Lecture Notes in Mathematics 1592

Karl Wilhelm Breitung

Asymptotic Approximations for Probability Integrals

Springer-Verlag

Karl Wilhelm Breitung

Asymptotic Approximations for Probability Integrals

Springer-Verlag
Berlin Heidelberg New York
London Paris Tokyo
Hong Kong Barcelona
Budapest

Author

Karl Wilhelm Breitung
Department of Civil Engineering
University of Calgary
2500 University Drive, N.W.
Calgary, Alberta T2N 1N4, Canada

Mathematics Subject Classification (1991): 41A60, 41A63, 60F10, 60G15, 60G70, 62N05, 62P99, 90B25

ISBN 3-540-58617-2 Springer-Verlag Berlin Heidelberg New York

CIP-Data applied for

Typesetting: Camera ready by author
SPIN: 10130174 46/3140-543210 - Printed on acid-free paper

Editorial Policy

§ 1. Lecture Notes aim to report new developments - quickly, informally, and at a high level. The texts should be reasonably self-contained and rounded off. Thus they may, and often will, present not only results of the author but also related work by other people. Furthermore, the manuscripts should provide sufficient motivation, examples and applications. This clearly distinguishes Lecture Notes manuscripts from journal articles which normally are very concise. Articles intended for a journal but too long to be accepted by most journals, usually do not have this "lecture notes" character. For similar reasons it is unusual for Ph. D. theses to be accepted for the Lecture Notes series.

§ 2. Manuscripts or plans for Lecture Notes volumes should be submitted (preferably in duplicate) either to one of the series editors or to Springer- Verlag, Heidelberg . These proposals are then refereed. A final decision concerning publication can only be made on the basis of the complete manuscript, but a preliminary decision can often be based on partial information: a fairly detailed outline describing the planned contents of each chapter, and an indication of the estimated length, a bibliography, and one or two sample chapters - or a first draft of the manuscript. The editors will try to make the preliminary decision as definite as they can on the basis of the available information.

§ 3. Final manuscripts should preferably be in English. They should contain at least 100 pages of scientific text and should include
- a table of contents;
- an informative introduction, perhaps with some historical remarks: it should be accessible to a reader not particularly familiar with the topic treated;
- a subject index: as a rule this is genuinely helpful for the reader.

Further remarks and relevant addresses at the back of this book.

Lecture Notes in Mathematics 1592

Preface

In this lecture note asymptotic approximation methods for multivariate integrals, especially probability integrals, are developed. It is a revised version of my *habilitationsschrift* "Asymptotische Approximationen für Wahrscheinlichkeitsintegrale". The main part of this research was done when I worked at the Technical University and the University of Munich.

The motivation to study these problems comes from my work in the research project "Zuverlässigkeitstheorie der Bauwerke" (reliability theory of structures) at the Technical University of Munich in the department of civil engineering. For the tolerant support of the mathematical research in this project I would like to thank Prof. Dr.-Ing. H. Kupfer.

I am grateful to Prof. Dr.-Ing. R. Rackwitz and Prof. Dr.-Ing. G.I. Schuëller that they made me clear the engineering topics of reliability theory and helped me in my research. Further I would like to thank my former colleagues at the Technical University of Munich and at the University of Munich which supported me during my work at these universities: Dr.-Ing. B. Fießler, Dr. M. Hohenbichler, Dr. A. Rösch, Dr. H. Schmidbauer and Dr. C. Schneider. Additionally I would like to express my gratitude to Prof. Dr. F. Ferschl for pointing out occasional errors and misprints in the original German version.

The major part of this revision was made when I stayed as visiting fellow at the University of New South Wales in 1991. I would like to thank especially Prof. A. M. Hasofer for his help and for his kind invitation to the University of New South Wales and to express my delight at having worked there.

For their help and discussions I thank wholeheartedly Prof. Dr. F. Casciati, Prof. Dr. L. Faravelli (both University of Pavia), Prof. Dr. P. Filip (FH Bochum), Prof. Dr. K. Marti (University of the Federal Armed Forces at Neubiberg) and Prof. Dr. W.-D. Richter (University of Rostock). Prof. Dr. M. Maes (University of Calgary) knows what I mean.

For eliminating the worst bugs in my English I thank Poul Smyth, the Irish poet of the Bermuda triangle, and for making a cover design Ringo Praetorius, the executioner of Schichtl at the Oktoberfest. Unfortunately the publisher and the series editors decided not to use this cover design.

Finally a short comment about the mathematical level of this note should be made. It is intended also for mathematically interested reliability engineers. Probably, therefore, the mathematicians will complain about the low level and the inclusion of too much elementary material and the engineers will go the other way.

Calgary, September 1994

Karl Wilhelm Breitung

Notation

The set of the natural numbers is denoted by $\mathbb{N}$ and the set of complex numbers by $\boldsymbol{C}$. The n-dimensional euclidean space is denoted by $\mathbb{R}^n$. For the set of the vectors in $\mathbb{R}^n$ with all components being positive we write $\mathbb{R}^n_+$. A vector in $\mathbb{R}^n$ is written as $\boldsymbol{x}$ and the zero vector $(0,\ldots,0)$ as $\mathbf{o}$. The transpose of $\boldsymbol{x}$ is written as $\boldsymbol{x}^T$. The unit vector in direction of the x_i-axis is denoted by $\boldsymbol{e}_i$. For the euclidean norm of a vector $\boldsymbol{x}$ we write $|\boldsymbol{x}|$ and for the scalar product of two vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ we use $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$. The subspace of $\mathbb{R}^n$ spanned by k vectors $\boldsymbol{a}_1,\ldots,\boldsymbol{a}_k$ is written as $\mathrm{span}[\boldsymbol{a}_1,\ldots,\boldsymbol{a}_k]$. For the orthogonal complement of a subspace $U \subset \mathbb{R}^n$ we write $U^\perp$.

An $n \times m$-matrix is written with bold Roman letters: $\boldsymbol{A}$,$\boldsymbol{B}$,... The n-dimensional unity matrix is denoted by $\boldsymbol{I}_n$ and an $n \times k$ matrix consisting of zeros by $\mathbf{o}_{n,k}$. The cofactor matrix $\boldsymbol{C}$ of an $n \times n$ matrix $\boldsymbol{A}$ is the $n \times n$ matrix $\boldsymbol{C} = ((-1)^{i+j} \det(\boldsymbol{A}_{ij}))_{i,j=1,\ldots,n}$ with $\boldsymbol{A}_{ij}$ being the $(n-1) \times (n-1)$ matrix obtained from $\boldsymbol{A}$ by deleting the i-th row and the j-th column. The rank of a matrix $\boldsymbol{B}$, i.e. the number of its linearly independent column vectors, is denoted by $\mathrm{rk}(\boldsymbol{B})$.

The probability of an event A is denoted by $\mathbb{P}(A)$. An one-dimensional random variable is denoted by a capital Roman letter: $X, Y, \ldots$ and for n-dimensional random vectors bold capital Roman letters are used: $\boldsymbol{X}, \boldsymbol{Y}, \ldots$. For the probability density function and the cumulative distribution function of a random variable we write p.d.f. and c.d.f. respectively. The expected value of a random variable X is written as $\mathbb{E}(X)$ and its variance as $\mathrm{var}(X)$. The covariance between X and Y is denoted by $\mathrm{cov}(X, Y)$.

A function $f : D \to \mathbb{R}$ on an open set $D \subset \mathbb{R}^n$ is called a C^1-function if all partial derivatives of first order exist and are continuous. Analogously by induction C^r-functions $(r > 1)$ are defined. A function $f : D \to \mathbb{R}$ on an open set $D \subset \mathbb{R}^n$ is called a C^r-function if all partial derivatives of order $r-1$ exist and are continuously differentiable. Further a function $f : D \to \mathbb{R}$ on a closed set $D \subset \mathbb{R}^n$ is called a C^r-function if there is an open set $U \subset \mathbb{R}^n$ with $D \subset U$ such that f is defined on U and f is according to the definition above a C^r-function.

A function $T : \mathbb{R}^n \to \mathbb{R}^m, \boldsymbol{x} \mapsto T(\boldsymbol{x}) = (t_1(\boldsymbol{x}),\ldots,t_m(\boldsymbol{x}))$ is called a C^r-vector function if all component functions $t_1(\boldsymbol{x}),\ldots,t_m(\boldsymbol{x})$ are C^r-functions. For a function $f : \mathbb{R}^n \to \mathbb{R}$ the first partial derivatives with respect to the variables x_i $(i = 1,\ldots,n)$ at the point $\boldsymbol{x}$ are denoted by $f^i(\boldsymbol{x})$ or $\frac{\partial f(\boldsymbol{x})}{\partial x_i}$ and the gradient of by $\nabla f(\boldsymbol{x})$. The second derivatives of this function with respect to the variables x_i and x_j $(i, j = 1,\ldots,n)$ at the point $\boldsymbol{x}$ are denoted by $f^{ij}(\boldsymbol{x})$ or $\dfrac{\partial^2 f(\boldsymbol{x})}{\partial x_i \partial x_j}$ and its Hessian by $\boldsymbol{H}_f(\boldsymbol{x})$.

For functions of the form $f(\boldsymbol{x}, \boldsymbol{y})$ by $\nabla_{\boldsymbol{x}} f(\boldsymbol{x}, \boldsymbol{y})$ the gradient with respect to the vector $\boldsymbol{x}$ is denoted. In the same way $\nabla_{\boldsymbol{y}} f(\boldsymbol{x}, \boldsymbol{y})$ means the gradient with respect to the second vector $\boldsymbol{y}$. The divergence of a C^1-vector function $\boldsymbol{u}(\boldsymbol{x})$ is denoted by $\mathrm{div}(\boldsymbol{u}(\boldsymbol{x}))$.

Contents

Chapter 1

Introduction

1.1 The Evaluation of Multivariate Integrals

In many fields of applied mathematics it is necessary to evaluate multivariate integrals. In a few cases analytic solutions can be obtained, but normally approximation methods are needed. The standard methods here are numerical and Monte Carlo integration. In spite of the fact that computation time today is accessible plentiful, such procedures are sufficient for many problems, but in a number of cases they do not produce satisfactory results. Such integrals appear for example in structural reliability, stochastic optimization, mathematical statistics, theoretical physics and information theory.

If we consider an integral in the form

$$\int_F f(\boldsymbol{x})\, d\boldsymbol{x} \tag{1.1}$$

with $F \subset \mathbb{R}^n$ and $f : F \to \mathbb{R}$, there are three main causes, which might make numerical or Monte Carlo integration difficult:

1. The dimension n of the integration domain is large.
2. The domain F has a complicated shape.
3. The variation of f in the domain F is large.

Often not only the integral itself, but its derivatives with respect to parameters are of interest. The most general case is that the integrand $f(\boldsymbol{x})$ as well as the integration domain F depend on parameters.

Another important point is that often not only one integral has to be computed, but the behavior of the integral under parameter changes is of interest. Then many evaluations of the integral are necessary. A similar problem occurs if some sort of optimization should be made. Further in reliability problems with random finite elements (see for example [4]) even now the necessary computing time can be prohibitive.

Methods for analytic approximations have been developed in different fields and sometimes due to the specialization nowadays in science some of them have been rediscovered at least once or twice. Therefore it is attempted here to give a list of the available textbooks in this field and certainly for the field of multivariate Laplace methods an overview of the relevant results.

The basic idea of such methods is that instead of integrating over the whole domain F, points or subsets are identified from whose neighborhoods the main contributions to the integral come from. Therefore instead of integrating the numerical problem is then to find these sets by some optimization method.

Such approximation methods are not a solution for all problems. Their efficiency and usefulness depends on the problem. If the underlying assumptions about the structure of the integral are not fulfilled to some degree, the use of other schemes might be better. In some cases then the use of asymptotic approximation methods, which are described in this book, perhaps in combination with the aforementioned methods, is advisable.

1.2 Structural Reliability

One field, in which such concepts had been very successful, is structural reliability. In the following we will give a short outline of this field.

The first proposals to use probabilistic models for structural reliability are from [98], [78] and [64], but not before the sixties such problems were studied more intensively. Then in the last thirty years in mechanical and civil engineering probabilistic methods were developed for the calculation of the reliability of components and structures, since pure deterministic methods were not adequate for a number of problems. Textbooks about such methods are [10], [16], [17], [39], [84], [89], [99], [124] and [128].

In structural reliability the computation of multivariate integrals is an important problem. The studies in this lecture note were motivated by the problems in this field, since standard methods did not lead to satisfactory results.

At the beginning the random influences acting on a structure were modelled simply by two time-invariant random variables, a load variable L and a resistance variable R. If $L \geq R$, the structure failed, if $L < R$, it remained intact.

But soon it was clear that such a model was far too simple even for components of a structure. Even if the time influence is neglected, for a sufficient description of the random influences a random vector $\boldsymbol{X} = (X_1, \ldots, X_n)$ with a large number n of components is needed. In general one part of this vector is composed of load variables and the other of resistance variables.

If now the p.d.f. $f(\boldsymbol{x})$ of the random vector $\boldsymbol{X}$ which models the random influences on a structure is known and the conditions for failure can be expressed as a function of the vector, the probability for failure can be calculated. Then the integration domain is given by a function $g(\boldsymbol{x})$ in the form $\{\boldsymbol{x}; g(\boldsymbol{x}) \leq 0\}$. The function $g(\boldsymbol{x})$ describes the state of the structural system under consideration. If $g(\boldsymbol{x}) > 0$, the system is intact and if $g(\boldsymbol{x}) \leq 0$, the system fails.

The failure domain is denoted by $F = \{\boldsymbol{x}; g(\boldsymbol{x}) \leq 0\}$, the safe domain by $S = \{\boldsymbol{x}; g(\boldsymbol{x}) > 0\}$ and the limit state surface, the boundary of F by $G = \{\boldsymbol{x}; g(\boldsymbol{x}) = 0\}$.

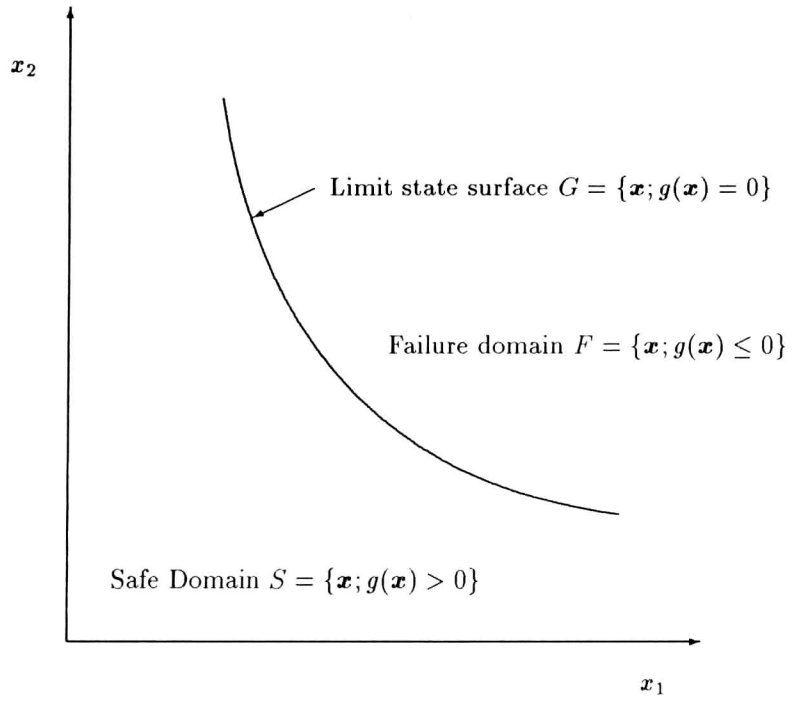

Figure 1.1: Failure domain, limit state surface and safe domain

Then the problem is to compute the probability of failure

$$\mathbb{P}(F) = \int_{g(\boldsymbol{x}) \leq 0} f(\boldsymbol{x})\, d\boldsymbol{x}. \tag{1.2}$$

Often we can distinguish two types of random variables in the random vector $\boldsymbol{X}$. Firstly random variables, which can not be changed by the design of the structure, such as loads acting on it; secondly variables, which can be influenced by the design, for example the strength of concrete.

It was tried first to use the standard methods for the calculation of such integrals mentioned above to compute these multivariate probability integrals. In numerical or Monte Carlo integration then the integral is approximated in

the form

$$\mathbb{P}(F) \approx \sum_{i=1}^{N} w(\boldsymbol{x}_i) f(\boldsymbol{x}_i). \tag{1.3}$$

Here the $\boldsymbol{x}_i$'s are points in F, determined by a deterministic or random mechanism and the $w(\boldsymbol{x}_i)$ are their respective weights. The main difficulties in computing this integral are that in general the probability density $f(\boldsymbol{x})$ is small in the failure domain and that the shape of the domain is not explicitly known, but given implicitly by the function $g(\boldsymbol{x})$; therefore standard techniques have difficulties in finding a scheme for creating enough points in F. Soon it became clear that in their usual form they were not suited to this problem, since they were too time consuming and inaccurate.

Certainly now with the increasing capacity of computers it is possible to solve more and more reliability problems, without spending any effort on refining methods. But the author thinks that it is absolutely wrong to say that we just have to wait until the computing capacity is large enough to make it possible to solve more complex problems. Since the main point in structural reliability and other fields is not only the computation of probabilities, but the gain of insight in the structure of the problem. If an approximate analytic solution is found, the important influences on the failure mechanism become often much clearer.

The usual formulation of the problem, compute $\mathbb{P}(F)$, given the density $f(\boldsymbol{x})$ and the limit state function $g(\boldsymbol{x})$, is misleading. In reality the probability distribution is known only approximately and also the limit state function is only an appproximation. The problem of estimating such a function by experimental design is considered in [58] and [32]. Der Kiureghian [44] and Maes [90] treat the problem of structural reliability under model uncertainties.

Furthermore in a problem of structural reliability the probability of failure depends on a number of design parameters in the structure. Some of these parameters can be changed by altering the design of the structure, others not. Therefore the real problem here is more complex than it appears at the first glance.

Therefore a more realistic probabilistic model is in the form

$$\mathbb{P}(F|\boldsymbol{\vartheta}) = \int_{g(\boldsymbol{x}|\boldsymbol{\vartheta})\leq 0} f(\boldsymbol{x}|\boldsymbol{\vartheta})\, d\boldsymbol{x}. \tag{1.4}$$

Here all functions depend on a parameter vector $\boldsymbol{\vartheta} = (\vartheta_1, \ldots, \vartheta_k)$.

1. $g(\boldsymbol{x}|\boldsymbol{\vartheta})$ is the limit state function for the parameter value $\boldsymbol{\vartheta}$.

2. $f(\boldsymbol{x}|\boldsymbol{\vartheta})$ is the probability density function of the random vector $\boldsymbol{X}$ for the parameter value $\boldsymbol{\vartheta}$.

The usual formulation is that the parameter vector $\boldsymbol{\vartheta}$ is assumed to be fixed and known. Then for this fixed vector $\boldsymbol{\vartheta} = \boldsymbol{\vartheta}_0$ the probability $\mathbb{P}(F|\boldsymbol{\vartheta}_0)$ is computed. But in the more general setting this is only one of the interesting

quantities. If the values of the parameters are known and can be influenced by the design of the structure, one problem can be to optimize the structure in some sense. In reliability the quantity to be minimized is usually the probability of failure, but there are restrictions on the possible parameter values.

The following quantities and distributions may be of interest:

1. The partial derivatives of $I\!P(F|\boldsymbol{\vartheta})$ with respect to the components $\vartheta_1, \ldots, \vartheta_k$ of the parameter vector.

2. The distribution of $I\!P(F|\boldsymbol{\vartheta})$ if there is a probability distribution of $\boldsymbol{\vartheta}$.

3. The asymptotic form of the conditional distribution of $\boldsymbol{X}$ under the condition $g(\boldsymbol{X}|\boldsymbol{\vartheta}) < 0$.

Until now only the time-invariant reliability problem is solved sufficiently. Often time-variant problems are transformed into a time-invariant form, for example by modelling a random process only by its extreme value distribution. This is related to the fact that the theory of multivariate stochastic processes and their extreme value theory is restricted mainly to Gaussian processes.

In probabilistic models for structural reliability we have two different probability theory and statistical inference. With methods of probability theory and other mathematical concepts the structure of the mathematical model is studied. The other problem is the relation of the model to reality. A common feature of the majority of books and papers in structural reliability is that the statistical problems are often neglected.

We will discuss shortly the problems of statistical inference in reliability modelling. Since in reliability problems data are often sparse or do not exist at all, conclusions are in many cases based on a lot of assumptions, whose correctness can not be proved. The results, which are obtained, are difficult to check. For example if scientist A computes a failure probability of 10^{-5} for a single building during one year and scientist B instead a probability of 10^{-6}, who is right? An approach for a solution might be the development of probabilistic models, which make more specific predictions (see for example [30] and [34]). This would mean that not only failure probabilities are computed, but the probability distributions of other events, which are connected with the occurrence of failures, but can be observed more frequently. By such an approach, we get an iterative process of model building, prediction, observation and model improvement.

In the book of Matheron [97] the general problem of probabilistic models in science is discussed. The importance of this book is that here it is made clear that the justification and the acceptance of these methods in science comes from the fact that they give for many problems satisfactory solutions, but not that they are correct descriptions of the reality. This might be quite unintelligible for statisticians who have never been involved in any applied work as they have lived in a pure academic environment. But anyone who works at problems which have some connection with the real world outside Academia will probably agree that this is not so wrong at all.

An alternative method in coping with uncertainty in structural reliability is for example fuzzy set theory (see [6]). In fuzzy set methods uncertainties, which are not of a probabilistic nature, can be modelled. A drawback is that there is no clear rule how to incorporate additional information as it is done for example in the Bayesian concept by Bayes' theorem.

A further alternative concept is convex modelling (see [9]). In convex modelling no specific probability distribution for the random influences is assumed and only bounds for admissible influences are derived.

Another problem of statistical inference in reliability is that the usual statistical estimation methods are focused on fitting distributions to the central part of the data and not to the tails. Here modified estimation procedures should be used (see [40]), which give more weight to the fit in the distribution tails, since in reliability calculation the risk is mainly in underestimating the extremes of random influences and not in making wrong estimates about their means.

1.3 Stochastic Optimization

A mathematical field, where similar problems occur, is stochastic optimization. Here a given stochastic system, for example a network, a technical structure or a queuing system, should be optimized in its performance, which is described by a function of the design parameters, but is in general not known analytically. Basic concepts of stochastic optimization can be found in the book of Rubinstein [123].

For a given stochastic system we have an integral in the form

$$I(\boldsymbol{\vartheta}) = \int\limits_{I\!R^n} R(\boldsymbol{x}|\boldsymbol{\vartheta}) f(\boldsymbol{x}|\boldsymbol{\vartheta})\, d\boldsymbol{x}, \tag{1.5}$$

and a set of k constraint functions

$$g_1(\boldsymbol{\vartheta}) = \ldots = g_k(\boldsymbol{\vartheta}) = 0. \tag{1.6}$$

Here $\boldsymbol{\vartheta}$ is an m-dimensional parameter vector, who can achieve values in a subset $V \subset I\!R^m$. $f(\boldsymbol{x}|\boldsymbol{\vartheta})$ with $\boldsymbol{x} \in I\!R^n$ is a p.d.f., $R(\boldsymbol{x}|\boldsymbol{\vartheta})$ is a function, which describes the performance of the system for the given values $\boldsymbol{\vartheta}$ and $\boldsymbol{x}$. The functions $g_1(\boldsymbol{\vartheta}), \ldots, g_k(\boldsymbol{\vartheta})$ give restrictions for the possible values of $\boldsymbol{\vartheta}$ and describe usually costs or available resources. Then $I(\boldsymbol{\vartheta})$ is the performance of the system for the parameter vector $\boldsymbol{\vartheta}$. In stochastic optimization for fixed values of the parameter $\boldsymbol{\vartheta}$ the values of at least some of the functions are not known exactly and have to be estimated by some procedure. Usually here the problem is considered that $g_i(\boldsymbol{\vartheta}) = I\!E(h_i(\boldsymbol{X}))$ and only random samples are available from which the expected value $I\!E(h_i(\boldsymbol{X}))$ has to be estimated.

In this formulation the task is to minimze the integral $I(\boldsymbol{\vartheta})$ under the restrictions in equation (1.6). For this purpose it is necessary to evaluate the function $I(\boldsymbol{\vartheta})$ and at least some of its sensitivities (derivatives, gradients, Hessians, etc.)

with respect to changes in the parameter vector ϑ. To compute these, various methods can be used. Here often Monte Carlo methods are adequate.

Applications of stochastic optimization methods in structural design are given in [96] and [95].

1.4 Large Deviations and Extreme Values

Large deviation theory studies the asymptotic behavior of the probabilities $I\!P(\lambda A)$ as $\lambda \to \infty$; here $I\!P(.)$ is a probability measure on a measurable space $(E, \mathcal{B})$ and A is an subset of E. If for example the underlying space is the n-dimensional Euclidean space $I\!R^n$, then such probabilities are given by n-dimensional integrals. The case of the standard normal probability measure is considered in the papers of Richter ([114], [14], [115], [116], [118] and [107]).

Similar questions appear if we study the asymptotic distribution of the sum $\sum_{i=1}^{n} X_i$ of random variables with mean zero (see [18], [19] and [119]).

Extreme value theory is concerned with the maxima (or minima) of sequences of random variables or of a stochastic process. If we consider a sequence $X_1, X_2, \ldots$ of i.i.d random variables, the classical question of extreme value theory was to find under which conditions by a suitable scaling and shifting a non-degenerate limit distributions exist for the sequences

$$Y_n = \max(X_1, \ldots, X_n) \text{ resp. } Z_n = \min(X_1, \ldots, X_n). \tag{1.7}$$

The classical textbook in this field is the book [68] of Gumbel written 1958. The textbook of Leadbetter, Lindgren and Rootzén [82] gives in the first part an overview of further development of the extreme value theory for sequences.

In the case of a stochastic process $X(t)$ we consider the random variables

$$Y(T) = \max_{0 \le t \le T} X(t) \text{ resp. } Z(T) = \min_{0 \le t \le T} X(t). \tag{1.8}$$

Here similar results as for sequences can be obtained, see [43] and [82]. The results in these books are usually derived by approximating the process by a suitably chosen sequence of random variables. An alternative method, based on sojourn times, is outlined in [12].

1.5 Mathematical Statistics

Similar questions arise in mathematical statistics. Here asymptotic methods play an important role in investigating large sample behavior. An overview of such applications can be found in the book of Barndorff-Nielsen and Cox [7]. Here a sequence of i.i.d. random variables $X_1, X_2, \ldots$ is given and for functions $f(X_1, \ldots, X_n)$ as $n \to \infty$ asymptotic approximations are sought. A standard example are the asymptotic behaviors of the sample mean $\bar{X} = n^{-1} \sum_{i=1}^{n} X_i$ and the sample variance $\bar{S} = (n-1)^{-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$, which converge under some regularity conditions to the respective moments of the X_i's.

If the random variable X_1 has the p.d.f. $f(x)$ the joint p.d.f. of the first n random variables is $\prod_{i=1}^{n} f(x_i)$ and the log likelihood function is $\ln(\prod_{i=1}^{n} f(x_i)) = \sum_{i=1}^{n} \ln(f(x_i))$. The asymptotic behavior of these function can be studied using the Laplace method. In the case of random vectors $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots$ we need again results about the asymptotic structure of multivariate integrals.

Additional importance have asymptotic approximation methods in Bayesian statistics, where the derivation of the posterior distribution requires often the evaluation of multivariate integrals. The first who used the Laplace method outlined in this book for such problems, i.e. the derivation of posterior distributions, was Lindley [86]. Further results are given in [131].

1.6 Contents of this Lecture Note

In chapter 2 first some results from linear algebra and analysis are given for reference and some new results about sufficient conditions for constrained extrema and multivariate parameter dependent integrals are derived. In the third chapter a short outline of the basic concepts of asymptotic analysis is given. Then in the fourth univariate and in the fifth multivariate Laplace type integrals are studied. In the sixth these results are applied to normal random vectors. In the seventh chapter it is shown that such approximations can be made also for non-normal random vectors. In the last chapter then asymptotic approximations for crossing rates of differentiable stationary Gaussian vector processes are derived.