Institute of Mathematical Statistics
LECTURE NOTES–MONOGRAPH SERIES

Analytic Statistical Models

Ib M. Skovgaard
Royal Veterinary and Agricultural University, Copenhagen

Volume 15

Institute of Mathematical Statistics
LECTURE NOTES–MONOGRAPH SERIES
Volume 15

Analytic Statistical Models

Ib M. Skovgaard
Royal Veterinary and Agricultural University, Copenhagen

Institute of Mathematical Statistics
Hayward, California

Institute of Mathematical Statistics

Lecture Notes–Monograph Series

The production of the *IMS Lecture Notes–Monograph Series* is managed by the IMS Business Office: Jessica Utts, IMS Treasurer, and Jose L. Gonzalez, IMS Business Manager.

Library of Congress Catalog Card Number: 90-86256

International Standard Book Number 0-940600-20-X

Printed in the United States of America

Preface

Asymptotic results given in the literature relating to parametric statistical inference may roughly be classified into one of the following three categories:

(1) The result is proved quite strictly under a list of specified conditions.

(2) The model is restricted to a certain class of 'well-behaved' models, such as the class of (curved) exponential families, and then the result is proved strictly under a few extra specified conditions.

(3) A 'heuristic proof' is given which is claimed to be valid under suitable regularity conditions.

All of these three approaches have their merits and demerits, most of which speak for themselves. The advantage of the third approach is that the progress of the asymptotic theory is quicker when strict proofs are not of direct concern. Often 'results' obtained in this way are verified in a strict setting later. The disadvantage of the heuristic approach is that it may be hard to draw the line between those results that are verifiable and those that are not. When applicable, the second approach often keeps the advantage of the third, because the extra conditions required are quite simple. At the same time strict results are obtained in this way; the only disadvantage seems to be the restriction on the range of models covered. The obvious advantages of the first approach are, to some extent, counteracted by the amount of work required to prove the results and, in particular, by the difficulties often involved in verifying the somewhat awkward conditions in specific applications.

It is the purpose here to extend the applicability of the second approach by the introduction of a class of statistical models (the analytic models) which is sufficiently well behaved to satisfy regularity conditions of the type typically met in theorems of asymptotic statistical inference, and at the same time sufficiently rich to contain many of the commonly used statistical models, including the (sufficiently smooth) curved exponential families. In this way the advantages of the three approaches sketched above may be combined, provided that a number of basic results are derived for the class of models. Thus, the main points are to define the class of analytic models, to derive its basic mathematical and probabilistic properties, to show that it contains a wide range of common statistical models, and to demonstrate its applicability in asymptotics.

The theory is adapted to likelihood based parametric statistical inference, but some of the techniques may be used in other connections also. Because of the dual role of the density function as the function used for inferential purposes and as the function describing the model, the likelihood based inference offers substantial simplifications, however.

It is by no means the intention to suggest that the asymptotic theory has been 'covered' even approximately by this treatment. The point has been to demonstrate that the class of analytic models may provide a suitable framework for further development of asymptotic theory of likelihood based inference for parametric statistical models. In particular, it is my hope that by use of this framework it will become easier to provide rigorous proofs of results of second and higher order statistical inference, without confining the results to curved exponential models. Especially in connection with conditioning on (approximately) ancillary statistics it is often questionable whether rigorous versions of heuristic derivations exist, because the conditioning to some extent 'destroys' the probabilistic properties available for independent replications. This problem has been a major motivation for the present work although it is not directly treated here; but possibly the result of Lemma 2.6.5 may be of help in this connection.

The layout of the material in chapters is sketched below. More detailed information about the content is given in the introduction to each of the chapters.

Chapter 1 contains auxiliary mathematical concepts and results, in particular in relation to multilinear functions. Throughout, the notation introduced here is used heavily. Chapter 2 begins with the definition of the class of analytic models and continues to explore its basic properties. Some of these properties are used later on, but others are included either for potential use or to increase the understanding of this class of models. Chapter 3 contains examples of analytic models and classes of such. A particular point here is to demonstrate that the analytic models contain the (sufficiently smooth) curved exponential families as well as a substantial amount of models that are not of this type. Chapters 4 and 5 contain examples of applications to asymptotic statistical theory. These are either examples of first order asymptotic results for general sequences of models (Chapter 4), or of higher-order asymptotic results for models for independent replications (Chapter 5). Chapter 4 contains first some general asymptotic results and then a number of examples for which the conditions for the results are verified. Some of these examples of special types of models may be of independent interest. Chapter 5 is confined to the general results because the conditions for these to hold are so simple that examples serve no purpose. This is so because only sequences of independent observations from the same model are in question here.

Because of the somewhat unusual notation employed, most readers probably need a quick reading of parts of the first four sections in Chapter 1, otherwise this chapter is mainly intended for reference. More precise guidelines are given at the beginning of the chapter. In Chapter 2, Sections 2-6 form the backbone of the theory. These sections give the definition of the class of analytic models, their basic properties, the main theorem (Theorem 2.4.2) deriving equivalent definitions, the definition of the essential quantity referred to as the index of a model, and some results used to prove that a model is analytic. This chapter is of a rather technical nature and the reader may find it illuminating to have a glance at the examples in Chapter 3 during the reading of Chapter 2. Chapters 3, 4 and 5 are largely independent of each other, although it is recommended to flip through the examples in Chapter 3 at an early stage in the process of reading.

Efforts have been made to make it possible to read parts of the monograph without reading everything from page 1. In particular, no references are made to previous proofs, except that within other proofs occasional references are made to techniques used previously or to quite specific notations. Thus it should be possible to skip proofs throughout, although this always implies some loss of insight, of course. Many definitions of notations are repeated frequently or referred to at least the first time they are met in each section, and as few abbreviations as feasible are introduced. While this may be somewhat annoying to the persistent reader, it hopefully facilitates more casual reading.

References to other theorems, definitions, etc., of the form Theorem 2.4 and equation (3.14), are within-chapter references, whereas, e.g., equation (1.2.24) and Definition 1.4.6 refer to Chapter 1.

An alphabetic list of notations is included, with reference to the introduction of each of the notations. As is common in statistics, random variables are denoted by upper case letters whereas the corresponding lower case letters denote points in the sample space, although this distinction is rather vague.

Readers familiar with the theory of tensor products may wonder why they are not introduced, since the theory developed here relies heavily on multilinear mappings. Indeed, parts of the theory could have been formulated more elegantly by use of the symmetric tensor product, but that has been avoided to limit the amount of mathematics introduced.

The present monograph serves as a dissertation for the Danish doctoral degree. A summary in Danish is included to meet the requirements in this connection.

A major part of this work was carried out while I visited the Department of Statistics at the University of British Columbia. Partial support in this connection was provided by the Natural Science and Engineering Research Council of Canada and the Danish Natural Science Research Council. I wish to thank, in particular, professor J. Zidek at the Department of Statistics at the University of British Columbia, Vancouver, for his arrangements to make this visit possible, and the entire staff at this department for their hospitality. It is also a pleasure to thank my colleagues at the Department of Mathematics at the Royal Veterinary and Agricultural University, Copenhagen, for their support of my work in general as well as for their efforts in connection with the extra amount of work during my leave of absence.

Table of contents