Lecture Notes in Mathematics

Edited by A. Dold and B. Eckmann

879

Vjačeslav V. Sazonov

Normal Approximation – Some Recent Advances

Springer-Verlag
Berlin Heidelberg New York 1981

Author

Vjačeslav V. Sazonov
Steklov Institute of Mathematics
42 Vavilov Street, 117966 Moscow, USSR

AMS Subject Classifications (1980): 60B12, 60E10, 60E99, 60F05

ISBN 3-540-10863-7 Springer-Verlag Berlin Heidelberg New York
ISBN 0-387-10863-7 Springer-Verlag New York Heidelberg Berlin

Printed in Germany

Printing and binding: Beltz Offsetdruck, Hemsbach/Bergstr.
2141/3140-543210

Preface

This course is devoted to the problem of estimation of the speed of convergence in the central limit theorem in R^k and in Hilbert space H . If $X_1, \ldots, X_n$ are independent identically distributed random variables with values in R^k or H and with $E|X_1|^2 < \infty$, P_n is the distribution of $n^{-1/2}\sum_{1}^{n}(X_j - EX_j)$ and Q is the normal probability measure with the same first and second moments as X_1 then the central limit theorem states that for a large class of functions f

$$\int f(x)(P_n - Q)(dx) \qquad (*)$$

tends to zero as $n \to \infty$; the problem is to construct upper bounds for the absolute value of $(*)$. This topic in the finite dimensional case was also covered in the recent monograph [8] . The difference between our approach and that of [8] is that we use mainly a method depending directly on convolutions rather then the method of characteristic functions. For the first time the method of convolutions in its explicit form was used by H.Bergström in 1944. In a number of respects it is simplier then the method of characteristic functions and leads to the goal more directly. We illustrate the method of convolutions in a simple example in §1 of Chapter I. In deriving the integral type estimates however we also apply the method of characteristic functions since we do not have proofs based on the method of convolutions only. Most of our estimates are expressed not in terms of absolute moments as in [8] but in terms of pseudomoments (for definitions see §1,2 of Chapter I). This makes the estimates sharper, the pseudomoments being distances from the distribution of summands to the corresponding normal measure. When the pseudomoments are small enough - less then some absolute constant - estimates for the variation distance are obtained which have the usual order $n^{-1/2}$. Special attention is paid to the constants involved in the estimates: in the k-dimensional case upper bounds of the form ck^{m} are obtained, where c is an absolute constant and m is a constant corresponding to the type of the estimate.

We do not try to prove results in full generality or to cover all the known achievements in the area. Our main aim is to present the main directions and methods (with the emphasis on the method of convolutions).

The Notes were written at the time the author was lecturing at UCLA in the spring of 1979 and at Moscow State University in 1979-1980. In the lectures Chapter I and §1 of Chapter II were covered (at UCLA Chapter I only). The content of Chapter II, §2 was also planned but was not delivered owing to lack of time.

I am greatly indebted to Professor Yu.V.Prohorov for his valuable advice and criticism during my work in this field. I am also grateful to Professor A.V.Balakrishnan for inviting me to UCLA and for his suggestion to write these Notes. While I was at UCLA we had many inspiring conversations about the theory of probability.

My special thanks go to Miss Stella Lozano of UCLA for typing a part of the first draft and to Mr. V.P.Sharma of ISI, Delhi and to Wendy Coutts of UBC, Vancouver for typing the final version of the manuscript.

V.V.Sazonov

Table of Contents

THE MAIN NOTATIONS

CHAPTER I : THE FINITE DIMENSIONAL CASE 1

§ 1. The two main methods and the main directions of extensions and improvements of classical results . . . 1

§ 2. The method of compositions. Estimates for convex and arbitrary Borel sets 11

§ 3. The method of characteristic functions. Integral type estimates 40

Comments on Chapter I 67

CHAPTER II : THE INFINITE DIMENSIONAL CASE 69

§ 1. An estimate for balls with a fixed centre 70

§ 2. A better speed estimate for balls with a fixed centre 82

Comments on Chapter II 100

REFERENCES 102

INDEX 105

The Main Notations

R^k	standard k-dimensional Euclidean space
H	a real separable Hilbert space
(x,y)	for $x,y \in R^k$ or H is the inner product of x and y $\left((x,y) = \sum_1^k x_j y_j,\ \text{for}\ x=(x_1,\dots,x_k),\ y=(y_1,\dots,y_k) \in R^k\right)$
$\lvert x\rvert = (x,x)^{\frac{1}{2}}$	
$\lVert a\rVert = \sum_1^k a_j$	for a nonnegative integral vector $a=(a_1,\dots,a_k) \in R^k$ (i.e. with nonnegative integer components a_j)
$a!$	for a nonnegative integral vector $a \in R^k$ is $a_1! a_2! \dots a_k!$.
D_a	for a nonnegative integral vector $a \in R^k$ is the differential operator $\partial^{\lVert a\rVert}/\partial x_1^{a_1} \dots \partial x_k^{a_k}$
x^a	for $x \in R^k$ and a nonnegative integral vector $a \in R^k$ is $x_1^{a_1} \dots x_k^{a_k}$
$x \le t$	for $x=(x_1,\dots,x_k)$, $y=(y_1,\dots,y_k) \in R^k$ means $x_j \le y_j$, $j=\overline{1,k}$
$x < y$	for $x,y \in R^k$ means $x \le y$, $x \ne y$
$\lvert V\rvert$	for a $k \times k$ matrix V is its determinant
$\lVert V\rVert$	for a $k \times k$ matrix is its norm, i.e. $\lVert V\rVert = \sup_{\lvert x\rvert \le 1} \lvert Vx\rvert$
I	identity $k \times k$ matrix
$S_r(x)$	open ball of radius r with centre at x
$S_{r,\varepsilon}(x)$	$= S_{r+\varepsilon}(x) \setminus S_r(x)$
χ_E	indicator function of a set E, i.e. $\chi_E(x) = 1$ or 0 according to $x \in E$ or $x \notin E$
E^c	complement of a set E
$\bar{E}$	closure of a set E
∂E	boundary of a set E
$E^{\varepsilon} = \cup\{S_{\varepsilon}(x) : x \in E\}$	
$E^{-\varepsilon} = \cup\{x : S_{\varepsilon}(x) \subset E\}$	
$\mathcal{C}$	class of all convex Borel subsets of R^k
P_X	probability measure corresponding to a random variable X
$\hat{K}$	for a probability measure or a distribution function K is its characteristic function, $\hat{K}(x) = \int \exp\{i(x,y)\} K(dy)$

$\|M\|$	for a signed measure M denotes its variation, i.e. $\|M\| = M^+ + M^-$, where M^+, M^- are the components of Jordan-Hahn decomposition of M
K^n	where K is a distribution function or a probability measure is the n-fold convolution of K with itself
$N_{\mu,V}$	normal distribution in R^k with mean μ and covariance matrix V
$N_T, T > 0$	is $N_{0,T^{-2}I}$
$N \quad =$	N_1
Φ	is normal (0,1) distribution function
$\phi_{\mu,V}$ (resp. ϕ_T, ϕ)	density function corresponding to $N_{\mu,V}$ (resp. N_T, N)
c (resp. $c(.)$)	with or without indices denote positive constants (resp. positive constants depending only on quantities in parentheses); the same symbol may stand for different constants.

If the area of an integration or the set over which $\max_x$, $\inf_x$, etc. is taken is not indicated it is understood that the integration is over the whole space and $\max_x$, $\inf_x$, etc. is with respect to the all possible values of x.