Translations of

MATHEMATICAL MONOGRAPHS

Volume 20

Statistical Problems with Nuisance Parameters

Ju. V. Linnik

American Mathematical Society

10.1090/mmono/020

Translations of

MATHEMATICAL MONOGRAPHS

Volume 20

Statistical Problems with Nuisance Parameters

Ju. V. Linnik

American Mathematical Society
Providence, Rhode Island

СТАТИСТИЧЕСКИЕ ЗАДАЧИ
С МЕШАЮЩИМИ ПАРАМЕТРАМИ

Ю. В. ЛИННИК

ТЕОРИЯ ВЕРОЯТНОСТЕЙ
И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

Издательство "Наука"
Главная Редакция
Физико-Математической Литературы
Москва 1966

Translated from the Russian by Scripta Technica

2000 *Mathematics Subject Classification.* Primary 62–XX.

Library of Congress Cataloging-in-Publication Data

Linnik, ÎU. V. (ÎUriĭ Vladimirovich), 1915–1972.
Statisticheskie zadachi s meshaîushchimi parametrami. English
Statistical problems with nuisance parameters, by Ju. V. Linnik. [Translated from the Russian by Scripta Technica].
p. cm. — (Translations of mathematical monographs, v. 20)
Includes bibliographical references.
1. Statistical hypothesis testing. 2. Estimation theory. I. Title. II. Series.

QA277.L513
519′.1 67030101

ISBN 978-0-8218-4687-2

10 9 8 7 6 5 4 3 2 1 13 12 11 10 09 08

PREFACE

The present book is devoted to the analytic theory of elimination of nuisance parameters in the testing of statistical hypotheses and to the theory of unbiased estimates. Our attention is concentrated on the analytic properties of tests and estimates and on the mathematical foundations for obtaining a test or unbiased estimate that is optimal in some sense or other. It does not, however, include either computational algorithms (which in many cases reduce to certain forms of linear programming) or tables. Thus the book does not contain individual statistical recipes for problems with nuisance parameters, but rather attempts to point out procedured for constructing such recipes.

In the introduction and later, we recall certain standard theorems on σ-algebras, probabilistic measures, and statistics. In Chapter I, we treat multiple Laplace transforms and describe the simpler properties and applications of analytic sheaves along the lines developed by H. Cartan. In later chapters these properties will be applied to the theory of exponential families. Chapter II gives the fundamentals of the theory of sufficient statistics for distributions in Euclidean spaces and exponential families associated with them (for repeated samples). Chapter III presents some of the problems themselves with nuisance parameters. Chapter IV treats the theory of similarity following J. Neyman, E. Lehmann, and H. Scheffé. Chapters V and VII–X discuss the recent researches of statisticians at Leningrad University in the theory of similar tests and unbiased estimates, particularly in connection with the Behrens-Fisher problem. In Chapter VI, an exposition is given of the remarkable method of R. A. Wijsman; however, this method does not yield all desirable tests. The role of the theory of sheaves of ideals of functions as an analytic foundation of the theory of similar tests and unbiased estimates for imcomplete exponential families is clarified in Chapters V and VII. Here exponential families are considered not only for repeated samples but for other cases as well.

In Chapter XI, an exposition is given of the problem of many small samples, and, in particular, of the researches of A. A. Petrov.

At the end of the book are several unsolved problems, which constitute only a small portion of the esthetically pleasing and varied problems that arise in analytical statistics. Our purpose of the present book is to draw the attention of persons interested in mathematical statistics to its analytical aspects.

A. M. Kagan, I. L. Romanovskaja, and V. N. Sudakov had a share in the writing of this book. Sections 2 and 3 of Chapter VII were written by the author in collaboration with A. M. Kagan, and section 4 of Chapter VIII with I. L. Romanovskaja. Section 2 of Chapter X was written by V. N. Sudakov. A considerable amount of help in the writing of Chapter I was provided by N. M. Mitrofanova and V. L. Eĭdlin.

I wish to express my gratitude to O. I. Rumjanceva and S. I. Čirkunova for their great help in the preparation of the manuscript.

Ju. V. Linnik

PREFACE TO THE AMERICAN EDITION

The American translation of this book takes account of several corrections of misprints and author's errors that were noticed by readers or by the author. It also contains a supplement to the book written by A. M. Kagan and V. P. Palamodov, expounding their important contributions published recently in "Teorija Verojatnostei i ee Primenenija". The answers to several questions raised at the end of the book are provided by the supplement. This new material includes a considerable advance in the theory of nonsequentially verifiable functions, the construction of all randomized similar tests for the Behrens-Fisher problem, and important progress in the estimation theory for incomplete exponential families, based on the introduction into statistics of the elements of homological algebra (in particular, flat modules).

The analytical sheaf theorems on which a large part of the book is based are replaced in the supplement by the Hörmander-Malgrange theory of linear differential operators with constant coefficients. This theory enables us to solve problems involving convex supports, rather than merely the polygonal ones discussed in the book. Thus we can now construct all similar tests for a linear hypothesis with unknown variances (least square method with unknown observation weights) and for many other problems of testing hypotheses and unbiased estimation. The optimization problems are thus reduced to purely analytic (variational) ones.

I am very grateful to the American Mathematical Society for publishing a translation of my book with the supplement. It is my pleasant duty to thank S. H. Gould and G. L. Walker for their interest in my book.

Ju. V. Linnik

This page intentionally left blank

TABLE OF CONTENTS

This page intentionally left blank

APPENDIX

UNSOLVED QUESTIONS

In the analytical formulation of certain problems presented in this book dealing with the elimination of nuisance parameters in statistical problems, certain questions and unsolved problems arise rather naturally. Here we arrange these according to the chapters dealing with the topic in question.

Chapter II. 1. Continuation of the work of Koopman, Dynkin and Brown on the classification of distributions admitting sufficient statistics of finite rank for the case in which the distribution densities can vanish. Weakening of the condition for existence of exponential families in the case when the densities do not vanish on the sample space.

2. Investigation of exponential families with "sliding carrier", i.e. families whose carrier depends on the parameters.

Chapter III. 1. Consider a repeated normal sample $x_1, \cdots, x_n$, where $x_i \in N(a, \sigma^2)$. Let us partition all almost-everywhere-continuous functions $\gamma(a, \sigma)$ into two classes:

I. "Verifiable functions" for which the hypothesis H_0: $\gamma(a, \sigma) = \gamma_0$ admits an invariant verification.

II. The remaining functions.

How can one describe the class I?

2. The same question for an infinite sample $x_1, x_2, \cdots$, and the application of sequential analysis, where the mean number of steps is bounded up to the final solution.

3. Generalization of question 1. Suppose that we have an exponential family of the form (5.2.3). Let us partition the collection of functions $\gamma(\theta_1, \cdots, \theta_s)$ into "verifiable" and "unverifiable" classes, just as in question 1. How can we describe the class of the form I?

Chapter IV. 1. How can we describe all the similar zones of distribution families of the form (4.3.5)?

2. Let $x_1, \cdots, x_n$ ($x_i \in N(0, 1)$) denote a repeated normal sample. Let $P_1(x_1, \cdots, x_n)$ and $P_2(x_1, \cdots, x_n)$ denote two independent polynomial statistics. Is it always possible to "uncouple" them by means of an orthogonal transformation, i.e. to reduce them to two statistics depending on the completeness of the different variables? (For a given sample size n and given degrees m_1 and m_2 of the polynomials the question can be solved in a finite number of steps for all such polynomials (see [41]).)

Chapter V. 1. Is it possible to weaken the condition for complexification of the relations (condition (Y_{I}) in §8)?

2. Is it possible to weaken the condition on the Jacobians (Y_{II})?

3. How can one describe in terms of generalized Laplace transformations (in the sense of the theory of generalized functions) the construction of nonsmooth cotests?

These questions are related to the following question in the theory of analytic functions.

4. Let Z denote a simply-connected complex polycylinder. Let O' denote a ring of functions that are holomorphic and that have a holomorphic continuation from the polycylinder to the Cartesian product of the half-planes containing Z. How can we describe the structure of the ideals of the ring O'? Under what conditions will the basis of the ideals be finite?

5. How can we construct cotests for the case in which the function h vanishes in a given region?

Formula (5.8.6) leads to the following analytic problem: Let $A_j(T_1, \cdots, T_s)$, $j = 1, 2, \cdots, r$; $r < s$, denote sufficiently smooth functions defined in the Euclidean space E_s of the arguments $T_1, \cdots, T_s$. Let U denote a simply-connected region contained in E_s and bounded by smooth surfaces. How can we describe the system of all functions $H_1, \cdots, H_r$ for which the sum of the convolutions

$$A_1 * H_1 + \cdots + A_r * H_r$$

exists and vanishes on U?

6. Is it possible to describe smooth cotests for exponential families with "sliding carrier" (cf. question 2 to Chapter II)?

7. The development of computational methods, in particular linear programming

methods, for finding optimal similar tests.

Chapter VI. 1. Generalization of Wijsman's results on the construction of all similar tests of the hypothesis H_0: $a/\sigma = \gamma_0$ for a repeated normal sample $x_1, \cdots$ $\cdots, x_n$ where $x_j \in N(a, \sigma^2)$. In what problems associated with tests of hypotheses on normal samples do parabolic differential equations appear? What can we do in other cases?

Chapter VII. 1. Questions of unbiased estimates of zero (UEZ's) analogous to the questions in Chapter V on cotests, i.e. questions 1–5 of Chapter V with the word "cotest" replaced by the expression "unbiased estimate of zero".

2. Generalization of the results of §2 dealing with inadmissible unbiased estimates. Ways of obtaining more general forms of inadmissible unbiased estimates for incomplete exponential families on a compact set of values of the parameters. Cases of inadmissibility for noncompact sets of values of the parameters.

3. Replacement of the variance as a function of the loss of an unbiased estimate with other sufficiently smooth functions. Conditions for inadmissibility of unbiased estimates.

4. Classification of the best unbiased estimates for incomplete exponential families from a standpoint of variance. From an analytic point of view this problem is connected with the following question from functional algebra. Let O denote the ring of all holomorphic functions defined on a compact simply-connected polycylinder and let I denote an ideal contained in O. Find the ring K of all linear differential operators L with constant coefficients such that $LI \subset I$.

5. Problems analogous to the preceding ones for the case in which we are using not the variance but other sufficiently smooth functions of the loss.

6. Construction of an analogue to Theorem 7.3.1 dealing with inadmissibility of a sample mean for scale parameters. Extension of Theorem 7.3.1 to observations connected with a homogeneous Markov chain.

Chapter VIII. 1. Investigation of the question of nonexistence, in general, for incomplete exponential families of similar zones that depend on sufficient statistics with sufficiently smooth boundaries.

2. Weakening of the conditions of Theorem 8.3.1. Does there exist a similar Fisher-Welch-Wald test if we require only continuity of the test boundary or only satisfaction of a Lipschitz condition for it?

Chapter IX. 1. Can we construct a test analogous to the simple randomized test in §3 for the case of nonidentical sample sizes?

2. Generalization of Theorem 9.4.1 on the characterization of the Bartlett-Scheffé test. Weakening of the conditions on the basic space of linear forms. Construction of an analogue of this theorem to characterize the simplest tests of the method of least squares with unknown weights.

Chapter X. 1. How can one construct a homogeneous unrandomized similar test for the Behrens-Fisher problem in the case of like parity of the sizes of the samples?

Chapter XI. 1. Investigations analogous to those made by Petrov for the scheme of many small samples from complete exponential families referred to in §1.

SUPPLEMENT

NEW RESULTS IN THE THEORY OF ESTIMATION AND TESTING HYPOTHESES FOR PROBLEMS WITH NUISANCE PARAMETERS

A. M. KAGAN AND V. P. PALAMODOV[1)]

The results expounded in this Supplement are mostly answers to questions put at the end of the book (see queries: 1 to Chapter III, 3 and 5 to Chapter V, 4 and 6 to Chapter VII). But §5 is an exception; there the problem of estimating a location parameter is considered when the "nuisance parameter" is the form of the function $F(x-\theta)$ itself, for which only several first distribution moments for $\theta = 0$ are known.

The results of §§1 and 2 are due to V. P. Palamodov [6, 7],[2)] of §3 to A. M. Kagan and V. P. Palamodov [2, 3], of §4 to A. M. Kagan [4, 12], of §5 to A. M. Kagan and A. L. Ruhin [5].

§1. INVARIANT VERIFICATION OF FUNCTIONS WHICH ARE POLYNOMIALS IN a AND $1/\sigma^2$ FOR NORMAL SAMPLES

This section gives the description of all polynomials $P(a, 1/\sigma^2)$ admitting invariant verification in the sense of §2, Chapter III (strictly speaking, that sense will be modified a little), on the evidence of a repeated sample $(x_1, \cdots, x_n)$ from a normal population $N(a, \sigma^2)$. The method of this section is purely analytical, in contrast, for instance, to that of E. Lehmann [13] establishing the non-verifiability of certain functions as a consequence of the indistinguishability of the corresponding families of distributions.

1) Editor's note. The translation of the Supplement was provided by the authors.

2) Authors' note. These refer to the Bibliography at the end of this Supplement, not to the main Bibliography.

Let $x_1, \cdots, x_n \in N(a, \sigma^2)$ be independent normal variables. In studying the question of the invariant verification we can restrict ourselves, without loss of generality, to those tests $\phi(\bar{x}, s)$ depending only upon the sufficient statistics

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i; \quad s^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2.$$

We have

$$E_{a,\sigma}\phi(\bar{x}, s) = \bar{\phi}(a, \sigma^2) = \check{\phi}(a, \xi)$$

$$= C_0 \int_{-\infty}^{\infty} \int_0^{\infty} \xi^{n/2} \exp[-\xi(s + (\bar{x} - a)^2)] s^{(n-3)/2} \phi(\bar{x}, s)\, d\bar{x}\, ds. \qquad \text{(S.1.1)}$$

Here we denoted $\xi = 1/\sigma^2$; C_0 is a constant.

Note that the integral (S.1.1) can be continued as an analytic function of two complex variables to the product of the complex plane C of the values of a and the complex halfplane C^+ of the values of ξ with $\operatorname{Re} \xi > 0$. This enables us to introduce the following definition.

Definition. A function $f(a, \xi)$ defined in $C \times C^+$ is called C-verifiable if there exists a test ϕ such that

$$\check{\phi}(a, \xi) = \psi(f(a, \xi)); \quad a \in C, \quad \xi \in C^+$$

for some $\psi \not\equiv \text{const}$.

Any real C-verifiable function is obviously verifiable in the sense of §2, Chapter III. The converse is not true; in [6] we give sufficient conditions for the C-verifiability of functions which are verifiable in the sense of §2, Chapter III.

Lemma S.1.1. *For any test* $\phi(\bar{x}, s)$

$$|\check{\phi}(a, \xi)| \le C_0 \cdot \left|\frac{\xi}{\operatorname{Re} \xi}\right|^{n/2} \exp\left[\frac{|\xi|^2}{\operatorname{Re} \xi} |\operatorname{Im} a|^2\right] \qquad \text{(S.1.2)}$$

for $a \in C$, $\xi \in C^+$ ($C_0, C_1, \cdots$ *in what follows are positive constants*).

Proof.

$$|\check{\phi}(a, \xi)| = C_0 |\xi|^{n/2} \times$$

$$\times \exp(-\operatorname{Re}(\xi a^2)) \int_{-\infty}^{\infty} \int_0^{\infty} \exp[-\xi(s+\bar{x}^2) + 2\xi a \bar{x}] s^{(n-3)/2} \phi(\bar{x}, s)\, ds\, d\bar{x}$$

$$\leq C_0 |\xi|^{n/2} \exp(-\operatorname{Re}(\xi a^2)) \int_{-\infty}^{\infty} \int_0^{\infty} \exp[-\operatorname{Re}\xi(s+\bar{x}^2) + 2\operatorname{Re}(\xi a)\bar{x}] s^{(n-3)/2}\, ds\, d\bar{x} \tag{S.1.3}$$

because $0 \leq \phi(\bar{x}, s) \leq 1$. The quantity

$$C_0 |\operatorname{Re}\xi|^{n/2} \cdot \exp\left[-\frac{\operatorname{Re}(\xi a)^2}{\operatorname{Re}\xi}\right] \int_{-\infty}^{\infty} \int_0^{\infty} \exp[-\operatorname{Re}\xi(s+\bar{x}^2) + 2\operatorname{Re}(\xi a)\bar{x}] s^{(n-3)/2}\, ds\, d\bar{x}$$

is the power function of the trivial test $\phi \equiv 1$ at the point $(\operatorname{Re}\xi, \operatorname{Re}(\xi a)/\operatorname{Re}\xi)$. Taking this into account, we can write the right side of (S.1.3) in the form

$$C_0 \left|\frac{\xi}{\operatorname{Re}\xi}\right|^{n/2} \exp\left[\frac{\operatorname{Re}(\xi a)^2}{\operatorname{Re}\xi} - \operatorname{Re}(\xi a^2)\right].$$

Putting $a = \alpha + i\beta$, $\xi = \zeta + i\eta$ we get

$$\frac{\operatorname{Re}(\xi a)}{\operatorname{Re}\xi} - \operatorname{Re}\xi a^2$$

$$= \frac{1}{\zeta}[(\alpha\zeta - \beta\eta)^2 - \zeta(\zeta(\alpha^2\beta^2) - 2\eta\alpha\beta)] = \frac{1}{\zeta}(\zeta^2 + \eta^2)\beta^2 = \frac{|\xi|^2}{\operatorname{Re}\xi}|\operatorname{Im} a|^2,$$

which leads to (S.1.2).

We shall say that $r(a, \xi) \in R$ if

$$r(a, \xi) = \frac{p(a, \xi)}{q(a, \xi)} = \frac{p_m(\xi) a^m + \cdots + p_0(\xi)}{q_k(\xi) a^k + \cdots + q_0(\xi)} \tag{S.1.4}$$

where p_i, q_j $(i = 0, 1, \cdots, m;\ j = 0, \cdots, k)$ belong to the ring A of the functions holomorphic in $\mathbf{C}^+$ and $p_m q_k \neq 0$. Without loss of generality we suppose that either $m > k$ or $m = k$ but $p_m(\xi)/q_k(\xi) \neq \text{const}$. Moreover, we can suppose that the number m is the least possible and the functions p_i and q_i do not vanish simultaneously.

Theorem S.1.1. 1°. *If a function* $r \in R$ *with* $m > 0$ *is* C*-verifiable, then* $k = 0$ *in* (S.1.4), *and the function* $q_0(\xi) \neq 0$ *on* $\mathbf{C}^+$. *Moreover, the power function of the test* $\psi(r(a, \xi))$ *is an entire function of order not exceeding* $2/m$.

2°. *If, moreover, the functions* $q_0, p_0, \cdots, p_m$ *and* p_0/q_0 *are analytic in the vicinity of zero, the function is of the form*

$$r(a, \xi) = r_2(\xi)a^2 + r_1(\xi)a + r_0(\xi);\ r_i = \frac{p_i}{q_0} \in A$$

times a constant; also, $r_2(0) = r_1(0) = 0$; $r_2'(0) = 1$ *and the image of the mapping* $r_2 : \mathrm{C}^+ \to \mathrm{C}$ *belongs to* C^+. *The function* ψ *is bounded in any closed angle belonging to* C^+.

In what follows we shall denote by C^* the complex plane compactified in the usual way.

Lemma S.1.2. *Let* $\rho: \omega \to \mathrm{C}^*$ *be a function meromorphic in a domain* $\omega \subset \mathrm{C}$ *and not a constant; let* Ω *be the range of its values. If* $\psi(\rho(\zeta))$, $\zeta \in \omega$ *is analytic in* ω, *then* ψ *is analytic in* Ω.

Proof. Let $z_0 \in \Omega$ be an arbitrary point distinct from ∞, and let $\zeta_0 \in \omega$ be a point at which $\rho(\zeta_0) = z$. If $\rho'(\zeta_0) \neq 0$, then in the neighborhood of the point z_0 we have $z = \rho(w(z))$ where $w(z)$ is holomorphic in z_0. Hence $\psi(z) = \psi(\rho(w(z)))$ which implies that ψ is holomorphic in z_0.

Since $\rho \not\equiv \text{const}$, the zeros of its derivative are isolated. Let ζ_0 be one of the zeros of ρ' and $U \subset \omega$ a closed bounded neighborhood of the point ζ_0, containing no other zeros of ρ' and no poles of ρ. Its image $V = \rho(U)$ is a bounded neighborhood of the point $z_0 = \rho(\zeta_0)$.

In the domain $V\setminus\{z_0\}$ the function ψ is bounded and was proved to be analytic. Hence it is analytic at the point z_0 also.

Now let ζ_0 be a pole of the function ρ. Choose a bounded closed neighborhood $U \subset \omega$ of the point ζ_0. Its image V is a neighborhood of ∞, and in the domain $V\setminus\{z_0\}$ the function ψ is bounded and analytic. Hence it is analytic at the point z_0 also. The lemma is proved.

We pass now to the proof of Theorem S.1.1. We suppose that $\psi \not\equiv \text{const}$.

It is easy to see that the range of values of the function $r: \mathrm{C} \times \mathrm{C}^+ \to \mathrm{C}^*$ always contains C; hence by Lemma S.1.2 the function ψ is entire.

If $k > 0$, then for a suitable ξ_0 the image of the mapping $r(a, \xi_0): \mathrm{C} \to \mathrm{C}^*$ contains the point ∞. Hence by Lemma S.1.2 the function ψ is analytic at that point. Since it is entire it must be a constant, which contradicts our assumptions. Hence $k = 0$. In a similar way we prove that q_0 has no zeros in C^+. Hence we get

$$r(a, \xi) = r_m(\xi) a^m + \cdots + r_0(\xi); \quad r_i = \frac{p_i}{q_0} \in A.$$

We fix the point ξ so that $r_m(\xi) \neq 0$. Then $r(a, \xi) \sim r_m(\xi) a^m$ for $a \to \infty$. The inequality (S. 1.2) implies for certain values of the constants A and B that

$$|\psi(r(a, \xi))| \leq C_0 \exp(A |\mathrm{Im}\, a|^2) \leq C_0 \exp(B\, |r(a, \xi)|^{2/m});$$

hence we deduce that the order of the entire function ψ does not exceed $2/m$. The first assertion of Theorem S. 1.1 is proved.

We pass to the proof of the second assertion. The functions $r_0, p_1, \cdots, p_m, q_0$ are analytic in the vicinity of zero. Hence the functions $r_i = p_i/q_0$ are analytic in the vicinity of zero, except perhaps at zero itself, where the only possible singularity is a pole. We shall show that in fact the coefficients r_i $(i = 1, \cdots, m)$ are analytic in the vicinity of zero and $r_i(0) = 0$. For each $i = 1, \cdots, m$ we have in the vicinity of zero

$$r_i(\xi) = \rho_i \xi^{\alpha_i} + \rho_i' \xi^{\alpha_i + 1} + \cdots,$$

with certain $\rho_i \neq 0$ and α_i. If all $\alpha_i > 0$ our assertion is proved. Suppose that $\alpha_i < 0$ for a certain i; consider the quantity

$$\alpha = \max_{i \geq 1} \left\{ -\frac{\alpha_i}{i} \right\}.$$

Let k be the largest number for which $-\alpha_k/k = \alpha$. For $\lambda \in \mathbf{C}$ put

$$a_\lambda(\xi) = \left[\frac{\lambda}{r_m(\xi)}\right]^{1/k}, \quad \xi > 0,$$

where the branch of the root can be chosen arbitrarily. Since $\alpha_k < 0$, the quantity $a_\lambda(\xi)$ is bounded for $\xi \to 0$ for any $\lambda \in \mathbf{C}$. Since $\xi > 0$, we get from (S. 1.2):

$$|\psi(r(a_\lambda(\xi), \xi))| \leq C \exp(B(\lambda)\, \xi). \tag{S. 1.5}$$

On the other hand,

$$r(a_\lambda(\xi), \xi) = \sum_{j>k} r_j(\xi) a_\lambda^j(\xi) + \lambda + \sum_{j<k} r_j(\xi) a_\lambda^j(\xi). \tag{S. 1.6}$$

In the first sum on the right-hand side of (S. 1.6)

$$|r_j(\xi)a_\lambda^j(\xi)| = \left\{\frac{|r_j(\xi)|^{1/j}}{|r_k(\xi)|^{1/k}}\right\} |\lambda|^{j/k}$$

$$\sim C\,\xi^{j(\alpha_j/j - \alpha_k/k)} \to 0$$

for $\xi \to 0$, in view of the choice of k. In the second sum

$$|r_j(\xi)a_\lambda^j(\xi)| \le C\xi^{j(\alpha_j/j - \alpha_k/k)}|\lambda|^{j/k} \le C(|\lambda|^{(k-1)/k} + 1)$$

for $0 < \xi < 1$, since $\alpha_j/j \ge \alpha_k/k$ for $j < k$ and the function $r_0(\xi)$ is bounded for $\xi \to 0$. We choose ξ so small that the first sum in (S.1.6) is smaller than 1. Then

$$|r(a_\lambda(\xi), \xi) - \lambda| \le C(|\lambda|^{(k-1)/k} + 1). \tag{S.1.7}$$

We take now a sufficiently large R. If the point λ runs over the circumference with the radius R, (S.1.7) implies that the point $r(a_\lambda(\xi), \xi)$, which depends continuously on λ, runs over a curve homeomorphic to that circumference and containing the circle

$$|z| \le R' = R - C(R^{(k-1)/k} + 1).$$

On the other hand, from (S.1.5) it follows that on this curve the function ψ is bounded by the constant $C\exp(B(\lambda)\xi)$. Since ξ can be chosen as small as we please, ψ is bounded by the constant C, which does not depend upon λ on the curve described above and hence in the circle of radius R'. As $R \to \infty$ so does R'; hence ψ is bounded on the whole plane; hence $\psi \equiv$ const. Since we assumed that $\psi \not\equiv$ const we see that all $\alpha_i > 0$.

Lemma S.1.3. *For each $\xi \in \mathbf{C}^+$ there exists an $\epsilon > 0$ such that the function ψ is bounded inside the angle*

$$|\operatorname{arc} z - \operatorname{arc} r_m(\xi)| \le \epsilon.$$

and, for odd m, also inside the angle

$$|\operatorname{arc} z - \operatorname{arc}(-r_m(\xi))| \le \epsilon. \tag{S.1.8}$$

Proof. Take an arbitrary $\xi \in \mathbf{C}^+$. Since $r_m(0) = 0$, it follows that $r_m(\xi) \not\equiv$ const. Therefore in the vicinity of the point ξ we can find points ξ_+ and ξ_- such that

$$\operatorname{arc} r_m(\xi_+) > \operatorname{arc} r_m(\xi) > \operatorname{arc} r_m(\xi_-),$$
$$\operatorname{arc} r_m(\xi_+) - \operatorname{arc} r_m(\xi_-) = \delta < \pi/2. \tag{S.1.9}$$

We shall show now that the function ψ is bounded inside the angle

$$\operatorname{arc} r_m(\xi_+) - \delta/3 \geq \operatorname{arc} z \geq \operatorname{arc} r_m(\xi_-) + \delta/3. \tag{S.1.10}$$

Consider the curves

$$\theta_\pm(a) = r(a, \xi_\pm), \quad \xi \in [0, \infty). \tag{S.1.11}$$

We have

$$\theta_\pm(a) - r_m(\xi_\pm)a^m = O(a^{m-1}); \quad \xi \to \infty.$$

This implies that the domain between the curves $\xi_\pm(a)$, taken together with a sufficiently large circle, contains the angle (S.1.10). By the inequality (S.1.2) the function $\psi(r(a, \xi_\pm))$ is bounded for real values of a. Hence the function ψ is bounded on the curves (S.1.11). But by (S.1.9) these curves are contained inside an angle less than $\pi/2$, while ψ was shown above to be an entire function of order not exceeding $2/m \leq 2$. Hence by the Phragmén-Lindelöf principle, the function ψ is bounded in the domain lying between the curves (S.1.11), as was to be proved. The case of odd m is dealt with by analogy with the preceding one.* To complete the proof of Theorem S.1.1, consider the set

$$\{\operatorname{arc} r_m(\xi), \ \xi \in C^+\}. \tag{S.1.12}$$

Since

$$r_m(\xi) = \rho_m \xi^{\alpha_m} + \rho'_m \xi^{\alpha_m + 1} + \cdots; \quad \alpha_m > 0$$

for sufficiently small ξ's, we have

$$\operatorname{arc} r_m(\xi) \sim \alpha_m \operatorname{arc} \xi + \operatorname{arc} \rho_m \quad \text{for } \xi \to 0. \tag{S.1.13}$$

Therefore if $\alpha_m > 1$, the set (S.1.12) contains an interval of length 2π. By Lemma S.1.3 it then follows that the function ψ is bounded inside the angle $2\pi - \epsilon$ for any $\epsilon > 0$. As ψ is an entire function of a finite order,. we have by the Phragmén-Lindelöf principle: $\psi = \text{const}$. Hence $\alpha_m = 1$. Multiplying $r(a, \xi)$ into $1/p_m$, we get $r'_m(0) = 1$.

---

*With ξ replaced by $-\xi$.

In this case it follows from (S. 1.13) that the set (S. 1.2) contains the interval $(-\pi/2, \pi/2)$. Then by Lemma S. 1.3 the function ψ is bounded inside the angle $|\text{arc } z| \leq \pi/2 - \epsilon$ for any $\epsilon > 0$. Hence ψ is a function of order at least 1 and of finite type. On the other hand, we have proved that its order does not exceed $2/m$. Hence, $m \leq 2$. The case $m = 1$ is excluded because for $m = 1$ Lemma S. 1.3 would imply the boundedness of the function ψ also in the angle $|\text{arc } z - \pi| \leq \pi/2 - \epsilon$ for any $\epsilon > 0$, which is impossible for nonconstant functions of finite order.

It remains only to verify that the image of the mapping r_2: $C^+ \to C$ belongs to C^+. Suppose it is not so, and that for a certain $\xi \in C^+$, $|\text{arc } r_2(\xi)| > \pi/2$. By Lemma S. 1.3 the function ψ is bounded inside the angle $|\text{arc } z - \text{arc } r_2(\xi)| \leq \epsilon$. Since this function is of the first order and bounded inside any angle of the form $|\text{arc } z| \leq \pi/2 - \epsilon$, again the Phragmén-Lindelöf principle implies that it vanishes identically. Theorem S. 1.1 is proved.

Theorem S. 1.2. *Let a polynomial* $p(a, \xi)$ *which is not a function of* ξ *only be* C*-verifiable. To be so, it is necessary and sufficient for* $p(a, \xi)$ *to be representable as a linear form of*

$$\xi(a^2 + Aa + B) \tag{S. 1.14}$$

where A *and* B *are real and* $A^2 - 4B \leq 0$.

Proof. Necessity. Let $p(a, \xi)$ be a C-verifiable polynomial, and let the test ϕ be such that $\check{\phi}(a, \xi) = \psi(p(a, \xi))$ where $\psi \not\equiv \text{const}$. By Theorem S. 1.1 the function ψ is an entire one and $p(a, \xi)$, after multiplication by a suitable constant, is of the form

$$p(a, \xi) = p_2(\xi)a^2 + p_1(\xi)a + p_0(\xi),$$

where

$$p_2(0) = p_1(0) = 0; \ p_2'(0) = 1. \tag{S. 1.15}$$

By the conditions of the theorem $p_2(\xi)$, $p_1(\xi)$, $p_0(\xi)$ are polynomials. We shall show that their order does not exceed 1. Suppose this is not true; then for a certain $a = a_0$ the polynomial $p(a_0, \xi)$ is of order $k > 1$ as a polynomial in ξ. Hence

$$p(a_0, \xi) = C\xi^k + O(|\xi|^{k-1}); \ \xi \to \infty. \tag{S. 1.16}$$

Let the point ξ move inside the angle $|\text{arc}\,\xi| \leq \pi/2 - \epsilon$. Then from (S. 1.2) it follows that $\psi(p(a_0, \xi))$ is bounded. On the other hand, the first term in (S. 1.16) runs over all the values of the angle $|\text{arc}\,\xi| \leq \pi - 2\epsilon$. Hence the function $p(a_0, \xi)$ runs over all the values of the complex plane, except perhaps in a certain circle and in the angle $|\text{arc}\,z - \pi| \leq 3\epsilon$. Hence the entire function ψ is bounded on the whole plane except perhaps in an angle $|\text{arc}\,z - \pi| \leq 3\epsilon$, which is as small as we please. Since this function is of finite order, the Phragmén Lindelöf principle implies that it is bounded on the whole plane and therefore is a constant. This contradiction proves that $p_2(\xi)$, $p_1(\xi)$ and $p_0(\xi)$ are linear functions of ξ.

From (S. 1.15) we get that $p_2(\xi) = \xi$; $p_1(\xi) = A\xi$, where A is a constant. Subtracting a suitable constant from $p(a, \xi)$ we can also annul the constant term of $p_0(\xi)$. Then $p(a, \xi)$ takes the form (S. 1.14).

Consider now the function $\text{arc}\,(a^2 + Aa + B)$ for real values of a. Suppose it to be nonconstant and let ϕ_1 and ϕ_2 be two distinct values of it. In that case, if ξ runs over the angle $|\text{arc}\,\xi| \leq \pi/2 - \epsilon$, and the point a runs over the real axis, the point $\xi(a^2 + Aa + B)$ will take on all the values inside the angles

$$|\text{arc}\,z - \phi_1| \leq \pi/2 - \epsilon \text{ and } |\text{arc}\,z - \phi_2| \leq \pi/2 - \epsilon.$$

From the inequality (S. 1.2) it follows that ψ is bounded inside these angles for any $\epsilon > 0$. Since $\phi_1 \neq \phi_2$ and ψ is an entire function of the first order, this is impossible in view of the Phragmén-Lindelöf principle. Hence $\text{arc}\,(a^2 + Aa + B) = \text{const}$. On the other hand, $\text{arc}\,(a^2 + Aa + B) = 0$ for $a \to \infty$. Hence $\text{arc}(a^2 + Aa + B) = 0$, A and B are real numbers and $a^2 + Aa + B$ is nonnegative for all real a, so that $A^2 - 4B \leq 0$.

Sufficiency. Consider the test

$$\phi(\bar{x}, s) = \begin{cases} (1 - t(\bar{x}^2/s))^{(n-3)/2}; & s \geq t\bar{x}^2 \\ 0 & ; s < t\bar{x}^2 \end{cases}$$

for a given $t > 0$. It can be shown that for a certain pair of numbers (y, s_0) with $s_0 > 0$ we have

$$E_{a,\xi}\,\phi(\bar{x} - y, s - s_0) = C_0 \exp(-r\xi^2(a^2 + Aa + B))$$

where $\tau \in (0, 1)$, so that $\xi(a^2 + Aa + B)$ is verifiable. We omit the details; they can be found in [6].

§2. THE DESCRIPTION OF ALL COTESTS FOR A CLASS OF EXPONENTIAL FAMILIES WITH POLYNOMIAL RELATIONS

This is closely related to §8, Chapter V, and we use here the notation and terminology of that chapter.

Consider the exponential family of densities with respect to Lebesgue measure

$$p(T, \theta) = C(\theta) h(T) \exp(\theta_1 T_1 + \cdots + \theta_s T_s) \tag{S.2.1}$$

with θ, $T \in R^s$. Before stating the conditions imposed upon $h(T)$ and the parametric set, we introduce the set $\omega \subset R^s$ determined in the following way. Construct the set of $\theta \in R^s$ for which for a certain $C = C(\theta)$ the condition

$$\{(\theta, T) < C\} \supset \operatorname{Supp} h = \mathfrak{T}$$

holds.

This set is a cone whose interior we denote by ω. It is important to remark that if $\theta \in \omega$ we can choose a constant B such that for a suitable $\epsilon > 0$

$$(\theta, T) \leq -\epsilon |T| + B; \quad T \in \mathfrak{T}. \tag{S.2.2}$$

We shall suppose that the family (S.2.1) satisfies the following requirements.

1.1. $\mathfrak{T} = \operatorname{Supp} h$ is a convex set.

1.2. ω is nonvoid.

1.3. For any $\epsilon > 0$

$$\| h(T) \exp(-\epsilon |T|) \|_{L_2} < \infty. \tag{S.2.3}$$

(Note that from (S.2.2) and (S.2.3) it follows that for any $\theta \in \omega$, $p(T, \theta)$ can be considered as a probability density.)

2.1. The parameter θ takes on values in an everywhere dense subvariety of a real algebraic variety $\Pi \cap \omega$.

2.2. The polynomial ideal I formed by the real polynomials of θ vanishing on $\Pi \cap \omega$ is a principal one, i.e. it consists of all polynomials of the form

$$G(\theta) P(\theta)$$

for a suitable $P(\theta)$.

Denote by N the variety of complex roots of $P(\theta)$ and by $\omega \times R_y^s$ the set $(\theta \in C^s;\ \operatorname{Re}\theta \in \omega)$.

3.1. Each connected component of the intersection $N \cap (\omega \times R_y^s)$ has at least one real point where $\operatorname{grad} P(\theta) \neq 0$.

Theorem S.2.1. *Under Conditions* 1.1–3.1 *each cotest can be represented uniquely by the formula*

$$\phi(t) = \frac{1}{h(T)} P(-\mathfrak{D})\Psi(T);\ \mathfrak{D} = \left[\frac{\partial}{\partial T_1}, \dots, \frac{\partial}{\partial T_s}\right] \tag{S.2.4}$$

where $\Psi(T)$ *is an (ordinary) function such that*

$$\left.\begin{array}{c}\operatorname{Supp}\Psi \subset \operatorname{Supp} h \\ \|\Psi \exp(-\epsilon|T|)\|_{L_2} < \infty \ \textit{for any}\ \epsilon > 0\end{array}\right\}. \tag{S.2.5}$$

The function Ψ *is uniquely determined.*

Conversely, each function of the type (S.2.4) *with* $\Psi(T)$ *satisfying the requirements* (S.2.5) *and lying between* α *and* $1-\alpha$ *for a value of* $\alpha \in (0, 1)$, *is a cotest.*

Proof. 1. Take a point $a \in \omega$. Then for certain constants $C > 0$ and B the condition

$$(a, T) \leq -C|T| + B \tag{S.2.6}$$

is fulfilled on the set $\operatorname{Supp} h$.

2. Set

$$\|\Phi\|_\epsilon = \|\Phi(T)\exp\epsilon|T|\,\|_{L_2}.$$

By a theorem of Hörmander [11], for any differential operator with constant coefficients there is a fundamental solution (in general, a generalized function) $E(T)$ under the condition:

$$\|E * \Phi\|_{-\epsilon} \leq C\|\Phi\|_\epsilon \tag{S.2.7}$$

for any $\Phi(T)$ for which the right-hand side of the formula is finite. Let $E(T)$ be such a solution for $P(-\mathfrak{D} + a)$.

Let $\phi(T)$ be a given cotest, put $\check{\phi} = \phi h$ and

$$\Psi(\tau) = E(T) \exp[-(a, T)] * \check{\phi}(T)$$
$$= \int E(T) \exp[-(a, T)] \check{\phi}(\tau - T)\, dT$$
$$= \exp[-(a, \tau)] \int E(T) \exp(a, \tau - T) \check{\phi}(\tau - T)\, dT.$$

In view of the inequality (S. 2.6), the function $\exp(a, T)\check{\phi}(T)$ belongs to the space L_2 with the weight $\exp(\epsilon|T|)$ for a certain $\epsilon > 0$. Therefore the property (S. 2.7) of the function E implies the inequality

$$\|\exp(a, \tau)\Psi(\tau)\|_{-\epsilon}$$
$$= \|E * \exp(a, T)\check{\phi}(T)\|_{-\epsilon} \leq C\|\exp(a, T)\check{\phi}(T)\|_{\epsilon}$$
$$\leq C'\|\check{\phi}\|_{-\epsilon} \text{ for all small } \epsilon > 0. \tag{S. 2.8}$$

Let us check the equality (S. 2.1). To this end we shall establish first that the function $E(T)\exp[-(a, T)]$ is the fundamental solution for the operator $p(-\mathfrak{D})$. By the Leibnitz differentiation formula we have

$$P(-\mathfrak{D})\{E(T)\exp(-(a, T))\}$$
$$= \exp(-(a, T))\sum_i \frac{a^i}{i!} P^{(i)}(-\mathfrak{D})E$$
$$= \exp(-(a, T))P(-\mathfrak{D} + a)E = \exp-(a, T)\cdot\delta = \delta,$$

from which we easily obtain

$$P(-\mathfrak{D})\Psi = P(-\mathfrak{D})(E(T)\exp(-(a, T)) * \check{\phi}) = \delta * \check{\phi} = \check{\phi}.$$

3. We must prove now that $\operatorname{Supp}\Psi \subset \operatorname{Supp} h$. Let $\tau \overline{\in} \operatorname{Supp} h$; then from the convexity of $\operatorname{Supp} h$ it follows that for a certain λ we have $(\lambda, T) > (\lambda, \tau)$ for $T \in \operatorname{Supp} h$. The construction of $E(T)$ leads to the relation

$$P(\mathfrak{D}_T + a)E(\tau - T) = 0 \text{ in the half-space } (\lambda, \tau - T) < 0.$$

By a theorem of V. P. Palamadov [8, 9] we have the representation

$$E(\tau - T) = \int_{\substack{P(\theta + a) = 0 \\ |\operatorname{Re}\theta| < 2\epsilon}} \exp(\theta, T)\mu(d\theta), \tag{S. 2.9}$$

where the condition $|\operatorname{Re}\theta| < 2\epsilon$ is secured by the special choice of the fundamental solution $E(T)$ for which $\|E * \phi\|_{-\epsilon} < C\|\phi\|_{\epsilon}$ (the property (S. 2.7)). The measure μ in (S. 2. 9) is such that the integral (S. 2. 9) converges absolutely as a generalized function in the half-space $(\lambda, \tau - T) < 0$.

We now have

$$\Psi(\tau) = \int E(\tau - T)\exp(-(a, \tau - T))\check{\phi}(T)\,dT$$

$$= \int_{\operatorname{Supp} h = \mathfrak{T}} E(\tau - T)\exp(-(a, \tau - T))\check{\phi}(T)\,dT$$

$$= \exp(-(a, \tau)) \int_{\substack{P(\theta + a) = 0 \\ |\operatorname{Re}\theta| < 2\epsilon}} (\exp(\theta + a, T), \check{\phi}(T))\mu(d\theta). \qquad \text{(S. 2.10)}$$

Here we have set

$$(\exp(\theta + a, T), \check{\phi}(T)) = \int_{\mathfrak{T}} \exp(\theta + a, T)\check{\phi}(T)\,dT.$$

The interchanging of the order of integration is permissible because for a sufficiently small $\epsilon' > 0$, in view of $|\operatorname{Re}\theta| < 2\epsilon$, we have:

$$\exp(\theta + a, T) = O(\exp(-\epsilon' |T|)) \text{ for } T \in \mathfrak{T}.$$

We shall now show that

$$(\exp(\theta_0 + a, T), \check{\phi}(T)) = 0 \text{ for } p(\theta_0 + a) = 0,\ |\operatorname{Re}\theta_0| < 2\epsilon.$$

For sufficiently small a we have $\theta_0 + a \in \omega \times R^s_y$. Let N' be a connected component of the intersection $N \cap (\omega \times R^s_y)$, containing the point $\theta_0 + a$. By the conditions of the theorem N' has at least one real point ζ at which grad $P(\zeta) \neq 0$. Since $P(\theta)$ is a polynomial with real coefficients, the condition grad $P(\zeta) \neq 0$ implies that the set N' contains an open $(s - 1)$-dimensional part ν of the set of the real zeros of $P(\theta)$. Now the cotest condition

$$E_\theta \phi = 0, \quad \theta \in \Pi \cap \omega$$

implies that

$$(\exp(\theta, T), \check{\phi}(T)) = 0 \text{ for } \theta \in \nu. \qquad \text{(S. 2.11)}$$

Since $(\exp(\theta, T), \check{\phi}(T))$ is analytic in $\omega \times R^s_y$, the relation (S. 2.11) holds by the principle of analytic continuation for the whole N', i.e.

$$(\exp(\theta_0 + a, T), \check{\phi}(T)) = 0$$

for $p(\theta + a) = 0$. Hence

$$\operatorname{Supp}\Psi \subset \mathfrak{T}.$$

4. We shall show that for any $\epsilon > 0$

$$\|\Psi\|_{-\epsilon} < \infty.$$

(S. 2.8) implies that

$$\|\exp(a, T)\Psi\|_{-\epsilon} < \infty.$$

First we shall prove that Ψ does not depend upon the choice of the point $a \in \omega$. Take

$$\Psi_a = \Psi \text{ and } \Psi_{a'} = \Psi'.$$

We have $\|\Psi \exp(a, T)\|_{-\epsilon} < \infty$ and $\|\Psi' \exp(a', T)\|_{-\epsilon} < \infty$. From this and from the convexity of the cone ω we deduce that the integrals

$$\widetilde{\Psi} = \int \Psi \exp(\theta, T)\, dT \text{ and } \widetilde{\Psi}' = \int \Psi' \exp(\theta, T)\, dT$$

converge absolutely for $\theta = b + c + a'$; $b \in \omega$. But

$$P(-\mathfrak{D})\Psi = P(-\mathfrak{D})\Psi'$$

implies that

$$P(\theta)(\widetilde{\Psi} - \widetilde{\Psi}') = 0 \text{ for } \theta \in \omega + a + a'.$$

Hence $\widetilde{\Psi} = \widetilde{\Psi}'$ and $\Psi = \Psi'$, and therefore

$$\|\Psi \exp(a, T)\|_{-\epsilon} < \infty \text{ for any } a \in \omega.$$

Since $|a|$ in ω can be made as small as we need, we have

$$\|\Psi\|_{-\epsilon} < \infty.$$

The argument of this section also shows that the function Ψ in (S. 2.1) is uniquely determined by the conditions (S. 2.5).

5. *Application to the Behrens-Fisher problem.* For the Behrens-Fisher problem we have

$$p(T, \theta) = C(\theta)\, h(T) \exp(\theta_1 T_1 + \cdots + \theta_4 T_4);$$

$$\omega = (\theta_2 < 0,\ \theta_4 < 0);$$

$$h(T) = (T_2 - T_1^2)^{(n-3)/2}(T_4 - T_3^2)^{(m-3)/2}.$$

The corresponding ideal is a principal one and is generated by the polynomial

$$P(\theta) = \theta_1\theta_4 - \theta_2\theta_3.$$

Thus the Conditions 1.1–2.2 obviously hold. Let us check the requirement 3.1. We take an arbitrary point $\theta \in N \cap (\omega \times R^s_y)$ and let it move into the real subspace while remaining during the whole motion in one and the same connected component.

1) Suppose that $\theta_3 \neq 0$. Consider the path formed by the points:

$$\theta_t = (\theta_1, \theta_2, t\theta_3, t\theta_4); \quad t \in \left[1, \frac{\theta_1}{\theta_3}\right].$$

Clearly all points θ_t belong to the same component of N. Consider now the path

$$\theta'_t = (\mathrm{Re}\,\theta_1 + it\,\mathrm{Im}\,\theta_1, \mathrm{Re}\,\theta_2 + it\,\mathrm{Im}\,\theta_2, \mathrm{Re}\,\theta_1 + it\,\mathrm{Im}\,\theta_1, \mathrm{Re}\,\theta_2 + it\,\mathrm{Im}\,\theta_2); \qquad t \in (0, 1).$$

This path takes θ_t into the point $(\mathrm{Re}\,\theta_1, \mathrm{Re}\,\theta_2, \mathrm{Re}\,\theta_3, \mathrm{Re}\,\theta_4)$ belonging to the real subspace.

2) If $\theta_3 = 0$ but $\theta_4 \neq 0$ the argument must be changed in an obvious way.

3) If $\theta_3 = 0$, $\theta_4 = 0$, the path

$$\theta_t = (\mathrm{Re}\,\theta_1 + it\,\mathrm{Im}\,\theta_1, \mathrm{Re}\,\theta_2 + it\,\mathrm{Im}\,\theta_2, 0, 0); \quad t \in (1, 0)$$

takes the point $(\theta_1, \theta_2, 0, 0)$ into $(\mathrm{Re}\,\theta_1, \mathrm{Re}\,\theta_2, 0, 0)$. It remains only to remark that for $P(\theta) = \theta_1\theta_4 - \theta_2\theta_3$, we have $\mathrm{grad}\, P(\theta) = 0$ only at the origin, which does not belong to ω.

Hence for the Behrens-Fisher problem all cotests are of the form

$$\phi = \frac{1}{h}\left[\frac{\partial^2}{\partial T_1 \partial T_4} - \frac{\partial^2}{\partial T_1 \partial T_3}\right]\Psi.$$

6. We can describe all the cotests for exponential families with an arbitrary number of polynomial relations. In that case the ideal I is not necessarily a principal one. Each cotest can be written in the form

$$\phi = \frac{1}{h}\sum_j P_j(-\mathfrak{D})\Psi_j,$$

where $P_j(\theta)$ are the generators of the ideal I and $\Psi_j(T)$ are in general, generalized functions with supports in $\mathfrak{T}$.

§3. CONDITIONS OF OPTIMAL UNBIASED ESTIMATION FOR INCOMPLETE EXPONENTIAL FAMILIES WITH POLYNOMIAL RELATIONS

We consider the problem of unbiased estimation of parametric functions of n independent observations of a random variable with the density with respect to Lebesgue measure:

$$f(x;\alpha)=\exp\{t_0(x)+c_1(\alpha)t_1(x)+\cdots+c_s(\alpha)t_s(x)+c_0(\alpha)\};\tag{S. 3.1}$$

here $s\le n$; the abstract parameter $\alpha\in A$. Introduce the natural parameters

$$\theta_i=c_i(\alpha)$$

and suppose that for $\alpha\in A$ the point $\theta=(\theta_1,\cdots,\theta_s)$ runs over an everywhere dense subset of an algebraic subvariety of the domain $\Omega\subset R^s$. This subvariety we shall write in the form $\Omega\cap\Pi$, where Π is an algebraic variety in R^s given by the polynomial relations

$$\left.\begin{array}{l}\Pi_1(\theta_1,\cdots,\theta_s)=0\\ \cdots\cdots\cdots\cdots\cdots\cdots\\ \Pi_r(\theta_1,\cdots,\theta_s)=0\end{array}\right\}\tag{S. 3.2}$$

with $r<s$.

The distribution of the repeated sample $(x_1,\cdots,x_n)=\vec{x}$ from the set (S. 3.1) is given in R^n by the density with respect to Lebesgue measure

$$f^{(n)}(\vec{x};\theta)=(C(\theta))^n\exp\left\{\sum_1^n t_0(x_i)+\theta_1\sum_1^n t_1(x_i)+\cdots+\theta_s\sum_1^n t_s(x_i)\right\}\tag{S. 3.3}$$

where $C(\theta)$ is determined by the norming condition. The sufficient statistics for the family (S. 3.3) are

$$T_1=\sum_1^n t_1(x_i);\ \cdots;\ T_s=\sum_1^n t_s(x_i).$$

We shall suppose that $T_1,\cdots,T_s$ are functionally independent; then the distribution of the vector $T=(T_1,\cdots,T_s)$ is given in R^s by the density with respect to Lebesgue measure

$$p(T;\theta)=C(\theta)h(T)\exp(\theta_1T_1+\cdots+\theta_sT_s),\tag{S. 3.4}$$

where $h(T)\ge 0$; $\theta\in\Omega\cap\Pi$. We denote by $\operatorname{Supp}h$ the support of the function

$h(T)$, i. e. $\operatorname{Supp} h = \{T: h(T) > 0\}$; let $\mathcal{T} = \operatorname{int} \operatorname{Supp} h$. Our condition for $h(T)$ consists in the following requirements:

$$\left.\begin{array}{c} h(T) \text{ is infinitely differentiable on } \mathcal{T} \\ \operatorname{mes} \operatorname{Supp} h = \operatorname{mes} \mathcal{T}. \end{array}\right\} \qquad (\text{S. } 3.5)$$

Let $g(T)$ be an unbiased estimate for a certain function $\gamma(\theta)$ depending only upon the vector of sufficient statistics T and having a finite variance for all values of $\theta \in \Omega \cap \Pi$. We shall investigate the conditions which are implied for the variety Π and the estimate itself by the optimality property of the estimate for all $\theta \in \Omega \cap \Pi$ in the class of unbiased estimates of the function $\gamma(\theta)$ with finite variances. Throughout this section we take the variance for the quality measure of the estimate. It is clear that the behavior of the function $g(T)$ outside $\mathcal{T}$ is of no importance for its properties as an estimate of $\gamma(\theta)$; therefore we shall suppose $g(T)$ to be defined only on $\mathcal{T}$. Denote by N the least (complex) algebraic variety in $\mathbf{C}^s$ containing $\Pi \cap \Omega$. Since Π is itself an algebraic variety, $\Pi \cap \Omega = N \cap \Omega$.

Theorem S. 3.1. *In order for the function* $g(T)$, $T \in \mathcal{T}$ *for which* $E_\theta g^2 < \infty$; $\theta \in \Omega \cap \Pi$ *to be the best unbiased estimate of the function* $E_\theta g = \gamma(\theta)$ *for the exponential family* (S. 3.5), $\theta \in \Omega \cap \Pi$, *under condition* (S. 3.5) *it is necessary and sufficient that*:

1) *In the space* $\mathbf{C}^s$ *of variables* $\theta_1, \cdots, \theta_s$ *there exists a linear system of coordinates* $\theta'_1, \cdots, \theta'_s$ *in which* N *is a cylinder of type* $L \times \nu$, *where* L *is the coordinate space* $\theta'_1 = \cdots = \theta'_m = 0$ *and* ν *is a certain set in the space*: $\theta'_{m+1} = \cdots = \theta'_s = 0$; $0 \leq m \leq s$.

2) *In the corresponding system of coordinates* $T'_1, \cdots, T'_s$ *in the space* R^s *of values* $(T_1, \cdots, T_s)$ *the function* $g(T)$ *depends only on* $T'_{m+1}, \cdots, T'_s$.

Proof. Let $\chi(T)$ be an arbitrary unbiased estimate of zero with a finite variance, i. e.

$$E_\theta(\chi) = 0; \ E_\theta \chi^2 < \infty; \ \theta \in \Omega \cap \Pi.$$

Lemma S. 3.1. *For* $g(T)$ *to be the optimal unbiased estimate of* $\gamma(\theta)$, $\theta \in \Omega \cap \Pi$, *it is necessary and sufficient that for each unbiased estimate of zero with a finite variance*

$$E_\theta(g\chi) = 0; \ \theta \in \Omega \cap \Pi.$$

In fact, let $g(T)$ be the best unbiased estimate of $\gamma(\theta)$ and $\chi(T)$ an arbitrary unbiased estimate of zero. Then for an arbitrary constant c

$$\mathfrak{D}_\theta(g+c\chi) = \mathfrak{D}_\theta(g) + 2c\,E_\theta(g\chi) + c^2\mathfrak{D}_\theta(\chi). \tag{S. 3.6}$$

As for all values of c

$$\mathfrak{D}_\theta(g+c\chi) \geq \mathfrak{D}_\theta(g),$$

then it is easy to deduce from (S. 3.6) that $E_\theta(g\chi) = 0$, $\theta \in \Omega \cap \Pi$. Conversely, let $E_\theta(g\chi) = 0$, $\theta \in \Omega \cap \Pi$ for any unbiased estimate of zero $\chi(T)$. If the estimate $g_1(T)$ is such that

$$E_\theta g_1 = \gamma(\theta),\ E_\theta g_1^2 < \infty,\ \theta \in \Omega \cap \Pi,$$

then $\chi = g_1 - g$ will be an unbiased estimate of zero. We have:

$$\mathfrak{D}_\theta(g_1) = E_\theta(g - \gamma(\theta) + \chi)^2 = \mathfrak{D}_\theta(g) + \mathfrak{D}_\theta(\chi) \geq \mathfrak{D}_\theta(g),$$

which proves Lemma S. 3.1.

We can now prove the sufficiency of the conditions of the theorem. Without loss of generality, we can assume that in the initial system itself the subspace L is given by the equations $\theta_1 = \cdots = \theta_m = 0$, and ν is a subset of the subspace $\theta_{m+1} = \cdots = \theta_s = 0$, and that the function $g(T)$ depends only upon $T_{m+1}, \cdots, T_s$. Let $\chi(T)$ be an arbitrary unbiased estimate of zero with $E_\theta(\chi^2) < \infty$, $\theta \in N \cap \Omega$. Take any point $\theta_0 = (\theta_{10}, \cdots, \theta_{s0}) \in N \cap \Omega$, denote by ω the section of Ω by the surface $\theta_1 = \theta_{10}, \cdots, \theta_m = \theta_{m0}$ and put

$$\Psi(T) = \Psi(T_{m+1}, \cdots, T_s) = \int \chi(T) h(T) \exp(\theta_{10}T_1 + \cdots + \theta_{m0}T_m) dT_1 \cdots dT_m.$$

The relation $N = L \times \nu$ implies that, together with the point θ_0, the set N contains the points $(\theta_{10}, \cdots, \theta_{m0}, \theta_{m+1}, \cdots, \theta_s)$ for all values of $\theta_{m+1}, \cdots, \theta_s$. Hence and from the condition $E_\theta(\chi) = 0$; $\theta \in \Omega \cap N$ we deduce that

$$\int \Psi(T) \exp(\theta_{m+1}T_{m+1} + \cdots + \theta_s T_s) dT_{m+1} \cdots dT_s = 0$$

for all $(\theta_{m+1}, \cdots, \theta_s) \in \omega$. By the uniqueness theorem for Laplace transforms, $\Psi(T) = 0$. We have further for $\theta_0 \in \Omega \cap N$:

$$E_{\theta_0}(g\chi) = C(\theta_0)\int g(T_{m+1}, \cdots, T_s)\, dT_{m+1} \cdots dT_s$$

$$\times \int \chi(T) h(T) \cdot \exp(\theta_{10} T_1 + \cdots + \theta_{m0} T_m)\, dT_1 \cdots dT_m$$

$$= C(\theta_0)\int g(T_{m+1}, \cdots, T_s) \exp(\theta_{m+1,\,0} T_{m+1} + \cdots + \theta_{s,\,0} T_s)\, dT_{m+1} \cdots dT_s$$

$$\times \int \chi(T) h(T) \exp(\theta_{10} T_1 + \cdots + \theta_{m0} T_m)\, dT_1 \cdots dT_m$$

$$= C(\theta_0)\int g(T)\Psi(T) \exp(\theta_{m+1,\,0} T_{m+1} + \cdots + \theta_{s0} T_s)\, dT_{m+1} \cdots dT_s = 0.$$

By Lemma S. 3.1, $g(T)$ is the best unbiased estimate of the function $\gamma(\theta)$.

The proof of the necessity of the conditions of Theorem S. 3.1 is more complicated and the following lemmas are required. Denote by $\mathfrak{D} = \mathfrak{D}(\mathfrak{T})$ the space of infinitely differentiable functions of T with compact supports lying in $\mathfrak{T}$. For any function $\phi(T) \in \mathfrak{D}$ the Laplace transform

$$\tilde{\phi}(\theta) = \int \exp(\theta, T)\, \phi(T)\, dT$$

is an entire function in $\mathbf{C}^s$.

Directly from Lemma S. 3.1 it follows that if $g(T)$ is the best unbiased estimate of the function $\gamma(\theta)$, $\theta \in \Omega \cap \Pi$ and $\chi(T) \in \mathfrak{D}$ is such that

$$\tilde{\chi}(\theta) = 0, \quad \theta \in \Omega \cap \Pi,$$

then

$$g\tilde{\chi}(\theta) = 0, \quad \theta \in \Omega \cap \Pi. \tag{S. 3.7}$$

Lemma S. 3.2. *If an entire function $\phi(\theta)$ is equal to zero on $\Omega \cap \Pi$, then it is equal to zero on the whole set N.*

Proof. Let M be an irreducible algebraic variety in $\mathbf{C}^s$. By a theorem of Whitney [17] there exists a proper subvariety M_* with the following property: the set $(M \setminus M_*) \cap R^s$ in the vicinity of each of its points θ is a real analytic subvariety of dimension d, and the vectors grad $f(\theta)$ (where $f(\theta)$ is an arbitrary real polynomial vanishing on M) form an $(s-d)$-dimensional space.

Now represent the variety N as the sum of irreducible subvarieties

$$N = N^1 \cup \cdots \cup N^l.$$

Applying to each N^λ the theorem of Whitney, we separate in it the exclusive subvariety M_*^λ. We remark that each of the sets $N^\lambda \setminus N_*^\lambda$ has a nonvoid intersection with Ω, since otherwise $\Omega \cap \Pi$ would be a part of $\bigcup_{\mu \neq \lambda} N^{\mu} \cup N^{\lambda}$, and N would not be the least algebraic variety containing $\Omega \cap \Pi$. We now fix the λ and choose a point $\theta \in (N^{\lambda} \backslash N^{\lambda}_{*}) \cap \Omega$. By the theorem of Whitney, in the vicinity of the point θ the set $N^{\lambda} \cap \Omega$ is a real analytic variety of dimension d; moreover there exist real polynomials $f_1, \cdots, f_{s-d}$, vanishing on $N^{\lambda} \cap \Omega$, whose gradients are linearly independent at the point θ. The polynomials $f_1, \cdots \cdots, f_{s-d}$ vanish on N^{λ}, since otherwise N would not be the least variety containing $\Omega \cap \Pi$. Hence there exists a complex neighborhood U of the point θ such that $N^{\lambda} \cap U$ is a complex analytic variety of dimension d. Since $\phi(\theta) = 0$ for $\theta \in \Omega \cap N^{\lambda}$, it follows that $\phi(\theta) = 0$ for $\theta \in U \cap N^{\lambda}$. But any complex irreducible variety is a connected analytic variety with the exception of a nowhere dense set. Hence $\phi(\theta) \equiv 0$, $\theta \in N^{\lambda}$. Since λ is any of the numbers $1, \cdots, l$, $\phi \equiv 0$ on N. Lemma S.3.2 is proved.

From (S.3.7) and Lemma S.3.2 it follows that for any $\chi \in \mathfrak{D}$, if $\tilde{\chi}(\theta) = 0$, $\theta \in \Omega \cap \Pi$ then $g\tilde{\chi}(\theta) = 0$, $\theta \in N$.

Let A be the ring of all polynomials of $\theta \in \mathbf{C}^s$ with complex coefficients. If $\mathfrak{I}$ is an ideal in A, and $\zeta \in \mathbf{C}^s$, we shall denote by $\mathfrak{I}_{\zeta}$ the ideal formed by the polynomials $p(\theta + \zeta)$, where $p(\theta) \in \mathfrak{I}$. Consider the set $\mathfrak{A} \subset A$ of polynomials $p(\theta)$ for which $g(T)$ is the best unbiased estimate of the function $\gamma(\theta)$ and is also a generalized solution in $\mathfrak{T}$ of the equation

$$p(\mathfrak{D})g = 0, \quad \mathfrak{D} = \left(\frac{\partial}{\partial T_1}, \cdots, \frac{\partial}{\partial T_s}\right).$$

It is easy to see that $\mathfrak{A}$ is an ideal. Denote by I the ideal in A formed by all polynomials vanishing on N.

Lemma S.3.3. *If* $\zeta \in N$, *then* $I_{\zeta} \subset \mathfrak{A}$.

Proof. Let $p(\theta) \in I_{\zeta}$ i.e. $p(\theta - \zeta) \in I$. Considering $g(T)$ as a generalized function in $\mathfrak{T}$, we have for an arbitrary function $\phi \in \mathfrak{D}$:

$$(p(\mathfrak{D})g, \phi \exp(\zeta, T)) = (g, p(-\mathfrak{D})\phi \exp(\zeta, T))$$

$$= (g, \exp(\zeta, T)p(-\mathfrak{D}-\zeta)\phi)(T) = \int g(T) \exp(\zeta, T) p(-\mathfrak{D} - \zeta)\phi(T)\, dT.$$

Since $p(\theta - \zeta) \in I$, the function $\widetilde{p(-\mathfrak{D} - \zeta)\phi} = p(\theta - \zeta)\tilde{\phi}(\theta)$ vanishes on $\Omega \cap \Pi$. But then we have

$$\int g(T) \exp(\zeta, T) p(-\mathfrak{D} - \zeta)\phi(T)\, dT = 0; \quad \theta \in \Omega \cap \Pi$$

because the expression to the left is equal to $g\tilde{\chi}(\theta)$, where $\chi = p(-\mathfrak{D}-\zeta)\phi$ satisfies the condition $\tilde{\chi}(\theta) = 0$, $\theta \in \Omega \cap \Pi$. Now from (S. 3.8) we deduce that the generalized function $p(\mathfrak{D})g$ vanishes on the functions of the form

$$\phi \exp(\zeta, T).$$

As any function from $\mathfrak{D}$ can be represented in this form, $p(\mathfrak{D})g = 0$ in $\mathfrak{T}$ i.e. $p(\theta) \in \mathfrak{A}$. Lemma S. 3.3 is thus proved.

Consider the ideal $\mathfrak{I} = \{I_\zeta\}_{\zeta \in N}$. By Lemma S. 3.3, each summand I_ζ belongs to $\mathfrak{A}$; hence $\mathfrak{I} \subset \mathfrak{A}$. Moreover, each polynomial from $\mathfrak{I}$ vanishes on the set $L = \bigcap_{\zeta \in N}(N - \zeta)$. But if $\theta \,\overline{\in}\, L$, then for a certain $\zeta \in N$, $\theta \,\overline{\in}\, (N - \zeta)$. Hence we can find a polynomial $f \in I_\zeta \subset \mathfrak{I}$ such that $f(\theta) \neq 0$. Thus L is the set of common roots of polynomials from $\mathfrak{I}$ and therefore is an algebraic variety.

Lemma S. 3.4. *The set* $L = \bigcap_{\zeta \in N}(N - \zeta)$ *is a linear subspace, and for each point* $\zeta \in L$

$$\mathfrak{I}_\zeta = \mathfrak{I}. \tag{S. 3.9}$$

Proof. Let $\zeta \in L$; if $\theta \in N$, then $\theta + \zeta \in N$ because $\zeta \in N - \theta$. Hence

$$\mathfrak{I}_\zeta = \{I_{\theta+\zeta}\}_{\theta \in N} \subset \{I_\theta\}_{\theta \in N} = \mathfrak{I}, \tag{S. 3.10}$$

i.e. $\mathfrak{I}_\zeta \subset \mathfrak{I}$. But the set of all roots of the polynomials from $\mathfrak{I}_\zeta$ is $L - \zeta$ and so from (S. 3.10) it follows that $L - \zeta \supset L$ i.e. $L \supset L + \zeta$ for all $\zeta \in L$. Hence L is a semigroup with respect to the operation of vector addition in $\mathbf{C}^s$. We shall now show that L is a linear subspace in $\mathbf{C}^s$. Let Λ be the largest linear subspace contained in L (we remark that L contains at least the origin of coordinates). If $\Lambda \neq L$, there exists a point $\zeta \in L \setminus \Lambda$. Now take an arbitrary point $\theta \in \Lambda$. Since L is a semigroup, it must contain the points $\theta + k\zeta$; $k = 1, 2, \cdots$.

Let $p(\theta)$ be an arbitrary polynomial vanishing on L; since $p(\theta + k\zeta) = 0$, $k = 1, 2, \cdots$, $p(\theta + k\zeta) = 0$ for all $\lambda \in \mathbf{C}^1$. The set of all straight lines $\theta + \lambda\zeta$, $\lambda \in \mathbf{C}^1$ forms a linear subspace Λ' spread on Λ and ζ. Since L is an algebraic variety, $\Lambda' \subset L$; i.e. Λ is not the largest linear subspace contained in L. This contradiction proves that $L = \Lambda$.

Since L is a linear subspace, it follows from $\zeta \in L$ that $-\zeta \in L$. Hence for any points $\theta \in N$ and $\zeta \in L$, we have $\theta - \zeta \in N$. Hence in (S. 3.10) we have an inverted inclusion and (S. 3.9) is proved.

Lemma S. 3.5. *The set of vectors*

$$\operatorname{grad} p(0), \quad p \in \mathfrak{I} \tag{S. 3.11}$$

coincides with the space $L^{\perp}$ *orthogonal to* L.

Proof. It is obvious that all the vectors (S. 3.11) belong to $L^{\perp}$ and that they form a linear space. We shall show that this linear space contains all the vectors from $L^{\perp}$. Suppose this is not so; then there exists a vector $a \in L^{\perp}$ which is orthogonal to all the vectors (S. 3.11). Let q be an arbitrary polynomial from I. For each point $\zeta \in N$ the polynomial $q(\theta + \zeta)$ belongs to $\mathfrak{I}$ and so the vector $\operatorname{grad} q(\zeta)$ is orthogonal to a. Hence $q_a'(\zeta) = 0$ on N i.e. $q_a' \in I$. Applying this argument to the polynomial q_a' we see that $q_a'' = 0$ on N, as was to be proved. Hence at each point $\zeta \in N$ all derivatives of the polynomial q in the direction of the vector a are equal to zero. This means that q vanishes on the whole line $\zeta + \lambda a$, $\lambda \in C^1$, On the other hand, N is the set of the roots of all polynomials from I. Hence the set N, together with each of its points ζ, contains the whole straight line $\{\zeta + \lambda a\}$. Hence the set $L \in \bigcap_{\zeta \in N}(N - \zeta)$ contains the straight line $\{\lambda a\}$, in contradiction to the orthogonality of a to L. This contradiction proves the lemma.

Lemma S. 3.6. *Each polynomial* $p(\theta)$ *vanishing on* L *belongs to* $\mathfrak{I}$.

Proof. Select the polynomials $p_1, \cdots, p_m \in \mathfrak{I}$ in such a way that the vectors $\operatorname{grad} p_i(0)$; $i = 1, \cdots, m$ form a basis in the space $L^{\perp}$. Construct a regular system of coordinates $(w_1, \cdots, w_m)$ in the vicinity of the origin so that $w_1 = p_1, \cdots, w_m = p_m$. Since the polynomial p vanishes on the coordinate subspace $w_1 = \cdots = w_m = 0$ (i.e. on L), it can be written in the form

$$p = \sum_1^m w_i f_i = \sum_1^m p_i f_i$$

where f_i are certain functions holomorphic at the origin. By (S. 3.9) we can also obtain the representation

$$p = \Sigma p_i' f_i ; \ p_i' \in \mathfrak{I}, \tag{S. 3.12}$$

where f_i are certain functions holomorphic at an arbitrary point ζ of L. Such a representation can obviously be obtained for each point $\zeta \in \overline{L}$, since we can always find a polynomial $p' \in \mathfrak{I}$ that does not vanish at ζ.

We fix now an arbitrary point $\zeta \in \mathbf{C}^s$. Denote by R_ζ the ring of all rational functions in $\mathbf{C}^s$ whose denominators do not vanish at the point ζ; by $\mathcal{H}_\zeta$ we denote the ring of all functions holomorphic in ζ. As is well known, the pair $(R_\zeta, \mathcal{H}_\zeta)$ is flat (see [15]). This means that $\mathcal{H}_\zeta$ is a flat R_ζ-module [15] and that for each R_ζ-module E the natural operation $E \to E \otimes_{R_\zeta} \mathcal{H}_\zeta$ is injective. We now take for E the factor-ring $R_\zeta|\mathcal{J}R_\zeta$ where $\mathcal{J}R_\zeta$ is the ideal in R_ζ formed by all the functions of the form

$$\Sigma q_j r_j;\ q_j \in \mathcal{J};\ r_j \in R_\zeta.$$

Since $\mathcal{H}_\zeta$ is a flat R_ζ-module, we have

$$E \times \mathcal{H}_\zeta \cong \mathcal{H}_\zeta|\mathcal{J}\mathcal{H}_\zeta$$

where $\mathcal{J}\mathcal{H}_\zeta$ is an analogous ideal in $\mathcal{H}_\zeta$. Hence we can assert that the natural mapping

$$R_\zeta|\mathcal{J}\mathcal{H}_\zeta \to \mathcal{H}_\zeta|\mathcal{J}\mathcal{H}_\zeta$$

is injective. By (S. 3.12) the polynomial p belongs to $\mathcal{J}\mathcal{H}_\zeta$ and hence belongs to $\mathcal{J}\mathfrak{R}_\zeta$, in view of the injectivity of this mapping. Hence

$$q_\zeta p \in \mathcal{J} \tag{S. 3.13}$$

where q_ζ is a certain polynomial distinct from zero in the point ζ. Consider an ideal in A generated by the polynomials q_ζ. By the Hilbert Nullstellensatz (see for instance [1]) the unity of the ring A belongs to that ideal i. e. $\Sigma_1^k h_i q_{\zeta_i} = 1$ with certain $h_i \in A$. Putting $\zeta = \zeta_i$ in (S. 3.13), multiplying by h_i and adding, we finally obtain $p \in \mathcal{J}$, which proves the lemma.

We can now complete the proof of the theorem. In the space $\mathbf{C}^s$ choose a rectangular system of coordinates $\theta_1', \cdots, \theta_s'$ in such a way as to make L coincide with the coordinate subspace in which $\theta_1' = \cdots = \theta_m'$. In that case $\theta_1', \cdots, \theta_m'$ vanish on L and by Lemma S. 3.6 belong to the ideal $\mathcal{J}$. Since $\mathcal{J} \subset \mathfrak{A}$, in the corresponding system of coordinates in R^s we have in the domain $\mathcal{T}$ the equations

$$\frac{\partial g}{\partial T_1'} = \cdots = \frac{\partial g}{\partial T_m'} = 0.$$

These equations in the space of generalized functions mean that the function g, up to its values on a set of measure zero, is constant with respect to the variables $T'_1, \cdots, T'_m$ in the domain $\mathcal{T}$. The representation $N = L \times \nu$ is obvious, because it follows from $L \subset N - \zeta$ that together with the point ζ the set N contains the whole variety $L + \zeta$.

Theorem S. 3.1 is proved.

It is interesting to compare the situation with respect to optimal unbiased estimation of parametric functions for complete and incomplete exponential families. For the former, by the Rao-Blackwell-Kolmogorov theorem, every statistic $g(T)$ depending only upon sufficient statistics is an optimal estimate of its mathematical expectation $E_\theta g$. For the incomplete exponential families, in view of Theorem S. 3.1, the optimality of $g(T)$ as an estimate of $E_\theta g$ means, roughly speaking, a kind of quasi-completeness with respect to some of parameters (the representation $N = L \times \nu$ is analogous to the completeness). As regards the optimal estimate itself, it depends on sufficient statistics, having the same indices as the parameters with the "quasi-completeness" property.

§4. THE SAMPLE MEAN AS THE ESTIMATE OF SCALE PARAMETERS

In this section we study the families of distribution functions of the form $F(x/\sigma)$ on the half-line $(0, \infty)$ depending upon the scale parameter $\sigma \in (0, \infty)$. Our purpose is to study the unbiased estimation of σ on the evidence of the sample $(x_1, \cdots, x_n)$ from the population characterized by $F(x/\sigma)$. As usual, we take the quadratic loss function.

Suppose the condition

$$\int_0^\infty x^2 dF(x) = \alpha_2 < \infty \qquad \text{(S. 4.1)}$$

holds. If we set

$$\alpha_1 = \int_0^\infty x\, dF(x),$$

the statistics $\alpha_1^{-1}\bar{x}$ will be an unbiased estimate of the parameter with a finite variance by (S. 4.1).

We remark first of all that in the class of all (not only unbiased) estimates of σ, the $\alpha_1^{-1}\bar{x}$ are always inadmissible except in the case of a degenerate distribution $F(x)$. In fact, we have

$$E_\sigma(c_1\bar{x} - \sigma)^2 = \sigma^2 E_1(c_1\bar{x} - 1)^2$$

and $\min_{c_1} E_1(c_1\bar{x} - 1)^2$ is attained, as is easily verified, for

$$c_1 = \frac{\alpha_1}{\alpha_1^2 + (\alpha_2 - \alpha_1^2)/n}.$$

Since $\alpha_2 \geq \alpha_1^2$ and, moreover, the equality sign holds for only the degenerate ones in the class of all estimates (except in the degenerate case), $\alpha_1^{-1}\bar{x}$ is always inadmissible.

When is $\alpha_1^{-1}\bar{x}$ admissible in the class of unbiased estimates of the scale parameter σ? To answer this question we shall first prove Lemma S. 4.1, where we use the notation $y = (x_2/x_1, \cdots, x_n/x_1)$.

Lemma S. 4.1. *Let*

$$s_n = c_n\bar{x}\frac{E_1(\bar{x} \mid y)}{E_1(\bar{x}^2 \mid y)} \tag{S. 4.2}$$

where the constant c_n *is determined from the condition*

$$c_n E_1\left\{\frac{E_1(\bar{x} \mid y)^2}{E_1(\bar{x}^2 \mid y)}\right\} = 1. \tag{S. 4.3}$$

Then

$$E_\sigma s_n = \sigma;\ E_\sigma(s_n - \sigma)^2 \leq E_\sigma(\alpha_1^{-1}\bar{x} - \sigma)^2$$

and equality holds in (S. 4.4) *if and only if, with probability* 1,

$$\frac{E_1(\bar{x} \mid y)}{E_1(\bar{x}^2 \mid y)} = b_n;\ b_n = \alpha_1^{-1}c_n^{-1}.$$

Proof. Put $\gamma = \gamma(y) = E_1(\bar{x} \mid y)/E_1(\bar{x}^2 \mid y)$. We have

$$E_\sigma(s_n) = c_n E_\sigma(\bar{x}\gamma) = c_n\sigma E_1(\gamma E_1(\bar{x} \mid y)) = \sigma c_n E_1\left\{\frac{E_1(\bar{x} \mid y)^2}{E_1(\bar{x}^2 \mid y)}\right\} = \sigma$$

by (S. 4.3);

$$E_\sigma(\alpha_1^{-1}\bar{x} - \sigma)^2 = E_\sigma(\alpha_1^{-1}\bar{x} - s_n)^2 + E_\sigma(s_n - \sigma)^2 + 2E_\sigma\{(\alpha_1^{-1}\bar{x} - s_n)(s_n - \sigma)\}.$$

Further,

$$
\begin{aligned}
&E_\sigma((\alpha_1^{-1}\bar{x} - s_n)(s_n - \sigma)) \\
&= \sigma^2 E_1((\alpha_1^{-1} - c_n y)(c_n y E_1(\bar{x}^2|y) - E_1(\bar{x}\,|\,y))) \\
&= \sigma^2(c_n - 1)\left\{\alpha_1^{-1}E_1(E_1(\bar{x}\,|\,y)) - c_n E_1\left\{\frac{E_1(\bar{x}\,|\,y)^2}{E_1(\bar{x}^2|\,y)}\right\}\right\} \\
&= \sigma^2(c_n - 1)\left\{1 - c_n E_1\left[\frac{E_1(\bar{x}|\,y)^2}{E_1(\bar{x}^2|\,y)}\right]\right\} = 0,
\end{aligned}
$$

by (S. 4.3). Thus

$$E_\sigma(\alpha_1^{-1}\bar{x} - \sigma)^2 = E_\sigma(\alpha_1^{-1}\bar{x} - s_n)^2 + E_\sigma(s_n - \sigma)^2 \geq E(s_n - \sigma)^2 \quad \text{(S.4.5)}$$

and the equality sign in (S. 4.5) holds for all $\sigma \in (0, \infty)$ simultaneously if and only if with probability 1

$$\alpha_1^{-1}\bar{x} = c_n \bar{x}\,\frac{E_1(\bar{x}\,|y)}{E_1(\bar{x}^2|\,y)};$$

i. e. if

$$E_1(\bar{x}\,|\,y) = b_n E_1(\bar{x}^2\,|\,y), \qquad b_n = \alpha_1^{-1}c_n^{-1}.$$

This proves Lemma S. 4.1.

From now on we shall suppose that all the moments of $F(x)$

$$\int_0^\infty x^k\,dF(x), \qquad k = 1, 2, \cdots, \quad \text{(S.4.6)}$$

are finite.

Theorem S. 4.1. *Let the function* $F(x)$ *satisfy the condition* (S. 4.6). *In order for the statistic* $\alpha_1^{-1}\bar{x}$ *to be an admissible estimate of* σ *in the class of the unbiased estimates in the sample sizes* $n = n_1$, $n = n_2$ (n_1, n_2 *are any numbers with* $n_2 > n_1 \geq 3$) *from the population given by* $F(x/\sigma)$, *it is necessary and sufficient for* $F(x)$ *to be either degenerate*:

$$F(x) = \begin{cases} 0, & x < x_0 \\ 1, & x_0 \leq x < \infty \end{cases} \qquad \textit{for some } x_0 > 0,$$

or a gamma-distribution:

$$F(x) = \begin{cases} 0, & x \le 0 \\ (\gamma^m/\Gamma(m)) \int_0^x x^{m-1} e^{-\gamma x}\, dx; & 0 < x < \infty \end{cases}$$

for some $m > 0$, γ.

Theorem S. 4.2. *If* $F(x)$ *satisfies the condition* (S. 4.6) *and* $\alpha_1^{-1}\bar{x}$ *is optimal in the class of the unbiased estimates of the parameter* σ *for a sample size* $n \ge 3$, *then* $F(x)$ *is either degenerate or a gamma-distribution.*

We omit the proofs, which proceed by means of functional equations. They are given in detail in [5].

§5. NONPARAMETRIC APPROACH TO THE ESTIMATION OF LOCATION PARAMETERS

Let $x_1, \cdots, x_n$ be a repeated sample from the population with the distribution function $F(x - \theta)$ satisfying the conditions

$$\int x\, dF = 0, \qquad \int x^2 dF < \infty. \tag{S. 5.1}$$

In this case the parameter $\theta \in R^1$ to be estimated on the evidence of the sample $(x_1, \cdots, x_n)$ means the mathematical expectation. For the well-known type of distribution function $F(x)$ E. Pitman [14] introduced, as early as 1938, the following estimate for θ:

$$t_n = \bar{x} - E_0(\bar{x} \mid x_2 - x_1, \cdots, x_n - x_1). \tag{S. 5.2}$$

For an absolutely continuous $F(x)$, $F(x) = \int_{-\infty}^x f(u)\, du$ the Pitman estimate, as mentioned in §3, Chapter VII, can be written in the form

$$t_n = \frac{\int_{-\infty}^{\infty} \xi \prod_1^n f(x_i - \xi)\, d\xi}{\int_{-\infty}^{\infty} \prod_1^n f(x_i - \xi)\, d\xi}. \tag{S. 5.3}$$

C. Stein proved [16] that under the condition

$$\int |x|^3\, dF(x) < \infty$$

the Pitman estimate is absolutely admissible.

The loss function is assumed to be quadratic throughout this section. We shall now consider in detail the situation when the form of the function $F(x)$ is unknown. If we suppose that $F(x)$ is allowed to be *arbitrary*, satisfying only the condition (S. 5.1), it is easy to see that there is no better estimate for θ than $\bar{x}$. However, in the case when for some integer $k > 1$ the first $2k$ moments of the distribution function $F(x)$ are known:

$$\mu_l = \int x^l \, dF(x), \qquad l = 1, 2, \cdots, 2k \tag{S. 5.5}$$

and

$$\mu_{2k} = \int x^{2k} \, dF < \infty \tag{S. 5/5}$$

the information about $F(x)$ contained in the moments (S. 5.4) can be used more efficiently than $\bar{x}$ for construction of the estimates of the parameter θ.

Under condition (S. 5.5) the set of all polynomials $Q(x_1, \cdots, x_n)$ of $x_1, \cdots$ $\cdots, x_n$ of degree not exceeding k forms a Hilbert space $L_k^{(2)}$ if the scalar product of the elements Q_1 and Q_2 is defined by:

$$(Q_1, Q_2) = E_0(Q_1 Q_2).$$

The subspace formed by the polynomials $Q \in L_k^{(2)}$ of the form

$$Q = Q(x_2 - x_1, \cdots, x_n - x_1)$$

will be denoted by Λ_k.

Consider the estimate

$$t_n^{(k)}(x_1, \cdots, x_n) = t_n^k = \bar{x} - \hat{E}_0(\bar{x} \mid \Lambda_k) \tag{S. 5.6}$$

where $\hat{E}_0(\cdot \mid \Lambda_k)$ is the operator of projection on the subspace Λ_k. Note that for the construction of the estimate (S. 5.6) we need to know only the first $2k$ moments of the distribution function, not the whole function.

Theorem S. 5.1. *For all* $\theta \in R^1$

$$E_\theta t_n^k = \theta; \;\; E_\theta (t_n^k - \theta)^2 \leq E_\theta(\bar{x} - \theta)^2, \tag{S. 5.7}$$

where the equality or the inequality in (S. 5.7) *are realized simultaneously for all* $\theta \in R^1$ *and the equality holds if and only if* $\hat{E}_0(\bar{x} \mid \Lambda_k) = 0$.

Proof. Since $Q \equiv 1 \in \Lambda_k$, we have

$$(\bar{x} - \hat{E}_0(\bar{x} \mid \Lambda_k), 1) = 0.$$

Hence

$$E_\theta(\bar{x} - \hat{E}_0(\bar{x} \mid \Lambda_k)) = E_\theta \bar{x} - E_\theta(\hat{E}_0(\bar{x} \mid \Lambda_k)) = \theta - E_0(\hat{E}_0(\bar{x} \mid \Lambda_k)) = 0.$$

Moreover,

$$E_\theta(\bar{x} - \theta)^2 = E_\theta(\bar{x} - \hat{E}_0(\bar{x} \mid \Lambda_k) - \theta + E_0(\bar{x} \mid \Lambda_k))^2$$

$$= E_\theta(t_n^k - \theta)^2 + 2\,E_\theta((t_n^k - \theta)\,\hat{E}_0(\bar{x} \mid \Lambda_k)) + E_\theta(\hat{E}_0(\bar{x} \mid \Lambda_k))^2 .$$

But

$$E_\theta((t_n^k - \theta)\,\hat{E}_0(\bar{x} \mid \Lambda_k)) = E_0(t_n^k \hat{E}_0(\bar{x} \mid \Lambda_k)) = 0$$

as $t_n^k = \bar{x} - \hat{E}_0(\bar{x} \mid \Lambda_k)$ is orthogonal to each function from Λ_k. Hence

$$E_\theta(\bar{x} - \theta)^2 = E_\theta(t_n^k - \theta)^2$$

$$= E_\theta(t_n^k - \theta)^2 + E_0(\hat{E}_0(\bar{x} \mid \Lambda_k))^2 \geq E_\theta(t_n^k - \theta)^2$$

and the equality sign (for all $\theta \in R^1$ simultaneously) holds under the condition

$$\hat{E}_0(\bar{x} \mid \Lambda_k) = 0.$$

In connection with Theorem S. 5.1 it is natural to raise the question: for what functions $F(x)$ is the estimate t_n^k better than $\bar{x}$? In other words, when does the knowledge of the first $2k$ moments of $F(x)$ enable us to improve upon the standard estimate $\bar{x}$.

Theorem S. 5.2. *If $F(x)$ satisfies condition (S. 5.5), then for $n \geq 3$ the estimate t_n^k will be better than the sample mean $\bar{x}$ as an estimate of the location parameter in all cases except when the first $(k + 1)$ moments of the distribution function $F(x)$, μ_l coincide with the corresponding moments of a normal law, so that*

$$\mu_l = \begin{cases} 0, & l \text{ odd} \\ (l-1)!!\,\sigma^l, & l \text{ even} \end{cases} \qquad 1 \leq l \leq k+1$$

for a certain σ^2.

Theorem S. 5.2 is an obvious consequence of the Theorem S. 5.1 and the following lemma.

Lemma S. 5.1. *If $n \geq 3$ then $\hat{E}_0(\bar{x} \mid \Lambda_k) = 0$ if and only if the first $(k+1)$*

moments of $F(x)$ *coincide with the corresponding moments of a normal law.*

Proof of Lemma S. 5.1. We proceed by induction. For $k = 1$ the lemma holds trivially. Since $\Lambda_k > \Lambda_{k+1}$, it follows from $\hat{E}_0(\bar{x} \mid \Lambda_k) = 0$ that $\hat{E}_0(\bar{x} \mid \Lambda_{k+1}) = 0$. We can assume now that the first k moments of $F(x)$ coincide with corresponding normal moments and we shall prove that then the moment μ_{k+1} coincides with the $(k+1)$st normal moment.

The condition

$$\hat{E}_0(\bar{x} \mid \Lambda_k) = 0 \tag{S. 5.8}$$

is equivalent to the set of conditions

$$\hat{E}_0(\bar{x} \mid \Lambda_{k-1}) = 0 \tag{S. 5.9}$$

$$\hat{E}_0(\bar{x}(x_{j_1} - x_1) \cdots (x_{j_k} - x_1)) = 0, \tag{S. 5.10}$$

where (S. 5.10) must hold for all the sets $(j_1, \cdots, j_k)$ of integers $2, \cdots, n$. The equivalence of the conditions (S. 5.8) and (S. 5.9)–(S. 5.10) follows from the fact that Λ_{k-1} and the functions $(x_{j_1} - x_1) \cdots (x_{j_k} - x_1)$ generate the whole Λ_k.

Let

$$(x_{j_1} - x_1) \cdots (x_{j_k} - x_1) = (x_{i_1} - x_1)^{\alpha_1} \cdots (x_{i_s} - x_1)^{\alpha_s}, \tag{S.5.11}$$

where $i_1, \cdots, i_s$ are mutually distinct and $\alpha_1 + \cdots + \alpha_s = k$. We have

$$(x_{i_1} - x_1)^{\alpha_1} \cdots (x_{i_s} - x_1)^{\alpha_s}$$

$$= \sum_{l_1, \cdots, l_s = 0}^{\alpha_1, \cdots, \alpha_s} (-1)^{l_1 + \cdots + l_s} C_{\alpha_1}^{l_1} \cdots C_{\alpha_s}^{l_s} x_1^{\alpha_1} \cdots x_{i_s}^{\alpha_s} x_1 k - \sum_{i=1}^{s} l_i. \tag{S.5.12}$$

Put $\Sigma_1^s l_i = l$; then from (S. 5.12) we get

$$\left.\begin{aligned} &E_0(\bar{x}(x_{i_1} - x_1)^{\alpha_1} \cdots (x_{i_s} - x_1)^{\alpha_s}) \\ &= \frac{1}{n} \sum_{l_1=0}^{\alpha_1} \cdots \sum_{l_s=0}^{\alpha_s} (-1)^l C_{\alpha_1}^{l_1} \cdots C_{\alpha_s}^{l_s} \mu_{l_1} \cdots \mu_{l_s} \mu_{k+1-l} \\ &+ \frac{1}{n} \sum_{q=1}^{s} \sum_{l_1=0}^{\alpha_1} \cdots \sum_{l_s=0}^{\alpha_s} (-1)^l C_{\alpha_1}^{l_1} \cdots C_{\alpha_s}^{l_s} \mu_{l_1} \cdots \mu_{l_{q-1}} \mu_{l_{q+1}} \cdots \mu_{l_s} \mu_{k-l}. \end{aligned}\right\} \tag{S. 5.13}$$

We shall consider the cases of odd and even values of $(k-1)$ separately.

1) $k-1$ odd. By the induction assumption, all odd moments up to the order $(k-1)$ are equal to zero. Hence from (S. 5.13) we get

$$n E_0(\bar{x}(x_{i_1}-x_1)^{\alpha_1}\cdots(x_{i_s}-x_1)^{\alpha_s})$$

$$= \Sigma^*(-1)^l C_{\alpha_1}^{l_1}\cdots C_{\alpha_s}^{l_s}\mu_{l_1}\cdots\mu_{l_s}\mu_{k+1-l}$$

$$+\sum_{q=1}^{s}\Sigma_q^*(-1)^l C_{\alpha_1}^{l_1}\cdots C_{\alpha_s}^{l_s}\mu_{l_1}\cdots\mu_{l_{q-1}}\mu_{l_{q+1}}\cdots\mu_{l_s}\mu_{k-l}, \quad \text{(S. 5.14)}$$

where the summation in Σ^* is taken over all even $l_1, \cdots, l_s$ and in Σ_q^* over even $l_1, \cdots, l_{q-1}, l_{q+1}, \cdots, l_s$ and odd l_q, the limits being indicated by (S. 5.13). In the sum Σ^* the number l is always even and therefore $(k+1-l)$ is odd. Moreover, if $k+1-l \le k-1$, then $\mu_{k+1-l}=0$ by the induction assumption. In the sum Σ_q^*, the number l is always odd; hence $(k-l)$ is also odd; since $k-l \le k-1$ we have $\mu_{k-l}=0$. Hence the condition (S. 5.10) is equivalent to the relation

$$\mu_{k+1}=0. \quad \text{(S. 5.15)}$$

2) $(k-1)$ is even. Consider again the relation (S. 5.14). We break Σ^* and $\Sigma_{q=1}^{s}\Sigma_q^*$ into subsums

$$\Sigma^* = S_0+S_2+\cdots+S_{k-1},$$

where S_{2m} is the part of the sum Σ^* corresponding to all the values $l_1, \cdots, l_s$, $0 \le l_1 \le \alpha_1, \cdots, 0 \le l_s \le \alpha_s$ for which $l_1+\cdots+l_s=2m$, and

$$\sum_{q=1}^{s}\Sigma_q^* = S_1+S_3+\cdots+S_k$$

where S_{2m+1} is the part of the double sum corresponding to the values $l_1, \cdots, l_s$ for which $l_1+\cdots+l_s=2m+1$.

Consider first the conditions

$$0<\alpha_1<k, \cdots, \quad 0<\alpha_s<k. \quad \text{(S. 5.16)}$$

We shall show that $S_{2m}+S_{2m+1}=0$ for $0<2m \le k-1$. Note that in view of the condition (S. 5.16) in the sum $\Sigma_{q=1}^{s} \Sigma_q^*$ we have $l_q + 1 \le h$; hence by the induction assumption

$$\mu_{l_q+1} = l_q \mu_{l_q-1} \sigma^2 \tag{S. 5.17}$$

for a certain $\sigma^2 > 0$. Then

$$S_{2m+1} = -\sum_{q=1}^{s} \sum_{l_1+\cdots+l_s=2m+1}^{*} {}_q \, C_{\alpha_1}^{l_1} \cdots C_{\alpha_s}^{l_s} \mu_{l_1} \cdots \mu_{l_{q-1}} \mu_{l_q+1} \mu_{l_{q+1}} \cdots \mu_{l_s} \mu_{k-1-2m}$$

$$= -\sum_{l_1+\cdots+l_s=2m}^{*} C_{\alpha_1}^{l_1} \cdots C_{\alpha_s}^{l_s} \mu_{l_1} \cdots \mu_{l_s} \mu_{k-1-2m} \sigma^2$$

$$\times \left[\frac{C_{\alpha_1}^{l_1+1}}{C_{\alpha_1}^{l_1}} (l_1 + 1) + \cdots + \frac{C_{\alpha_s}^{l_s+1}}{C_{\alpha_s}^{l_s}} (l_s + 1) \right]. \tag{S. 5.18}$$

But

$$\frac{C_{\alpha_1}^{l_1+1}}{C_{\alpha_1}^{l_1}} (l_1 + 1) + \cdots + \frac{C_{\alpha_s}^{l_s+1}}{C_{\alpha_s}^{l_s}} (l_s + 1) = \sum_{q=1}^{s} (\alpha_q - l_q) = k - 2m.$$

Since

$$\mu_{k-1-2m} \sigma^2 (k - 2m) = \mu_{k+1-2m}$$

(recall that $m > 0$), we get from (S. 5.18)

$$S_{2m} + S_{2m+1} = 0. \tag{S. 5.19}$$

In view of (S. 5.19) the condition (S. 5.10) is equivalent to

$$S_0 + S_1 = 0. \tag{S. 5.20}$$

But

$$S_0 = \mu_{k+1},$$

$$S_1 = -(\alpha_1 + \cdots + \alpha_s) \mu_2 \mu_{k-1} = -k \sigma^2 \mu_{k-1}.$$

Hence (S. 5.10) is equivalent to the equality

$$\mu_{k+1} = k\sigma^2 \mu_{k-1}. \tag{S. 5.21}$$

The induction assumption together with (S. 5.21) gives

$$\mu_{k+1} = k!!\, \sigma^{k+1}.$$

Now let one of the numbers $\alpha_1, \cdots, \alpha_s$ be equal to k; then all the other numbers vanish. Without loss of generality we can assume that

$$\alpha_{i_1} = k, \quad \alpha_{i_2} = \cdots = \alpha_{i_s} = 0. \tag{S. 5.22}$$

In this case (S. 5.14) reduces to

$$n E_0(\bar{x}(x_{i_1} - x_1)^k) = \sum_{\substack{l=0 \\ l \text{ even}}}^{k-1} C_k^l \mu_l \mu_{k+1-l} - \sum_{\substack{l=1 \\ l \text{ odd}}}^{k} C_k^l \mu_{l+1} \mu_{k-l}.$$

In the second sum we put $l' = k - l$; then

$$n E_0(\bar{x}(x_{i_1} - x_1)^k) = \sum_{\substack{l=0 \\ l \text{ even}}}^{k-1} C_k^l \mu_l \mu_{k+1-l} - \sum_{\substack{l'=0 \\ l' \text{ even}}}^{k-1} C_k^{l'} \mu_{k+1-l'} \mu_{l'} = 0,$$

so that under condition (S. 5.22), the condition (S. 5.14) is always satisfied.

We remark that for $n \geq 3$ the subspace Λ_k always contains the function $(x_{i_1} - x_1)^{\alpha_1} \cdots (x_{i_s} - x_1)^{\alpha_s}$ under condition (S. 5.16).

Hence we have established that for $n \geq 3$ the condition (S. 5.8) is equivalent to the coincidence of the first $(k+1)$ moments of the distribution function $F(x)$ with the moments of a normal law. This completes the proof of Lemma S. 5.1 and Theorem S. 5.2.

BIBLIOGRAPHY

[1] B. L. van der Waerden, *Moderne Algebra*. Vol. 2, Springer, Berlin, 1931; 4th ed., 1959; Russian transl., GITTL, Moscow, 1947. MR 2, 120; MR 31 #1292.

[2] A. M. Kagan and V. P. Palamodov, *Conditions of optimal unbiased estimation of parametric functions for incomplete exponential families with polynomial ties*, Dokl. Akad. Nauk SSSR 1967. (Russian)

[3] ———, *Incomplete exponential families and variance unbiased minimum estimates*. I, Teor. Verojatnost. i Primenen. 12 (1967), 34–49. (Russian)

[4] A. M. Kagan, *Sample mean as an estimate of the shift parameters*, Dokl. Akad. Nauk SSSR 169 (1966), 1006–1008 = Soviet Math. Dokl. 7 (1966), 1041–1043. MR 33 # 6747.

[5] A. M. Kagan and A. L. Ruhin, *On the theory of the estimation of a scale parameter*, Teor. Verojatnost. i Primenen. 12 (1967). (Russian)

[6] V. P. Palamodov, *On verifiable functions*, Teor. Verojatnost. i Primenen. 12 (1967). (Russian)

[7] ———, *Testing multidimensional polynomial hypotheses*, Dokl. Akad. Nauk SSSR 172 (1966), 291–293 = Soviet Math. Dokl. 7 (1966), 95–97.

[8] ———, *On systems of differential equations with constant coefficients*, Dokl. Akad. Nauk SSSR 148 (1963), 523–526 = Soviet Math. Dokl. 4 (1963), 133–136. MR 29 # 1442.

[9] ———, Ph. D. Thesis, Moscow State University, Moscow, 1965. (Russian)

[10] G. M. Fihtengol'c, *A course on differential and integral calculus*. Vol. II, "Nauka", Moscow, 1966; German transl., Hochschülbucher für Mathematik, Band 62, 2nd ed., VEB Deutscher Verlag, Berlin, 1966.

[11] L. Hörmander, *Linear partial differential operators*, Die Grundlehren der math. Wissenschaften, Band 116, Academic Press, New York and Springer-Verlag, Berlin, 1963. MR 28 # 4221.

[12] A. M. Kagan, *On the estimation theory of location parameter*, Sankhya (1966).

[13] E. Lehmann, *On the non-verifiability of certain parametric functions*, Teor. Verojatnost. i Primenen. 10 (1965), 758–760. (Russian summary) MR 32 # 8445.

[14] E. J. G. Pitman, *The estimation of the location and scale parameters of a continuous population of any given form*, Biometrika 30 (1938), 391–421.

[15] J.-P. Serre, *Géométrie algébrique et géométrie analytique*, Ann. Inst. Fourier, Grenoble 6 (1955–56), 1–42. MR 18, 511.

[16] C. Stein, *The admissibility of Pitman's estimator of a single location parameter*, Ann. Math. Statist. 30 (1959), 970–979. MR 22 # 278.

[17] H. Whitney, *Elementary structure of real algebraic varieties*, Ann. of Math. (2) 66 (1957), 545–556. MR 20 # 2342.

[18] E. Lehmann, *Testing statistical hypotheses*, Wiley, New York, and Chapman & Hall, London, 1959. MR 21 # 6654.

BIBLIOGRAPHY

BAHADUR, R. R.

1. *Sufficiency and statistical decision functions,* Ann. Math. Statistics 25 (1954), 423–462. MR 16, 154.

BASU, D.

2. *On statistics independent of sufficient statistics,* Sankhyā 20 (1958), 223–226. MR 21 #4494.

BESICOVITCH, A. S.

3. *On diagonal values of probability vectors of infinitely many components,* Proc. Cambridge Philos. Soc. 57 (1961), 759–766. MR 23 #A4154.

BLACKWELL, D.

4. *Conditional expectation and unbiased sequential estimation,* Ann. Math. Statistics 18 (1947), 105–110. MR 8, 478.

BOCHNER, C. and MARTIN, W. T.

5. *Several complex variables,* Princeton Univ. Press, Princeton, N. J., 1948; Russian transl., IL, Moscow, 1951. MR 10, 366.

BRÉNY, H.

6. *L'état actuel du problème de Behrens-Fisher,* Trabajos Estadíst. 6 (1955), 111–113. MR 17, 868.

BROWN, L.

7. *Sufficient statistics,in the case of independent random variables,* Ann, Math. Statistics 35 (1964), 1456–1474.

BOURBAKI, N.

8. *Espaces vectoriels topologiques, Éléments de Mathématiques,* no. 1189. 1229, Hermann, Paris, 1953, 1955; Russian transl., IL, Moscow, 1959. MR 14, 880; MR 17, 1109.

WALD, A.

9. *Testing the difference between the means of two normal populations with unknown standard deviations,* Selected Papers in Probability and Statistics, McGraw-Hill, New York, 1955, pp. 669–695. MR 16, 435.

WIDDER, D.

10. *The Laplace transform,* Princeton Math. Series, vol. 6, Princeton Univ. Press, Princeton, N. J., 1946.

WIJSMAN, R. A.

11. *Incomplete sufficient statistics and similar tests,* Ann. Math. Statistics 29 (1958), 1028–1045. MR 21 #5256.

VOLODIN, I. N.

12. *On the distinction between the Poisson and Pólya distributions when a large number of small samples is available,* Teor. Verojatnost. i Primenen. 10 (1965), 364–367. (Russian) MR 31 #6299.

GLEASON, A. M.

13. *Finitely generated ideals in Banach algebras,* J. Math. Mech. 13 (1964), 125–132. MR 28 #2458.

DANTZIG, G.

14. *On the non-existence of tests of "Student's" hypothesis having power function independent of σ,* Ann. Math. Statistics 11 (1940), 186–191. MR 1, 348.

DARMOIS, G.

15. *Sur les lois de probabilité à estimation exhaustive,* C. R. Acad. Sci. Paris 260 (1935), 1265–1266.

DOETSCH, G.

16. *Handbuch der Laplace-Transformation* Vols. I, II, Birkhäuser, Basel, 1950, 1955. MR 13, 230; MR **18**, 35.

DOWKER, C. H.

17. *Lectures on sheaf theory,* Tata Institute, Bombay, 1956, 1962. MR 19, 301.

DOOB, J. L.

18. *Stochastic processes,* Wiley, New York, 1953; Russian transl., IL, Moscow, 1956. MR 15, 445; MR 19, 71.

DYNKIN, È. B.

19. *Necessary and sufficient statistics for a family of probability distributions,* Uspehi Mat. Nauk 6 (1951), no. 1(41), 68–90; English transl., Selected Transl. Math. Stat. and Prob., vol. 1, Amer. Math. Soc., Providence, R.I., 1961, pp. 17–40. MR 12, 839.

ZINGER, A. A.

20. *Independence of quasi-polynomial statistics and analytical properties of distributions,* Teor. Verojatnost. i Primenen. 3 (1958), 265–284. (Russian) MR 21 #941.

21. *On a problem of A. N. Kolmogorov,* Vestnik Leningrad. Univ. 11 (1956), no. 1, 53–56. (Russian) MR 17, 863.

ZINGER, A. A. and LINNIK, Ju. V.

22. *Characterization of the normal distribution,* Teor. Verojatnost. i Primenen. 9 (1964), 692–695. (Russian) MR 30 #607.

KAGAN, A. M. and LINNIK, Ju. V.

23. *A class of families admitting similar zones,* Vestnik Leningrad. Univ. Ser. Mat. Meh. Astronom. 19 (1964), 25–36. (Russian)

KAGAN, A. M. and ŠALAEVSKIĬ, O. V.

24. *The Behrens-Fisher problem for the existence of similar regions in an algebra of sufficient statistics,* Dokl. Akad. Nauk SSSR 155 (1964), 1250–1252 = Soviet Math. Dokl. 5 (1964), 556–558. MR 28 #4627a.

CARTAN, H.

25. *Idéaux des fonctions analytiques de n variables complexes,* Ann. Sci. École Norm. Sup. (3) 61 (1944), 149–197. MR 7, 290.

26. *Sur les matrices holomorphes de n variables complexes,* J. Math. Pures Appl. 19 (1940), 1–26. MR 1, 312.

27. *Idéaux et modules de fonctions analytiques de variables complexes,* Bull. Soc. Math. France 78 (1950), 29–64. MR 12, 172.

28. *Variétés analytiques complexes et cohomologie,* Colloque sur les Fonctions de Plusieurs Variables, Bruxelles 1953, Georges Thone, Liège and Masson & Cie, Paris, 1953, pp. 41–55. MR 16, 235.

29. *Variétés analytiques réelles et variétés analytiques complexes,* Bull. Soc. Math. France 85 (1957), 77–99. MR 20 #1339.

KENDALL, M. G.

30. *The evergreen correlation coefficient,* Contributions to Probability and Statistics, Stanford Univ. Press, Stanford, Calif., 1960, pp. 274–277. MR 22 #11457.

KOLMOGOROV, A. N.

31. Izv. Akad. Nauk SSSR Ser. Mat. 6 (1942), 3–32. MR 4, 221.

32. *Unbiased estimates,* Izv. Akad. Nauk SSSR Ser. Mat. 14 (1950), 303–326; English transl., Amer. Math. Soc. Transl. (1) 11 (1962), 144–170. MR 12, 116.

CRAMÉR, H.

33. *Mathematical methods of statistics,* Princeton Math. Series, vol. 9, Princeton Univ. Press, Princeton, N. J., 1946; Russian transl., IL, Moscow, 1948. MR 8, 39.

KOOPMAN, B. O.

34. *On distributions admitting a sufficient statistic,* Trans. Amer. Math. Soc. 39 (1936), 399–409.

KULLBACK, S.

35. *Information theory and statistics,* Wiley, New York and Chapman & Hall, London, 1959. MR 21 #2325.

LEHMANN, E.

36. *Testing statistical hypotheses,* Wiley, New York and Chapman & Hall, London, 1959; Russian transl., IL, Moscow, 1963 and "Nauka", Moscow, 1964. MR 21 #6654.

LEHMANN, E. and SCHEFFÉ, H.

37. *Completeness, similar regions and unbiased estimation.* I, Sankhyā 10 (1950), 305–340. MR 12, 511.

LINNIK, Ju. V.

38. *On polynomial statistics in connection with the analytical theory of differential equations,* Vestnik Leningrad. Univ. 11 (1956), no. 1, 35–48; English transl., Selected Transl. Math. Stat. and Prob., vol. 1, Amer. Math. Soc., Providence, R.I., 1961, pp. 171–206. MR 17, 983.

39. *Polynomial statistics and polynomial ideals,* Calcutta Math. Soc. Golden Jubilee Commemoration Volume (1958–1959), Part I, Calcutta Math. Soc., Calcutta, 1963, pp. 95–98. MR 29 #2830.

40. *On the theory of statistically similar regions,* Dokl. Akad. Nauk SSSR 146 (1962), 300–302 = Soviet Math. Dokl. 3 (1962), 1297–1299. MR 25 #3584.

41. *Sur certaines questions de la statistique analytique,* Ann. Fac. Sci. Univ. Clermont Math. No. 8 (1962), 53–61.

42. *Complex variables in problems with nuisance parameters and finite rank sufficient statistics,* Dokl. Akad. Nauk SSSR 149 (1963), 1026–1028 = Soviet Math. Dokl. 4 (1963), 512–513. MR 27 #3041.

43. *Remarks on the Fisher-Welch-Wald test,* Dokl. Akad. Nauk SSSR 154 (1964), 514–516 = Soviet Math. Dokl. 5 (1964), 118–120.

44. *Randomized homogeneous tests for the Behrens-Fisher problem,* Izv. Akad. Nauk SSSR Ser. Math. 28 (1964), 249–260; English transl., Selected Transl. Math. Stat. and Prob., vol. 6, Amer. Math. Soc., Providence, R.I., 1966, pp. 207–217. MR 28 #5521.

45. *On the construction of optimal similar solutions of the Behrens-Fisher problem,* Trudy Mat. Inst. Steklov. 79 (1965), 40–53 = Proc. Steklov Inst. Math. no. 79 (1965), 41–56.

46. *On A. Wald's test for the comparing of two normal samples,* Teor. Verojatnost. i Primenen 9 (1964), 16–30. (Russian) MR 28 #5520.

47. *Characterization of tests of the Bartlett-Scheffé type,* Trudy Mat. Inst. Steklov. 79 (1965), 32–39 = Proc. Steklov Inst. Math. no. 79 (1965), 32–40.

48. *An application of a theorem of H. Cartan in mathematical statistics,* Dokl. Akad. Nauk SSSR **160** (1965), 1248–1249 = Soviet Math. Dokl. 6 (1965), 291–293. MR 31 #6298.

LINNIK, Ju. V., ROMANOVSKAJA, J. L. and ŠALAEVSKIĬ, O. V.

49. *Remarks on the theory of the Fisher-Welch-Wald test,* Teor. Verojatnost. i Primenen. 10 (1965), 727–730. (Russian) MR 32 #8446.

LINNIK, Ju. V., ROMANOVSKIĬ, J. V. and SUDAKOV, V. N.

50. *A non-randomized homogeneous test in the Behrens-Fisher problem,* Dokl. Akad. Nauk SSSR 155 (1964), 1262–1264 = Soviet Math. Dokl. 5 (1964), 570–572. MR 28 #4627b.

LINNIK, Ju. V. and ŠALAEVSKIĬ, O. V.

51. *On the analytic theory of tests for the Behrens-Fisher problem,* Dokl. Akad. Nauk SSSR 150 (1963), 26–27 = Soviet Math. Dokl. 4 (1963), 580–582. MR 27 #3042.

ŁOJASIEWICZ, S.

52. *Sur le problème de division,* Studia Mathematica **18** (1959), 87–136. MR 21 #5893.

LUKACZ, E.

53. *Characterization of populations by properties of suitable statistics,* Proc. Third Berkeley Sympos. Math. Stat. and Prob. 1954–1955, vol. 2, University of California Press, Berkeley, 1956, pp. 195–214. MR 18, 942.

LJAPUNOV, A. A.

54. *On completely additive vector-valued functions,* Izv. Akad. Nauk SSSR Ser. Mat. 4 (1940), 465–468. (Russian) MR 2, 315.

MALGRANGE, B.

55. *Lectures on the theory of functions of several complex variables,* Tata Institute, Bombay, 1962.

NEYMAN, J.

56. *Sur la vérification des hypothèses statistiques composées,* Bull. Soc. Math. France 63 (1935), 346–366.

57. *Un théorème d'existence,* C. R. Acad. Sci. Paris 222 (1946), 843–845. MR 7, 457.

58. *Current problems of mathematical statistics,* Proc. Internat. Congress Math. 1954, vol. 1, Noordhoff, Groningen and North-Holland, Amsterdam, 1957, pp. 349–370. MR 20 #1374.

OKA, K.

59. *Sur les fonctions analytiques de plusieurs variables.* VII: Sur quelques notions arithmétiques, Bull. Soc. Math. France 78 (1950), 1–27. MR 12, 18.

PETROV, A. A.

60. *Verification of statistical hypotheses on the type of distribution based on small samples,* Teor. Verojatnost. i Primenen. 1 (1956), 248–271. (Russian) MR 19, 76.

RAO, C. R.

61. *Information and the accuracy attainable in the estimation of statistical parameters,* Bull. Calcutta Math. Soc. 37 (1945), 81–91. MR 7, 464.

62. *Some theorems on minimum variance estimates,* Sankhyā 12 (1952), 27–42. MR 14, 1103.

ROMANOVSKAJA, J. L.

63. *On the Fisher-Welch-Wald test,* Sibirsk. Mat. Ž. 5 (1964), 1343–1359. (Russian) MR 30 #661.

ROMANOVSKIĬ, I. V. and SUDAKOV, V. N.

64. *On the existence of independent partitions,* Trudy Mat. Inst. Steklov. 79 (1965), 5–10 = Proc. Steklov Inst. Math. no. 79 (1965), 1–7.

RÜCKERT, W.

65. *Zur Eliminationsproblem der Potenzreihenideale,* Math. Ann. 107 (1932), 259–281.

SAKS, S.

66. *Théorie de l'intégrale,* Warsaw, 1933; English transl., rev. ed., Dover, New York, 1964; Russian transl., IL, Moscow, 1949.

STEIN, Ch.

67. *A two-sample test for a linear hypothesis whose power is independent of the variance,* Ann. Math. Statistics 16 (1945), 243–258. MR 7, 213.

SVERDRUP, E.

68. *Similarity, unbiasedness, minimaxibility and admissibility of statistical test procedures,* Skand. Aktuarietidskr. 36 (1953), 64–86. MR 15, 453.

HALMOS, P. R.

69. *Measure theory,* Van Nostrand, Princeton, N. J., 1950; Russian transl., IL, Moscow, 1953. MR 11, 504; MR 16, 22.

HALMOS, P. R. and SAVAGE, L. J.

70. *Application of the Radon-Nikodym theorem to the theory of sufficient statistics,* Ann. Math. Statistics 20 (1949), 225–241. MR 11, 42.

HARDY, G. H., LITTLEWOOD, J. E. and PÓLYA, G.

71. *Inequalities,* 2nd ed., Cambridge Univ. Press, New York, 1952; Russian transl., IL, Moscow, 1948. MR 13, 727; MR 18, 722.

HARDY, B.

72. *Some properties of an angular transformation of the correlation coefficient,* Biometrika 43 (1956), 219–224. MR 17, 981.

HOGG, R. and CRAIG, A.

73. *Sufficient statistics in elementary distribution theory,* Sankhyā 17 (1956), 209–216. MR 19, 188.

FEĬGEL'SON, T. S.

74. *On a simple method of establishing independence of statistics,* Vestnik Leningrad Univ. Ser. Mat. Astronom. 19 (1964), no. 3, 157–158. (Russian) MR 29 #4131.

FUKS, B. A.

75. *Introduction to the theory of analytic functions of several complex variables,* Fizmatgiz, Moscow, 1962; English transl., Transl. Math. Monographs, vol. 8, Amer. Math. Soc., Providence, R. I. 1963; reprint 1965. MR 27 #4945; MR 29 #6049.

76. *Special chapters in the theory of analytic functions of several complex variables,* Fizmatgiz, Moscow, 1963; English transl., Transl. Math. Monographs, vol. 14, Amer. Math. Soc., Providence, R. I., 1965. MR 30 #4979.

ŠALAEVSKIĬ, O. V.

77. *On the non-existence of regularly varying tests for the Behrens-Fisher problem,* Dokl. Akad. Nauk SSSR 151 (1963), 509–510 = Soviet Math. Dokl. 4 (1963), 1043–1045. MR 27 #2037.

CHATTERJEE, S. K.

78. *On an extension of Stein's two sample procedure to the multi-normal problem,* Calcutta Statist. Assoc. Bull. 8 (1959), 121–148. MR 21 #4501.

SCHEFFÉ, H.

79. *On solutions of the Behrens-Fisher problem based on the t-distribution,* Ann. Math. Statistics **14** (1943), 35–44. MR **4**, 221.

SHOHAT, J. A. and TAMARKIN, J. D.

80. *The problem of moments,* Math. Surveys, vol. 1, Amer. Math. Soc., Providence, R. I., 1943; rev. ed., 1947. MR **5**, 5.

ISBN 978-0-8218-4687-2

9 780821 846872

MMONO/20.S

AMS *on the* Web
www.ams.org