $\bar{\mathbb{R}}(\mathscr{E},W)\cap\{f:f\leqslant r_2+\varepsilon\}$. $\mathscr{E}$ 对 $\mathscr{F}$ 的亏值 (deficiency) $\delta(\mathscr{E},\mathscr{F})$, 是对全部 W和全部 r_2 使此交集为非空集的最小 ε 值.

定义 1 中所考虑的 W 是包括所有介于 0 到 1 之间的损失函数 W. 这当然包括了许多统计问题,即使我们缩小范围只考虑满足上述紧性条件的 Z,甚至只考虑 Z 为有限的情况,它所包括的统计问题仍然很多(按照 Le Cam [1986] 书中给的一般定义, 这种限制并不影响 $\delta(\mathscr{E}, \mathscr{F})$ 值). Lehmann [1988] 认为这个亏值定义所包括的统计问题实在太多,他的观点当然不无道理,可是我们无从知道每一个实验者所想要的 W,我们采纳这个定义是为顾全各种不同W的可能性.

下面是实验距离的定义.

定义 2 两个实验 $\mathscr{E}$ 和 $\mathscr{F}$ 的距离 $\Delta(\mathscr{E},\mathscr{F})$ 为两亏值 $\delta(\mathscr{E},\mathscr{F})$ 和 $\delta(\mathscr{F},\mathscr{E})$ 中较大的一个.

总结来说,这个定义的意义是: 如果只考虑满足 $0\leqslant W\leqslant 1$ 的损失函数 W, 则任何能从两实验之一获得的结果(风险),都能在 $\Delta(\mathscr{E},\mathscr{F})$ 距离范围内从另外一个实验得到.

上述距离定义与另外一个距离定义有密切关系. 这种关系在 Le Cam [1986] 书中总结为第 123 页上的一个定理, 其意义如下: 假定我们做一个实验 $\mathscr{E}$,在 $\mathscr{E}$ 的样本空间 $\mathscr{X}$ 里获得一个观察值 x, 则我们可应用随机化方法,按照 $\mathscr{F}$ 的测度空间$(\mathscr{Y},\mathscr{A}_2)$ 上的一个概率分布 $K_x(dy)$ 在样本空间 $\mathscr{Y}$ 里构造一个观察值 Y, 而使 Y 的分布接近于原来 $\mathscr{F}$ 里设的分布 Q_θ. 所谓接近是根据如下的距离定义: 令 KP_θ 为一测度,其定义为

$$(KP_\theta)(A)=\int K_x(A)P_\theta(dx),\quad A\in\mathscr{A}_2.$$

定义测度 Q_θ 和 KP_θ 之间的距离为

$$\|Q_\theta-KP_\theta\|=\sup_f\left|\int fdQ_\theta-\int fdKP_\theta\right|,$$

这里 f 为任意满足 $|f|\leqslant 1$ 的 $\mathscr{A}_2$ 可测函数, 则在相当弱的条件之下,可以证明亏值 $\delta(\mathscr{E},\mathscr{F})$ 就是

$$\inf_K\sup_\theta\frac{1}{2}\|Q_\theta-KP_\theta\|.$$

这里的 inf 是对所有的随机化概率 K 和所有 K 的极限来取. (证明见 Le Cam [1986])

我们将称此处所用的范数

$$\|\mu\|=\sup_f\left\{\int fd\mu:|f|\leqslant 1\right\}$$

为 L_1- 范数, 它的另一个名称为总变差范数 (total variation norm). L_1-范数具有特殊的统计意义: $1-\dfrac{1}{2}\|P-Q\|$ 是检验 P 对 Q 的两类错误概率和之极小值. 因此, $\delta(\mathscr{E},\mathscr{F})\leqslant\varepsilon$ 的意

义就是：只要损失函数W取值于[0,1]中，则对 $\mathscr{F}$ 上的任何风险函数，我们都能在 ε 范围内，在 $\mathscr{E}$ 上找到一个同样好的风险函数，并且，在实验 $\mathscr{E}$ 完成后，运用随机化 K，在 2ε 误差范围内能重新构造出 $\mathscr{F}$ 中的分布 Q_θ.

在上面，我们称Δ为“距离”，其实它只是一个伪距离，因为两个很不同的实验 $\mathscr{E}$ 和 $\mathscr{F}$ 之间的距离 $\Delta(\mathscr{E},\mathscr{F})$ 可能是零. 因此，若 $\Delta(\mathscr{E},\mathscr{F})=0$，我们称 $\mathscr{E}$ 和 $\mathscr{F}$ 为等价或有相同类型. 在实验类型的空间上，Δ变成了距离.

我们相信读者不难体会到Δ具有确定的统计意义. 虽然如 Lehmann 所说在Δ定义下的决策问题集是稍许大了些，但是每一个从事实验设计的人皆知道，要选出一个使每一个人都赞成的决策问题子集不是一件容易的事.

在下节，我们将讨论Δ的另一面，证明距离Δ与许多统计工作者爱用的似然比分布有密切关系.

§2.2 似然比，Blackwell 表达式

本节假设参数空间Θ为含有 k 个元素的有限集.

现考虑实验 $\mathscr{E}=\{P_\theta: \theta\in\Theta\}$，其中 P_θ 为测度空间 $(\mathscr{X},\mathscr{A})$ 上的概率. 令 $S=\sum\limits_{\theta\in\Theta}P_\theta$. 因为 S 控制每一个 P_θ，因此存在拉冬-尼可丁密度 $f_\theta=\dfrac{dP_\theta}{dS}$. 我们可以算出 f_θ 在 $x\in\mathscr{X}$ 的值. 这些密度值可用一个 k 维向量来表示:

$$v(x)=\{f_\theta(x):\theta\in\Theta\},\quad v(x)\in\mathscr{R}^k,$$

不失一般性，可以假设 $f_\theta(x)\geqslant 0$，$\sum\limits_\theta f_\theta(x)=1$. 令

$$U(\Theta)=\left\{u=[u_\theta:\theta\in\Theta]:u_\theta\geqslant 0,\sum_\theta u_\theta=1\right\}$$

为 k 维单纯形，则 $v(x)$ 是 $U(\Theta)$ 中的一点. S 经变换 $x\rightsquigarrow v(x)$

所造成的象是单纯形 $U(\Theta)$ 上的一个测度，记它为 m，而个别 P_θ 的象则是 $U(\Theta)$ 上另外一个概率测度 P'_θ，由此很容易证明 $\frac{dP'_\theta}{dm}$ 就是 u_θ。于是我们得到了另外一个实验 $\mathscr{E}'=\{P'_\theta:\theta\in\Theta\}$，其中 $P'_\theta(du)=u_\theta m(du)$。

众所周知，从 $\mathscr{X}$ 到 $U(\Theta)$ 的变换 $x\rightsquigarrow v(x)$ 是一个充分统计量，因此按照第一节距离定义，$\mathscr{E}$ 和 $\mathscr{E}'$ 是等价的。

事实上，任何根据 $v(x)$ 的信息而达到的风险，也能从已知 x 的信息而达到。因此，$\delta(\mathscr{E},\mathscr{E}')=0$。反过来说，称 $x\rightsquigarrow v(x)$ 是充分的，就是说在给定 v 时的条件期望不依赖 θ 值。在相当弱的条件下(将 $\mathscr{X}$ 空间完备化即可)，它的意义是：给定 v，x 的条件分布独立于 θ。在此情况下，将 P_θ 转换到 P'_θ 所要用的随机化概率 K 不依赖于 θ。最后，如果我们不用通常的 Markov 核 K 而用 Markov 核的极限，则所用的“弱条件”可以完全去掉。

根据上面的构造，m 的总质量为 k，对每个 $\theta\in\Theta$，$\int u_\theta dm=1$，我们称测度 m 为实验 $\mathscr{E}$ 的典型测度。为了明确表示 m 是从 $\mathscr{E}$ 得来的，有时将 m 记为 $m_{\mathscr{E}}$。$\mathscr{E}$ 的等价实验 $\mathscr{E}'=\{P'_\theta:\theta\in\Theta\}$ 称为 Blackwell 典型表达式，其中 $P'_\theta(du)=u_\theta m(du)$。

设 $\mathscr{M}$ 集为 $U(\Theta)$ 上满足下列条件的所有正测度 μ 组成：对每一个坐标函数 $u_\theta,\theta\in\Theta,\mu\in\mathscr{M}$ 满足 $\int u_\theta\mu(du)=1$。我们可用其中任何一个 μ 来定义一个实验，记它为 $\mathscr{F}=\{Q_\theta:\theta\in\Theta\}$，$Q_\theta(du)=u_\theta\mu(du)$。

在第一节中提到，任何两个实验 $\mathscr{E}$ 和 $\mathscr{F}$ 的距离可以用 $\Delta(\mathscr{E},\mathscr{F})$ 来度量，也可用任何度量测度 $m_{\mathscr{E}}$ 和 $m_{\mathscr{F}}$ 的距离来度量。这里我们将采用 Dudley 以及其他一些作者用的“对偶 Lipschitz 距离 (dual Lipschitz distance)”。对 $U(\Theta)$ 中的任何两个向量 u' 和 u''，令

$$|u'-u''|=\sup_\theta\{|u'_\theta-u''_\theta|:\theta\in\Theta\}.$$

对 $U(\Theta)$ 上的任何两个测度 μ' 和 μ''，对偶 Lipschitz 距离定