Springer Series in Statistics

Lucien Le Cam, Grace Lo Yang

Asymptotics in Statistics

Some Basic Concepts

Second Edition

统计学中的渐近性

基本概念

第2版

Springer

世界图书出版公司
www.wpcbj.com.cn

Lucien Le Cam
Grace Lo Yang

Asymptotics in Statistics

Some Basic Concepts

Second Edition

Springer

Lucien Le Cam
Department of Statistics
University of California, Berkeley
Berkeley, CA 94720
USA

Grace Lo Yang
Department of Mathematics
University of Maryland
College Park, MD 20742
USA
gly@math.umd.edu

Library of Congress Cataloging-in-Publication Data
Le Cam, Lucien M. (Lucien Marie), 1924–
Asymptotics in statistics : some basic concepts / Lucien Le Cam, Grace Lo Yang.
—2nd ed.
p. cm. — (Springer series in statistics)
Includes bibliographical references and index.
ISBN 0-387-95036-2 (alk. paper)
1. Mathematical statistics—Asymptotic theory. I. Yang, Grace Lo. II. Title.
III. Series.
QA276 .L336 2000
519.5—dc21 00-030759

图书在版编目（CIP）数据

统计学中的渐近性：基本概念：第2版 = Asymptotics in Statistics：Some basic Concepts Second Edition：英文/（美）卡姆（Cam，L. L.）著.—影印本.—北京：世界图书出版公司北京公司，2015.6

ISBN 978-7-5100-9795-9

Ⅰ.①统…　Ⅱ.①卡…　Ⅲ.①统计学—英文　Ⅳ.①O29

中国版本图书馆 CIP 数据核字（2015）第 121035 号

Asymptotics in Statistics：Some Basic Concepts Second Edition

统计学中的渐近性：基本概念　第2版

著　　者： Lucien Le Cam and Grace Lo Yang
责任编辑： 刘　慧　岳利青
装帧设计： 任志远

出版发行： 世界图书出版公司北京公司
地　　址： 北京市东城区朝内大街 137 号
邮　　编： 100010
电　　话： 010-64038355（发行）　64015580（客服）　64033507（总编室）
网　　址： http://www.wpcbj.com.cn
邮　　箱： wpcbjst@vip.163.com
销　　售： 新华书店
印　　刷： 三河市国英印务有限公司
开　　本： 711mm×1245 mm　1/24
印　　张： 13
字　　数： 249.6 千
版　　次： 2015 年 7 月第 1 版　　2015 年 7 月第 1 次印刷
版权登记： 01-2015-1559
ISBN 978-7-5100-9795-9　　　**定价：** 52.00 元

Springer Series in Statistics

Springer
New York
Berlin
Heidelberg
Barcelona
Hong Kong
London
Milan
Paris
Singapore
Tokyo

Springer Series in Statistics

Andersen/Borgan/Gill/Keiding: Statistical Models Based on Counting Processes.
Berger: Statistical Decision Theory and Bayesian Analysis, 2nd edition.
Bolfarine/Zacks: Prediction Theory for Finite Populations.
Borg/Groenen: Modern Multidimensional Scaling: Theory and Applications
Brockwell/Davis: Time Series: Theory and Methods, 2nd edition.
Chen/Shao/Ibrahim: Monte Carlo Methods in Bayesian Computation.
Efromovich: Nonparametric Curve Estimation: Methods, Theory, and Applications.
Fahrmeir/Tutz: Multivariate Statistical Modelling Based on Generalized Linear Models.
Farebrother: Fitting Linear Relationships: A History of the Calculus of Observations 1750-1900.
Federer: Statistical Design and Analysis for Intercropping Experiments, Volume I: Two Crops.
Federer: Statistical Design and Analysis for Intercropping Experiments, Volume II: Three or More Crops.
Fienberg/Hoaglin/Kruskal/Tanur (Eds.): A Statistical Model: Frederick Mosteller's Contributions to Statistics, Science and Public Policy.
Fisher/Sen: The Collected Works of Wassily Hoeffding.
Good: Permutation Tests: A Practical Guide to Resampling Methods for Testing Hypotheses, 2nd edition.
Gouriéroux: ARCH Models and Financial Applications.
Grandell: Aspects of Risk Theory.
Haberman: Advanced Statistics, Volume I: Description of Populations.
Hall: The Bootstrap and Edgeworth Expansion.
Härdle: Smoothing Techniques: With Implementation in S.
Hart: Nonparametric Smoothing and Lack-of-Fit Tests.
Hartigan: Bayes Theory.
Hedayat/Sloane/Stufken: Orthogonal Arrays: Theory and Applications.
Heyde: Quasi-Likelihood and its Application: A General Approach to Optimal Parameter Estimation.
Huet/Bouvier/Gruet/Jolivet: Statistical Tools for Nonlinear Regression: A Practical Guide with S-PLUS Examples.
Kolen/Brennan: Test Equating: Methods and Practices.
Kotz/Johnson (Eds.): Breakthroughs in Statistics Volume I.
Kotz/Johnson (Eds.): Breakthroughs in Statistics Volume II.
Kotz/Johnson (Eds.): Breakthroughs in Statistics Volume III.
Küchler/Sørensen: Exponential Families of Stochastic Processes.
Le Cam: Asymptotic Methods in Statistical Decision Theory.
Le Cam/Yang: Asymptotics in Statistics: Some Basic Concepts, 2nd edition.
Longford: Models for Uncertainty in Educational Testing.
Miller, Jr.: Simultaneous Statistical Inference, 2nd edition.
Mosteller/Wallace: Applied Bayesian and Classical Inference: The Case of the Federalist Papers.
Parzen/Tanabe/Kitagawa: Selected Papers of Hirotugu Akaike.
Politis/Romano/Wolf: Subsampling.

(continued after index)

Dedicatory Note

On April 25, 2000, as the last editorial touches were added to our book, Lucien Le Cam passed away. At Berkeley, he worked without stopping on the final revision of our book up until the time when he was hospitalized, just four days before his death.

His passing is a great loss to the statistical profession. He was a fundamental thinker in mathematical statistics and a principal architect of the modern asymptotic theory of statistics, following and extending the path of Neyman and Wald. Among the numerous seminal concepts that he introduced, the Le Cam distance between experiments provides a coherent statistical theory that links asymptotics and decision theory.

I was privileged and extremely fortunate to be his student and later his collaborator on several projects, including this book. Statisticians all know that Professor Le Cam was a brilliant mathematician. His students, colleagues, and friends also know of his kindness, generosity, integrity and patience.

This book is dedicated to the memory of Professor Le Cam. We feel his loss keenly. He will be greatly missed.

Grace Lo Yang
College Park, Maryland

May 2000

Preface to the Second Edition

This is a second edition, "reviewed and enlarged" as the French used to say. The 1990 edition was generally well received, but complaints about the conciseness of the presentation were heard. We have tried to accommodate such concerns, aiming at being understood by a "good" second-year graduate student.

The first edition contained some misleading typos. They have been removed, but we cannot guarantee that the present version is totally free of misprints.

Among substantial changes, we have introduced a Chapter 4 on gaussian and Poisson experiments. The gaussian ones are used everywhere because of mathematical tractability. Poisson experiments are less tractable and less studied, but they are probably more important. They will loom large in the new century.

The proof of asymptotic sufficiency of the "centerings" of Chapter 6 has been reworked and simplified. So has Chapter 7, Section 4. Section 5 of the same chapter has been augmented by a derivation of the Cramér-Rao lower bound imitated from Barankin [1949].

We thought of adding material on the von Mises differentiable functions and similar entities; we did not. The reader can refer to the recent book by van der Vaart [1998]. Similarly, we did not cover the use of "empirical processes" even though, following Dudley's paper of 1978, they have been found most valuable. A treatment of that material is expected to appear in a book by Sara van de Geer. Also, we did not include material due to David Donoho, Iain Johnstone, and their school. We found ourselves unprepared to write a distillate of the material. We did touch briefly on "nonparametrics," but not on "semiparametrics." This is because we feel that the semiparametric situation has not yet been properly structured.

We hope that the reader will find this book interesting and challenging, in spite of its shortcomings.

The material was typed in LaTeX form by the authors themselves, borrowing liberally from the 1990 script by Chris Bush. It was reviewed anonymously by distinguished colleagues. We thank them for their kind encouragement. Very special thanks are due to Professor David Pollard who took time out of a busy schedule to give us a long list of suggestions. We did not follow them all, but we at least made attempts.

We wish also to thank the staff of Springer-Verlag for their help, in particular editor John Kimmel, who tried to make us work with all deliberate speed. Thanks are due to Paul Smith, Te-Ching Chen and Ju-Yi-Yen, who helped with the last-minute editorial corrections.

Lucien Le Cam
Berkeley, California
Grace Lo Yang
College Park, Maryland

February 2000

Preface to the First Edition

In the summer of 1968 one of the present authors (LLC) had the pleasure of giving a sequence of lectures at the University of Montreal. Lecture notes were collected and written out by Drs. Catherine Doléans, Jean Haezendonck and Roch Roy. They were published in French by the Presses of the University of Montreal as part of their series of *Séminaires de Mathématiques Supérieures*. Twenty years later it was decided that a Chinese translation could be useful, but upon prodding by Professor Shanti Gupta at Purdue we concluded that the notes should be updated and rewritten in English and in Chinese. The present volume is the result of that effort.

We have preserved the general outline of the lecture notes, but we have deleted obsolete material and sketched some of the results acquired during the past twenty years. This means that while the original notes concentrated on the LAN situation we have included here some results of Jeganathan and others on the LAMN case. Also included are versions of the Hájek-Le Cam asymptotic minimax and convolution theorems with some of their implications. We have not attempted to give complete coverage of the subject and have often stated theorems without indicating their proofs.

What we have attempted to do is to present a few concepts and tools in an elementary manner refering the reader to the general literature for further information. We hope that this will provide the reader with a way of thinking about asymptotic problems in statistics that is somewhat more coherent than the traditional reliance upon maximum likelihood.

We wish to extend our thanks to the Presses of the University of Montreal for the permission to reuse some copyrighted material and to Springer-Verlag for the production of the present volume.

We also extend all our thanks to Professor Kai-Tai Fang whose efforts with Science Press are very much appreciated.

The English version of the manuscript was typed at Berkeley by Ms. Chris Bush whose patience and skill never cease to amaze us.

As Chris can attest, producing the typescript was no simple task. We were fortunate to have the help of Ruediger Gebauer and Susan Gordon at Springer-Verlag. We are very grateful for the assistance of Mr. Jian-Lun Xu in the preparation of the Chinese version.

Lucien Le Cam
Berkeley, California
Grace Lo Yang
College Park, Maryland

October 1989

Contents

1

Introduction

In this volume we describe a few concepts and tools that we have found useful in thinking about asymptotic problems in statistics. They revolve largely around the idea of approximating a family of measures, say, $\mathcal{E} = \{P_\theta; \theta \in \Theta\}$ by other families, say, $\mathcal{F} = \{Q_\theta; \theta \in \Theta\}$ that may be better known or more tractable.

For example, contemplate a situation where the statistician observes a large number of independent, identically distributed variables $X_1, X_2, \ldots, X_n$ that have a common Cauchy distribution with density

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

on the line.

Let $P_{\theta,n}$ be the joint distribution of $X_1, \ldots, X_n$. Let Z_n be another variable that has a gaussian distribution $G_{\theta,n}$ with expectation θ and variance $2/n$ on the real line.

The theory expounded in later chapters says that, for n large, the two families $\mathcal{E}_n = \{P_{\theta,n}; \theta \in \Re\}$ and $\mathcal{F}_n = \{G_{\theta,n}; \theta \in \Re\}$ are, for most statistical purposes, very close to each other.

For another example, suppose that $Y_1, Y_2, \ldots Y_n$ are independent with a common density $[1 - |x - \theta|]^+$ on the line. Let $Q_{\theta,n}$ be their joint distribution and let $H_{\theta,n}$ be gaussian with mean θ and variance $1/(n \log n)$. Then, for n large, $\{Q_{\theta,n}; \theta \in \Re\}$ and $\{H_{\theta,n}; \theta \in \Re\}$ are very close to each other.

Chapter 2 introduces distances that are intended to make precise the above "close to each other." The ideas behind the possible introduction of such distances go back to Wald [1943]. They are also related to the "comparison of experiments" described by Blackwell [1951] and others. Here, following Blackwell, we shall use the name "experiment" for a family $\mathcal{E} = \{P_\theta; \theta \in \Theta\}$ of probability measures P_θ carried by a σ-field $\mathcal{A}$ of subsets of a set $\mathcal{X}$. The set Θ is often called the "parameter space." It is convenient to think of each θ as a theory that provides a stochastic model P_θ for the observation process to be carried out by the experimenter.

Note that in the preceding example of comparison between the Cauchy $\{P_{\theta,n}; \theta \in \Re\}$ and the gaussian $\{G_{\theta,n}; \theta \in \Re\}$, the parameter space is the same $\Theta = \Re$ but the Cauchy observations are in a space $\mathcal{X} = \Re^n$ while the gaussian Z_n is one dimensional. The distance introduced in Chapter 2 gives a number for any pair $\mathcal{E} = \{P_\theta; \theta \in \Theta\}$ and $\mathcal{F} = \{Q_\theta; \theta \in \Theta\}$ provided that they have the same parameter space Θ. Chapter 2 also gives, for Θ finite, a standard representation of experiments indexed by Θ in the form of their Blackwell canonical representation. It shows that, for finite fixed Θ, convergence in the sense of our distance is equivalent to convergence of the distribution of likelihood ratios.

Chapter 3 is about some technical problems that occur in the convergence of likelihood ratios. They are often simplified drastically by the use of a condition called "contiguity." That same chapter also introduces Hellinger transforms and Hellinger distances. They are particularly useful in the study of experiments where one observes many independent observations.

Chapter 4 is about gaussian experiments and Poisson experiments. By "gaussian," we mean gaussian shift experiments where the covariance structure is supposed to be known and is not part of the parameter space. The parameter space describes the possibilities available for the expectations of the gaussian distributions. The system is described in a general, infinite-dimensional setup to accommodate nonparametric or semi-parametric studies. Poisson experiments occur in many studies concerned with "point processes." They are not as mathematically tractable as the gaussian experiments, but they may be more important. Both Poisson and gaussian experiments are cases of infinitely divisible experiments, the gaussian case representing a form of degeneracy obtainable by passages to the limit. Both Poisson and gaussian experiments can be obtained in situations where one has many *independent* observations.

Some of the limit theorems available in that situation form the object of Chapter 5. It gives, in particular, a form of what Hájek and Šidák were friendly enough to call "Le Cam's three lemmas."

Chapter 6 is about the LAN conditions. LAN stands for local asymptotic normality, which really means local asymptotic approximation by a gaussian shift experiment linearly indexed by a k-dimensional space. In addition to detailing some of the consequences of the LAN conditions around a point, the chapter contains the description of a method of construction of estimates: One starts with a good auxiliary estimate θ_n^* and picks a suitable set of vectors $\{u_{n,i}; i = 0, 1, 2, \ldots, k\}$ with $u_{n,0} = 0$ and $\{u_{n,i}; i = 1, \ldots k\}$, a basis of the parameter space $\Re^k$. Then one fits a quadratic to the logarithms of likelihood ratios computed at the points $\theta_n^* + u_{n,i} + u_{n,j}$, $i, j = 0, 1, \ldots, k$. One takes for the new estimate T_n the point of $\Re^k$ that maximizes the fitted quadratic.

In the LAN case, the estimate so constructed will be asymptotically minimax, asymptotically sufficient. It will also satisfy Hájek's convolution theorem, proved here by van der Vaart's method. The chapter ends with a description of what happens in the locally asymptotically *mixed* normal (LAMN) case. Here, we cite mostly results taken from Jeganathan's papers, referring to the books by Basawa and Prakasa Rao, [1980], Basawa and Scott [1983], Prakasa Rao [1987], and Greenwood and Shiryaev [1985] for other results and examples.

Chapter 7 comes back to the case of independent observations, describing what the LAN conditions look like in that case and more particularly in the standard independent identically distributed case. Most statisticians have heard of a theory of maximum likelihood based on Cramér's conditions. The theory obtainable by application of the results of Chapter 6 is somewhat similar but more in tune with the concepts of Chapter 2. Its conditions are weaker than those used by Cramér. A sufficient condition is an assumption of differentiability in quadratic mean, discussed at some length here. The theory also works in other cases, as shown by several examples. We have also in-