A unified asymptotic expansion for distributions of quadratic functionals in nonlinear regression models

Holger Dette
Ruhr-Universität Bochum
Fakultät für Mathematik
44780 Bochum
Germany
email: holger.dette@ruhr-uni-bochum.de
FAX: +49 2 34 32 14 559

Yuri Grigoriev
Dept. of Mathematics
Siberian Transport University
D. Kovalchuc Str. 191
Novosibirsk 630049
Russia
email: grigoriev@online.nsk.su

July 5, 2001

Abstract

In the common nonlinear regression model asymptotic expansions for quadratic statistics are considered, which are used in the construction of confidence regions and statistical tests of hypotheses for the unknown parameter. One purpose of this paper is to present a unified treatment of the problem, which can applied to a substantially broader class of quadratic statistics as recently discussed in Grigoriev and Ivanov (1993) or Ivanov (1997). In particular the method proposed in this paper yields second order asymptotic expansions for the statistics introduced by Hamilton, Watts and Bates (1982), Pazman (1992) and Grigoriev (1994). A second purpose of the paper is to use these results for the classification of certain properties (invariance, u_n-representability) of several quadratic statistics proposed in the literature for inference in nonlinear regression models.

Keywords and Phrases: nonlinear regression, least squares estimation, asymptotic expansion, invariance

1 Introduction

Let $(\mathbb{R}^n, \mathcal{B}^n)$ denote the n-dimensional space equipped with its Borel field and consider the statistical experiments $(\mathbb{R}^n, \mathcal{B}^n, P_\theta^n, \theta \in \Theta)$ generated by the nonlinear regression model

$$y = \eta(\theta) + e = (\eta(x^1, \theta), \ldots, \eta(x^n, \theta))^T + e, \tag{1.1}$$

with the observed vector $y \in \mathbb{R}^n$, the vector of unknown parameter $\theta = (\theta^1, \ldots, \theta^m)^T \in \Theta \subseteq \mathbb{R}^m, m \leq n$, and the random vector $e = (e^a)_{a=1}^n$ with i.i.d. components e^a. We suppose that e^a has a distribution not depending on the parameter θ with expected value zero, variance $\gamma_2 = \sigma^2$ and existing third and fourth order cumulants γ_3 and γ_4, respectively. The set Θ is a convex and open set and for fixed $x = (x^1, \ldots, x^n)$ the mapping $\theta \mapsto \eta$ is supposed to be continuous with continuous third (or fourth) order derivatives (if required) such that the rank of the $n \times m$ matrix

$$F = F(\theta) = \left(\frac{\partial \eta^a}{\partial \theta^i}\right)_{i=1,\ldots,m}^{a=1,\ldots,n} = (F_i^a)_{i=1,\ldots,m}^{a=1,\ldots,n} \tag{1.2}$$

is m for all $\theta \in \Theta$. Here $\eta^a(\theta) = \eta(x^a, \theta)$ denotes the ath component of the vector $\eta(\theta)$, $x^a \in \mathcal{X}$ is the ath value of explanatory variable and $\mathcal{X}$ is the design space with sigma field containing all one point sets and containing at least m points. Let $\hat{\theta}_n = \hat{\theta}_n(y^1, \ldots, y^n)$ denote the least squares estimator of the unknown parameter $\theta \in \Theta$ obtained from the observation $y = (y^1, \ldots, y^n)$ by the condition

$$S(\hat{\theta}_n) = \inf_{\tau \in \Theta} S(\tau), \quad S(\theta) = \|y - \eta(\theta)\|^2, \tag{1.3}$$

where $\|\cdot\|^2 = \langle \cdot, \cdot \rangle$ and the inner product $\langle \cdot, \cdot \rangle$ on $\mathbb{R}^n$ is defined with respect to the matrix $n^{-1}\delta_{ab}$ (here and throughout this paper $\delta_{ab} = \delta^{ab} = \delta_a^b$ denotes Kronecker's symbol and simultaneously the identity matrix).
For the sake of simplicity a function $g(\theta)$ evaluated at the least squares estimator $\theta = \hat{\theta}_n$ will be denoted with $\hat{g}$, e.g. $\hat{S} = S(\hat{\theta}_n)$ (by formula (1.3)) or $\hat{\eta} = \eta(\hat{\theta}_n)$ (by formula (1.1)). For a particular value $\theta_0 \in \Theta$ we use the notation $\eta = \eta(\theta_0)$. Following differential geometric convention we denote a matrix $A = (A_{ij})_{i=1,\ldots,n_1}^{j=1,\ldots,n_2}$ with A_{ij}. If $n_1 = n_2$, then A^{ij} denotes simultaneously elements of A^{-1} and the matrix A^{-1} itself (i.e. $A = A_{ij}, A^{-1} = A^{ij}$) the specific meaning will be clear from the context. We will also make substantial use of Einstein's rule; for example $A^{ij}B_{jk} = A_{ij}B^{jk}$ denotes the matrix AB [and simultaneously the element in the position (i,k)] and $A_{ij}B^{ij} = A^{ij}B_{ij} = \text{trace}(AB^T)$. Throughout this paper we use differential geometric notations for quantities, which are connected with the expectation surface

$$E^m = \{\eta(\theta) | \theta \in \Theta\}. \tag{1.4}$$

Generalizing the notation of the Fisher information matrix we introduce

$$F_{i_1 \ldots i_k} = \left(\frac{\partial^k}{\partial \theta^{i_1} \ldots \partial \theta^{i_k}} \eta^a(\theta)\right)^{a=1,\ldots,n},$$

$$\Pi_{(i_1,\ldots,i_k)(j_1,\ldots,j_\ell)} = \langle F_{i_1,\ldots,i_k}, F_{j_1,\ldots,j_\ell} \rangle,$$

and more generally we define

$$\Pi_{(\alpha_1)(\alpha_2)\ldots(\alpha_k)} = \frac{1}{n} \sum_{a=1}^{n} F_{(\alpha_1)}^a F_{(\alpha_2)}^a \cdots F_{(\alpha_k)}^a,$$

where for $s = 1, \ldots, k$

$$\alpha_s = (i_1, \ldots, i_{r_s}) \in \mathbb{N}_0^{r_s}$$

is a multi index with $|\alpha_s| = \sum_{j=1}^{r_s} i_j$. We denote by

$$T^m(\hat{\theta}_n) = \{x \in \mathbb{R}^n \,|\, x = \hat{\eta} + \hat{F}t, t \in \mathbb{R}^m\} \tag{1.5}$$

the tangent space of the expectation surface E^m at the point $\hat{\eta} = \eta(\hat{\theta}_n)$, where the matrix $\hat{F} = F(\hat{\theta}_n)$ is defined in (1.2) [note that with this notation $T^m(\theta)$ is the tangent space of E^m at the point $\eta(\theta)$]. It is worthwhile to mention that the matrix $\omega_{ij} = \sigma^{-2}\Pi_{(i)(j)}$ is the metric tensor of the expectation surface E^m at the point θ and that the matrix

$$M_{ij} = \Pi_{(i)(j)}$$

is proportional to the Fisher information matrix, provided that some conditions of regularity are satisfied [see Borovkov (1998)].

For a fixed $\theta_0 \in \Theta$ statistical tests of the hypotheses

$$H_0 : \theta = \theta_0 \quad \text{versus} \quad H_1 : \theta \neq \theta_0$$

are usually based on statistics $T_n = T_n(y^1, \ldots, y^n; \theta_0)$ which are quadratic functionals of the observations and converge weakly to a χ^2_m distribution under the null hypothesis. If $\chi^2_{1-\alpha}(m)$ denotes the $(1-\alpha)$-quantile of the (central) χ^2_m-distribution, and the hypothesis H_0 is rejected if $T_n > \chi^2_{1-\alpha}(m)$, then under the null hypothesis

$$P^n_{\theta_0}\Big\{T_n < \chi^2_{1-\alpha}(m)\Big\} = 1 - \alpha + o(1) \tag{1.6}$$

as $n \to \infty$. Similary, a confidence region for the unknown parameter θ is obtained from the acceptance regions of the above test and given by

$$C_{1-\alpha} = \{\theta \in \Theta \mid T_n(\theta) \leq \chi^2_{1-\alpha}(m)(1 + \Delta_n n^{-1})\}\ , \tag{1.7}$$

where $\Delta_n = \Delta_n(\theta)$ denotes Bartlett's adjustment, which eliminates the term of order n^{-1} in the asymptotic expansion, that is

$$P^n_{\theta_0}\Big\{\theta_0 \in C_{1-\alpha}\Big\} = 1 - \alpha + o(n^{-1}). \tag{1.8}$$

It is the purpose of the present paper to present unified second order asymptotic expansions for a wide class of quadratic statistics, which can be used for the construction of tests and confidence regions as indicated above. To be precise, we define Ψ_n as the class of all quadratic statistics (weakly convergent to a χ^2_m-distribution) $T_n : \mathbb{R}^n \to \mathbb{R}$, which admit at the point $\theta_0 \in \Theta$ the stochastic expansion

$$T_n = T_{0n} + T_{1n}n^{-1/2} + T_{2n}n^{-1} + \xi n^{-3/2}. \tag{1.9}$$

Here T_{0n}, T_{1n}, T_{2n} are statistics which will be specified below and T_{1n}, T_{2n} can be represented in the form

$$\begin{aligned} T_{1n} &= \frac{1}{\sigma^2}\sum_{i=1}^{2} \alpha_i A_i, \\ T_{2n} &= \frac{1}{\sigma^2}\sum_{j=1}^{9} \beta_j B_j, \end{aligned} \tag{1.10}$$

where the random variables A_i and B_j are sums of the i.i.d. errors e^a in the model (1.1) and $\xi = \xi(\theta)$ denotes a random variable with the property

$$\sup_{\theta \in Q} P_\theta^n \{|\xi| > c(\log n)^{5/2}\} = o(n^{-3/2}) \tag{1.11}$$

for some compact subset $Q \subset \Theta$ and a constant $c > 0$. The constants α_i, β_j are independent of n, characterize the statistic T_n and are called *structural parameters* of the statistic T_n. Note that the statistic T_n can depend on the least squares estimate $\hat{\theta}_n$, i.e. $T_n = T_n(\hat{\theta}_n, \theta)$, and the constant c in (1.11) is usually different for any statistic under consideration. It is demonstrated in Section 6 that (to the knowledge of the authors) the class Ψ_n contains all quadratic statistics proposed in the literature for testing the hypothesis $H_0 : \theta = \theta_0$ in the nonlinear regression model (1.1).
Asymptotic expansions for several quadratic statistics were recently derived by Grigoriev and Ivanov (1992, 1993) and Ivanov (1997). However these results can not be used to cover all statistics included in the class Ψ_n considered in this paper . For example the statistic of Hamilton, Watts and Bates (1982) considered in Section 6 cannot be treated by the results obtained in Grigoriev and Ivanov (1992, 1993) and Ivanov (1997). One goal of this paper is to apply the method of virtual coefficients introduced by Grigoriev and Ivanov (1993) to obtain asymptotic expansions for all statistics in the class Ψ_n which is substantially larger than the class considered by Ivanov (1997). A further purpose of the paper is to use these results for the characterization of certain properties (such as invariance, u_n-representability) of the statistics in the class Ψ_n.
The remaining part of this paper is organized as follows. In Section 2 we introduce a virtual vector v_n and investigate its asymptotic distribution. The asymptotic expansion for the distribution of quadratic statistics in the class Ψ_n is given in Section 3, while Section 4 introduces several concepts of statistical invariance. These concepts are used in Section 5 to study properties of the quadratic statistics in the class Ψ_n. Finally, Section 6 illustrates these methods in several examples. Note that our approach is similar to the method used by Grigoriev and Ivanov (1993) but can be applied to a substantially larger class of statistics for inference in the nonlinear model. To the knowledge of the authors the examples discussed in Section 6 included all quadratic statistics, which have been proposed in the literature so far in this context .

2 Asymptotic expansions for the distribution of the virtual vector

Let $u_n = \sqrt{n}(\hat{\theta}_n - \theta)$ denote an affine transformation of the least squares estimator, which admits the stochastic expansion [see Ivanov and Zwanzig (1983)]

$$u_n = h_{0n} + h_{1n} n^{-1/2} + h_{2n} n^{-1} + h_{3n} n^{-3/2} \ , \tag{2.1}$$

where

$$\sup_{\theta \in Q} P_\theta^n \{|h_{3n}| > c(\log n)^{5/2}\} = o(n^{-3/2})$$

[see also formula (1.11)] and $h_{\nu n} = (h^i_{\nu n})^{i=1,\dots,m}$, $\nu = 0, 1, 2$, are homogeneous vector polynomials of degree $\nu + 1$ in the random variables

$$b_{i_1 \dots i_k} = \langle e, F_{i_1 \dots i_k} n^{1/2} \rangle \tag{2.2}$$

$(i_1 + \ldots + i_k = \nu + 1)$ with coefficients uniformly bounded in n. We recall the definition of the Fisher information matrix $M = \Pi_{(i)(j)}$, denote its inverse by $\Lambda^{ij} = M^{-1}$, introduce random variables

$$x^i_{i_1 \dots i_k} = \Lambda^{i\alpha} b_{\alpha i_1 \dots i_k}, \quad k \geq 0 \tag{2.3}$$

and define the quantities A_i, B_j in the representation (1.10) as follows:

$$\begin{aligned}
A_1 &= \Pi_{(i)(j)} x^i_k x^j x^k, \\
A_2 &= \Pi_{(i)(jk)} x^i x^j x^k, \\
B_1 &= \Pi_{(i)(j)} x^i_{k\ell} x^j x^k x^\ell, \\
B_2 &= \Pi_{(i)(j)} x^i_\ell x^j x^\ell_k x^k, \\
B_3 &= \Pi_{(i)(k\ell)} x^i_j x^j x^k x^\ell, \\
B_4 &= \Pi_{(ik)(\ell)} x^i_j x^j x^k x^\ell, \\
B_5 &= \Pi_{(ik)(\alpha)} \Pi_{(j\beta)(\ell)} \Lambda^{\alpha\beta} x^i x^j x^k x^\ell, \\
B_6 &= \Pi_{(i\alpha)(k)} \Pi_{(j\beta)(\ell)} \Lambda^{\alpha\beta} x^i x^j x^k x^\ell, \\
B_7 &= \Pi_{(ik)(\alpha)} \Pi_{(j\ell)(\beta)} \Lambda^{\alpha\beta} x^i x^j x^k x^\ell, \\
B_8 &= \Pi_{(ij)(k\ell)} x^i x^j x^k x^\ell, \\
B_9 &= \Pi_{(i)(jk\ell)} x^i x^j x^k x^\ell \,.
\end{aligned} \tag{2.4}$$

Note the difference to Grigoriev and Ivanov (1993) who did not consider each term of the sum $\beta_5 B_5 + \beta_6 B_6 + \beta_7 B_7$ with a separate factor, but used $\beta_5 = \beta_6 = 4\beta_7$ for their asymptotic analysis. Thus our results contain the results of these authors as a special case and we will demonstrate in Section 6 that there are important situations where the asymptotic analysis of this paper is applicable in contrast to the expansion derived by Grigoriev and Ivanov (1993) [e.g. the statistics introduced by Hamilton, Watts and Bates (1982), Pazman (1992) and Grigoriev (1994)]. Moreover, it is demonstrated in Section 6 that all quadratic statistics considered in the literature for the problem of testing simple hypotheses or for the problem of constructing confidence regions in the nonlinear model belong to the class Ψ_n defined in Section 1, i.e. they admit a stochastic expansion of the form (1.9) and (1.10).
Following Grigoriev (1994) or Ivanov (1997) the terms $h^i_{\nu n}$ in the expansion (2.1) can be rewritten as

$$\begin{aligned}
h^i_{0n} &= x^i, \\
h^i_{1n} &= x^i_j x^j - \frac{1}{4} a_{\alpha ij} \Lambda^{i\alpha} x^j x^k, \\
h^i_{2n} &= \frac{1}{2} x^i_{jk} x^j x^k + x^i_j x^j_k x^k - \frac{1}{12} a_{\alpha jk\ell} \Lambda^{i\alpha} x^j x^k x^\ell,
\end{aligned} \tag{2.5}$$

$$
\begin{aligned}
&-\frac{1}{4}a_{\alpha jk}\Lambda^{\alpha\beta}x^i_\beta x^j x^k - \frac{1}{2}a_{\alpha jk}\Lambda^{i\alpha}x^j_\ell x^k x^\ell, \\
&+\frac{1}{8}a_{\alpha\beta j}a_{\gamma k\ell}\Lambda^{i\alpha}\Lambda^{\beta\gamma}x^j x^k x^\ell,
\end{aligned}
$$

where the quantities a_{ijk} and $a_{ijk\ell}$ are defined by

$$
\begin{aligned}
a_{ijk} &= 2(\Pi_{(i)(jk)} + \Pi_{(j)(ik)} + \Pi_{(k)(ij)}), \\
a_{ijk\ell} &= 2\Big(\Pi_{(i)(jk\ell)} + \Pi_{(j)(ik\ell)} + \Pi_{(k)(ij\ell)} + \Pi_{(\ell)(ijk)} + \Pi_{(ij)(k\ell)} + \Pi_{(ik)(j\ell)} + \Pi_{(i\ell)(jk)}\Big).
\end{aligned}
\tag{2.6}
$$

The following definition was introduced by Dette and Grigoriev (2000) and characterizes a broad class of quadratic functionals.

Definition 2.1. *A statistic T_n is called u_n-representable at the point θ if it admits a stochastic asymptotic expansion of the form*

$$
T_n = T_n' + o_p(n^{-1}), \tag{2.7}
$$

where

$$
T_n' = T_{0n} + T_{1n}n^{-1/2} + T_{2n}n^{-1}
$$

and

$$
\begin{aligned}
T_{0n} &= \Pi_{(i)(j)}u^i_n u^j_n, \\
T_{1n} &= c_1\Pi_{(i)(jk)}u^i_n u^j_n u^k_n, \\
T_{2n} &= \Big(c_2\Pi_{(ij)(k\ell)} + c_3\Pi_{(i)(jk\ell)} + c_4\Lambda^{rs}\Pi_{(r)(ij)}\Pi_{(s)(k\ell)}\Big)u^i_n u^j_n u^k_n u^\ell_n.
\end{aligned}
$$

The vector $c = (c_i)_{i=1}^4 \in \mathbb{R}^4$ characterizes the u_n-representable statistic T_n. The coefficients of the vector c are called structural coefficients of the u_n-representable statistic T_n.

Observing the equations (2.1) - (2.5) it is straightforward to show that any u_n-representable statistic also admits a stochastic expansion of the form (1.9) and (1.10) and therefore belongs to the class Ψ_n defined in Section 1. However, the converse inclusion is not true. A typical example is the statistic of Neyman and Pearson which admits an expansion of the form (1.9) but is not u_n-representable [see the statistic $T_n^{(1)}$ in Section 6 for more details].
Ivanov and Zwanzig (1983) modified techniques of Pfanzagl (1973) and Michel (1975) to obtain the (asymptotic) distribution of the least squares estimator in nonlinear regression models. These expansions can be used for the determination of the (asymptotic) distribution of u_n-representable statistics [see e.g. Bardadym and Ivanov (1985) or Dette and Grigoriev (2000)]. In order to solve this problem for non u_n-representable statistics Grigoriev and Ivanov (1992, 1993) proposed a technique of virtual expansions and used this method to derive asymptotic expansions for the Neyman-Pearson, Wald- und Kullback-Leibler statistic [note that these statistics admit an

asymptotic expansion of the form (1.9) with $\beta_5 = \beta_6 = 4\beta_7$; see Ivanov (1997)]. However, for the general class Ψ_n of statistics defined by (1.9) this method has to be modified. To this end we introduce a general v_n-vector

$$v_n = \sum_{\nu=0}^{2} \tilde{h}_{\nu n} n^{-\nu/2} + o_p(n^{-1}), \tag{2.8}$$

where (for appropriate constants $\pi_1, \pi_2, \rho_1, \ldots, \rho_{16}$)

$$\begin{aligned}
\tilde{h}^i_{0n} &= x^i, \\
\tilde{h}^i_{1n} &= \pi_1 x^i_j x^j + \pi_2 a_{\alpha jk} \Lambda^{i\alpha} x^j x^k, \\
\tilde{h}^i_{2n} &= (\rho_1 x^i_{\alpha\beta} x^\beta + \rho_2 x^i_\beta x^\beta_\alpha) x^\alpha, \\
&+ \Big(\rho_3 \Pi_{(s)(\alpha\beta)} \Lambda^{js} x^i_j + \rho_4 \Pi_{(j)(\alpha r)} \Lambda^{ir} x^j_\beta, \\
&+ \rho_5 \Pi_{(r)(j\alpha)} \Lambda^{ir} x^j_\beta + \rho_6 \pi_{(\alpha)(s\beta)} \Lambda^{js} x^i_j + \rho_7 \Pi_{(\alpha)(rj)} \Lambda^{ir} x^j_\beta\Big) x^\alpha x^\beta, \\
&+ \Big(\rho_8 \Pi_{(r)(j\gamma)} \Pi_{(s)(\alpha\beta)} + \rho_9 \Pi_{(j)(r\gamma)} \Pi_{(\alpha)(s\beta)} + \rho_{10} \Pi_{(\gamma)(rj)} \Pi_{(s)(\alpha\beta)}, \\
&+ \rho_{11} \Pi_{(r)(j\gamma)} \Pi_{(\alpha)(s\beta)} + \rho_{12} \Pi_{(\gamma)(rj)} \Pi_{(\alpha)(s\beta)} + \rho_{13} \Pi_{(j)(r\gamma)} \Pi_{(s)(\alpha\beta)}\Big) \Lambda^{ir} \Lambda^{js} x^\alpha x^\beta x^\gamma, \\
&+ \Big(\rho_{14} \Pi_{(r\alpha)(\beta\gamma)} + \rho_{15} \Pi_{(r)(\alpha\beta\gamma)} + \rho_{16} \Pi_{(\alpha)(r\beta\gamma)}\Big) \Lambda^{ir} x^\alpha x^\beta x^\gamma.
\end{aligned}$$

Observing the representation (2.5) a straightforward calculation shows that for the choice

$$\pi_1 = 1,\ \pi_2 = -\tfrac{1}{4},\ \rho_1 = \frac{1}{2}, \tag{2.9}$$

$$\begin{aligned}
\rho_2 &= 1, & \rho_3 &= -\tfrac{1}{2}, & \rho_4 &= -1, \\
\rho_5 &= -1, & \rho_6 &= -1, & \rho_7 &= -1, \\
\rho_8 &= \tfrac{1}{2}, & \rho_9 &= 1, & \rho_{10} &= \tfrac{1}{2}, \\
\rho_{11} &= 1, & \rho_{12} &= 1, & \rho_{13} &= \tfrac{1}{2}, \\
\rho_{14} &= -\tfrac{1}{2}, & \rho_{15} &= -\tfrac{1}{6}, & \rho_{16} &= -\tfrac{1}{2},
\end{aligned}$$

the virtual vector v_n defined in (2.1) and (2.8) coincides with the vector u_n defined in (2.5). Moreover, it is of independent interest to compare this vector with the virtual vector introduced by Grigoriev and Ivanov (1993) or Ivanov (1997), which contains the virtual coefficients $(\pi_1, \pi_2, \tau_1, \ldots, \tau_6)$ and is given by

$$\begin{aligned}
\tilde{h}_{0n} &= x^i, \\
\tilde{h}^i_{1n} &= \pi_1 x^i_j x^j + \pi_2 a_{\alpha jk} \Lambda^{i\alpha} x^j x^k, \\
\tilde{h}^i_{2n} &= \tau_1 x^i_{jk} x^j x^k + \tau_2 x^i_j x^j_k x^k + \tau_3 a_{\alpha jk\ell} \Lambda^{i\alpha} x^j x^k x^\ell \\
&+ \tau_4 a_{\alpha jk} \Lambda^{\alpha\beta} x^i_\beta x^j x^k + \tau_5 a_{\alpha jk} \Lambda^{i\alpha} x^j_\ell x^k x^\ell \\
&+ \tau_6 a_{\alpha\beta j} a_{\gamma k\ell} \Lambda^{i\alpha} \Lambda^{\beta\gamma} x^j x^k x^\ell,
\end{aligned} \tag{2.10}$$

where $a_{ijk\ell}$ and a_{ijk} are defined in (2.6).

The common procedure of deriving asymptotic expansions for the distribution of the statistic $T_n = T_n' + o_p(n^{-1})$ consists of two steps [see e.g. Ivanov (1997)]. In a first step one derives (assuming appropriate conditions of regularity) the estimates

$$P_\theta^n\Big\{T_n' < z - \delta_n\Big\} + o(n^{-3/2}) \le P_\theta^n\Big\{T_n < z\Big\} \le P_\theta^n\Big\{T_n' < z + \delta_n\Big\} + o(n^{-3/2}) \tag{2.11}$$

uniformly with respect to $\theta \in Q$, where $Q \subset \Theta$ denotes a compact set specified below, and the sequence δ_n is given by $\delta_n = cn^{-3/2}(\log n)^{5/2}$ [see equation (1.11)]. In a second step we use that the statistic T_n' is v_n-representable and calculate the probability

$$P_\theta^n\Big\{T_n' < z \pm \delta_n\Big\} = \int I\Big\{x \mid T_n'(\theta + n^{-1/2}x) < z \pm \delta_n\Big\} dF_n(x) \ , \tag{2.12}$$

where F_n denotes the distribution function of the statistic v_n. The asymptotic expansion is now obtained from (2.11) and (2.12) by an expansion of the integral. This procedure can be applied for any statistic $T_n = T_n' + o_p(n^{-1})$, where T_n' can be expressed as a function of v_n. Substituting the expansion (2.8) for v_n in a representation of the form (2.7) we obtain a stochastic expansion of the form (1.9) and (1.10), where the structural coefficients in (1.10) are given by

$$\begin{aligned}
\alpha_1 &= 2\pi_1, \\
\alpha_2 &= 12\pi_2 + c_1, \\
\beta_1 &= 2\rho_1, \\
\beta_2 &= 2\rho_2 + \pi_1^2, \\
\beta_3 &= 2(\rho_3 + \rho_4) + 4\pi_1\pi_2 + \pi_1 c_1, \\
\beta_4 &= 2(\rho_5 + \rho_6 + \rho_7) + 8\pi_1\pi_2 + 2\pi_1 c_1, \\
\beta_5 &= 2(\rho_8 + \rho_9 + \rho_{10}) + 16\pi_2^2 + 8\pi_2 c_1, \\
\beta_6 &= 2(\rho_{11} + \rho_{12}) + 16\pi_2^2 + 8\pi_2 c_1, \\
\beta_7 &= 2\rho_{13} + 4\pi_2^2 + 2\pi_2 c_1 + c_4, \\
\beta_8 &= 2\rho_{14} + c_2, \\
\beta_9 &= 2(\rho_{15} + \rho_{16}) + c_3.
\end{aligned} \tag{2.13}$$

The representation (2.13) has two important consequences. On the one hand it follows that for any statistic $T_n \in \Psi_n$ there exists a vector v_n with virtual coefficients such that (2.13) holds. Consequently an asymptotic expansion for the distribution of any statistic in the class Ψ_n is available, which depends only on the virtual coefficients $\pi_1, \pi_2, \rho_1, \ldots, \rho_{16}$. Secondly, using the equations (2.13) we can express these coefficients in terms of the structural parameters $\alpha_1, \alpha_2, \beta_1, \ldots, \beta_9$ and consequently the asymptotic expansion of any statistic T_n in the class Ψ_n can be specified in terms of its corresponding structural parameters. Note that the technique of virtual coefficients introduced in this paper can be applied to a broader class of statistics as considered in Grigoriev and Ivanov (1993). The main result of this section gives an asymptotic expansion for the asymptotic distribution of the statistic v_n defined in (2.8). To be precise let C^m denote the class of all convex Borel sets of $\mathbb{R}^m$ and define $\varphi(\cdot;\Sigma)$ as the density of an $m-$dimensional multivariate normal distribution with mean zero and covariance matrix Σ.

Theorem 2.2: *If the assumptions I — VIII in Grigoriev and Ivanov (1993) are satisfied for a compact set $Q \subset \Theta$, then*

$$\sup_{\theta\in Q}\sup_{C\in C^m}\left|P_\theta^n\Big\{v_n(\theta)\in C\Big\} - \int_C \varphi(y;\sigma^2\Lambda(\theta))\{1+M_{1n}(\theta;y)n^{-1/2}+M_{2n}(\theta;y)n^{-1}\}dy\right|$$

$$= O\Big(\frac{\log^2 n}{n^{3/2}}\Big)\ , \tag{2.14}$$

where the random variables $M_{\nu n}(\theta;y)$ are polynomials in $y=(y^1,\dots,y^m)$ of degree 3ν $(\nu=1,2)$ with coefficients uniformly bounded with respect to $\theta\in Q$, $n\in\mathbb{N}$ and explicitly given in the Appendix.

The proof of Theorem 2.2 is omitted because it proceeds along the lines of Linnik and Mitrofanova (1963), Pfanzagl (1973), Chibishov (1973) and Michel (1975) and is similar to the one given in Ivanov and Zwanzig (1983) (note that in contrast to the first named authors we have to use results for asymptotic expansions for sums of non i.i.d. random variables). The polynomials M_{1n} and M_{2n} are calculated similary as in Michels (1975) [see also Ivanov (1997)] and are given in the Appendix for the sake of completeness. Note that in general the polynomial M_{2n} consists of 56 sums and differs from the one obtained in Grigoriev and Ivanov (1993). A slight simplification of M_{2n} is obtained in the case $v_n=u_n$, where M_{2n} contains only 40 sums. Similary, in the case of a symmetric error distribution the polynomial M_{2n} reduces to 30 sums.

3 A unified asymptotic expansion for the distribution of quadratic statistics in the class Ψ_n

In order to provide asymptotic expansions for the distribution of statistics in the class Ψ_n we introduce 16 functions $P_j=P_j(\theta)$ defined by

$$\begin{aligned} P_1 &= \frac{\gamma_4}{\sigma^4}\Lambda^{is}\Lambda^{jr}\Pi_{(i)(j)(r)(s)}, \\ P_2 &= \frac{\gamma_3^2}{\sigma^6}\Lambda^{is}\Lambda^{jr}\Lambda^{kt}\Pi_{(i)(k)(t)}\Pi_{(s)(j)(r)}, \\ P_3 &= \frac{\gamma_3^2}{\sigma^6}\Lambda^{is}\Lambda^{jr}\Lambda^{kt}\Pi_{(i)(j)(k)}\Pi_{(s)(r)(t)}, \\ P_4 &= \frac{\gamma_3}{\sigma^2}\Lambda^{is}\Lambda^{jr}\Pi_{(is)(j)(r)}, \end{aligned} \tag{3.1}$$

$$
\begin{aligned}
P_5 &= \frac{\gamma_3}{\sigma^2}\Lambda^{is}\Lambda^{jr}\Pi_{(ij)(r)(s)}, \\
P_6 &= \frac{\gamma_3}{\sigma^2}\Lambda^{is}\Lambda^{jr}\Lambda^{kt}\Pi_{(i)(jr)}\Pi_{(k)(i)(s)}, \\
P_7 &= \frac{\gamma_3}{\sigma^2}\Lambda^{is}\Lambda^{jr}\Lambda^{kt}\Pi_{(t)(rs)}\Pi_{(k)(i)(j)}, \\
P_8 &= \frac{\gamma_3}{\sigma^2}\Lambda^{is}\Lambda^{jr}\Lambda^{kt}\Pi_{(t)(rk)}\Pi_{(i)(j)(s)}, \\
P_9 &= \sigma^2\Lambda^{is}\Lambda^{jr}\Lambda^{kt}\Pi_{(kt)(j)}\Pi_{(ir)(s)}, \\
P_{10} &= \sigma^2\Lambda^{is}\Lambda^{jr}\Lambda^{kt}\Pi_{(ik)(j)}\Pi_{(rt)(s)}, \\
P_{11} &= \sigma^2\Lambda^{is}\Lambda^{jr}\Lambda^{kt}\Pi_{(i)(js)}\Pi_{(k)(rt)}, \\
P_{12} &= \sigma^2\Lambda^{is}\Lambda^{jr}\Lambda^{kt}\Pi_{(i)(jr)}\Pi_{(s)(kt)}, \\
P_{13} &= \sigma^2\Lambda^{is}\Lambda^{jr}\Lambda^{kt}\Pi_{(i)(jk)}\Pi_{(s)(rt)}, \\
P_{14} &= \sigma^2\Lambda^{is}\Lambda^{jr}\Pi_{(is)(jr)}, \\
P_{15} &= \sigma^2\Lambda^{is}\Lambda^{jr}\Pi_{(ij)(rs)}, \\
P_{16} &= \sigma^2\Lambda^{is}\Lambda^{jr}\Pi_{(i)(jrs)},
\end{aligned}
$$

where γ_3 and γ_4 denote the third and fourth cumulant of the error distribution, respectively. The proof of the following Theorem 3.1 is similar to the one given in Grigoriev and Ivanov (1993) [see also Bardadym and Ivanov (1985), Ivanov (1997) and Dette and Grigoriev (2000), who obtained similar results for a smaller class of quadratic statistics].

Theorem 3.1: *If the assumptions of Theorem 2.2 are satisfied, then for any $z_0 > 0$ (in the case $m = 1$) and any $z_0 = 0$ (in the case $m > 1$) as $n \to \infty$*

$$
(3.2) \qquad \sup_{\theta\in Q}\sup_{z\geq z_0}\Big|P_\theta^n\Big\{T_n < z\Big\} - G_m(z) - \frac{1}{n}\sum_{j=0}^{3}\sum_{k=1}^{16}\lambda_{jk}P_k(\theta)G_{m+2j}(z)\Big| = O\Big(\frac{(\log n)^{5/2}}{n^{3/2}}\Big),
$$

where T_n is an arbitrary statistic in the class Ψ_n, the functions $P_k(\theta)$ are defined in (3.1) and uniformly bounded with respect to $\theta \in Q$ and $n \in \mathbb{N}$, G_r denotes the cumulative distribution function of the central χ_r^2-distribution with r degrees of freedom and the coefficients

$$
\lambda_{jk} = \lambda_{jk}(\alpha_1, \alpha_2, \beta_1, \ldots, \beta_9)
$$

characterize the specific statistic under consideration and are given in Table 3.1.

Remark 3.2: As indicated in Section 2 the coefficients λ_{jk} in the expansion (3.2) can also be expressed in terms of the virtual parameter $\pi_1, \pi_2, \rho_1, \ldots, \rho_{16}$. However, these representations are tedious and omitted for the sake of brevity.

Remark 3.3: It is worthwhile to mention two remarkable properties of the coefficients in the expansion (3.2). From Table 3.1 it is easy to see that the coefficients λ_{jk} satisfy the equations

$$\sum_{j=0}^{3} \lambda_{jk} = 0, \qquad k = 1, \ldots, 16. \tag{3.3}$$

Consequently, if we denote

$$a_{jn}(\theta) = \sum_{k=1}^{16} \lambda_{jk} P_k(\theta), \qquad j = 0, \ldots, 3, \tag{3.4}$$

then it follows that

$$\sum_{j=0}^{3} a_{jn}(\theta) = 0. \tag{3.5}$$

We finally note that these identities can be alternatively obtained by similar arguments as given in Chandra and Ghosh (1979) and that Theorem 3.1 and Table 3.1 extend the analoguous result and table given in Grigoriev and Ivanov (1993) and Ivanov (1997) to all statistics with a stochastic expansion of the form (1.9) and (1.10).

Remark 3.4: If the statistic T_n is u_n-representable then the coefficients λ_{jk} in the asymptotic expansion (3.2) depend on the structural parameters only through the coefficients $c_1, \ldots, c_4$ in the expansion (2.7) and are given in Table 3.2. This follows easily from Table 3.1 by substituting in the equations (2.13) the results from (2.9), which gives

$$\begin{aligned}
&\alpha_1 = 2, \qquad \alpha_2 = c_1 - 3, \\
&\beta_1 = 1, \qquad \beta_2 = 3, \quad \beta_3 = c_1 - 4, \\
&\beta_4 = 2(c_1 - 4), \qquad \beta_5 = \beta_6 = 5 - 2c_1, \\
&\beta_7 = \tfrac{1}{4}(5 - 2c_1) + c_4, \; \beta_8 = c_2 - 1, \quad \beta_9 = c_3 - \frac{4}{3}.
\end{aligned} \tag{3.6}$$

Recalling the definition of the coefficients $a_{jn} = a_{jn}(\theta)$ in (3.4) we obtain from Theorem 3.1, (3.5) and (3.4) the representation

$$P_\theta^n\Big\{T_n \geq t\Big\} = \bar{G}_m(t) + \frac{1}{n}\sum_{j=0}^{3} a_{jn}(\theta)\bar{G}_{m+2j}(t) + o(n^{-1}), \tag{3.7}$$

where $\bar{G}_r(t) = 1 - G_r(t)$. If g_r denotes the density of the (central) χ_r^2-distribution we can use the well known identity [see Abramowitz and Stegun (1964)]

$$\bar{G}_{r+2}(t) = \bar{G}_r(t) + 2g_{r+2}(t) \tag{3.8}$$

to express the second term in the expansion (3.7) in terms of the densities $g_j(t)$. To this end we note that (3.8) implies

$$\bar{G}_{m+2k}(t) = \bar{G}_m(t) + 2\sum_{j=1}^{k} g_{m+2j}(t), \qquad k \geq 0 \tag{3.9}$$

and obtain from (3.7) and (3.4)

$$P_\theta^n\Big\{T_n \geq t\Big\} \;=\; \bar{G}_n(t) + \frac{2}{n}\sum_{j=1}^{3} B_{jn}(\theta) g_{m+2j}(t) + o(n^{-1}). \tag{3.10}$$

k/j	0	1	2	3
1	$\frac{1}{8}$	$-\frac{1}{4}$	$\frac{1}{8}$	
2	$-\frac{1}{8}$	$\frac{3}{8}$	$-\frac{3}{8}$	$\frac{1}{8}$
3	$-\frac{1}{12}$	$\frac{1}{4}$	$-\frac{1}{4}$	$\frac{1}{12}$
4	$\frac{1}{4}\alpha_1$	$-\frac{1}{2}\alpha_1$	$\frac{1}{4}\alpha_1$	
5		$-\frac{1}{2}\alpha_1$	$\frac{1}{2}\alpha_1$	
6	$-\frac{1}{4}\alpha_1$	$\frac{1}{4}(3\alpha_1+\alpha_2)$	$-\frac{1}{4}(3\alpha_1+2\alpha_2)$	$\frac{1}{4}(\alpha_1+\alpha_2)$
7		$\frac{1}{2}\alpha_1$	$-\alpha_1-\frac{1}{2}\alpha_2$	$\frac{1}{2}(\alpha_1+\alpha_2)$
8		$\frac{1}{2}(\alpha_1+\alpha_2)$	$-(\alpha_1+\alpha_2)$	$\frac{1}{2}(\alpha_1+\alpha_2)$
9		$\frac{1}{2}(\alpha_1+\alpha_2)^2$ $-\frac{1}{2}(\beta_3+\beta_5)$	$-(\alpha_1+\alpha_2)^2$ $+\frac{1}{2}(\beta_3+\beta_5)$	$\frac{1}{2}(\alpha_1+\alpha_2)^2$
10		$\frac{1}{2}(\alpha_1+\alpha_2)^2$ $-(\beta_3+\beta_5)$ $-\frac{1}{2}(\beta_2+\beta_4+\beta_6)$	$-(\alpha_1+\alpha_2)^2$ $+(\beta_3+\beta_5)$ $+\frac{1}{2}(\beta_2+\beta_4+\beta_6)$	$\frac{1}{2}(\alpha_1+\alpha_2)^2$
11		$\frac{1}{2}(\alpha_1+\alpha_2)^2$ $-\frac{1}{2}(\beta_2+\beta_4+\beta_6)$	$-(\alpha_1+\alpha_2)^2$ $+\frac{1}{2}(\beta_2+\beta_4+\beta_6)$	$\frac{1}{2}(\alpha_1+\alpha_2)^2$
12	$-\frac{1}{8}\alpha_1^2$	$\frac{1}{8}(\alpha_1+\alpha_2)^2$ $+\frac{1}{4}\alpha_1^2-\frac{1}{2}\beta_7$	$-\frac{1}{4}(\alpha_1+\alpha_2)^2$ $-\frac{1}{8}\alpha_1^2+\frac{1}{2}\beta_7$	$\frac{1}{8}(\alpha_1+\alpha_2)^2$
13	$-\frac{1}{4}\alpha_1^2+\frac{1}{2}\beta_2$	$\frac{1}{4}(\alpha_1+\alpha_2)^2+\frac{1}{2}\alpha_1^2$ $-(\beta_2+\beta_7)$ $-\frac{1}{2}(\beta_4+\beta_6)$	$-\frac{1}{2}(\alpha_1+\alpha_2)^2-\frac{1}{4}\alpha_1^2$ $+\frac{1}{2}(\beta_2+\beta_4+\beta_6)+\beta_7$	$\frac{1}{4}(\alpha_1+\alpha_2)^2$
14	$\frac{1}{8}\alpha_1^2$	$-\frac{1}{4}\alpha_1^2-\frac{1}{2}\beta_8$	$\frac{1}{8}\alpha_1^2+\frac{1}{2}\beta_8$	
15	$\frac{1}{4}\alpha_1^2-\frac{1}{2}\beta_2$	$-\frac{1}{2}\alpha_1^2+\frac{1}{2}\beta_2-\beta_8$	$\frac{1}{4}\alpha_1^2+\beta_8$	
16		$-\frac{3}{2}(\beta_1+\beta_9)$	$\frac{3}{2}(\beta_1+\beta_9)$	

Table 3.1: Coefficients λ_{jk} in the expansion (3.2) for statistics $T_n \in \Psi_n$.

Here the coefficients $B_{jn}(\theta)$ are defined by

(3.11) $$B_{jn}(\theta) = -\sum_{i=0}^{j-1} a_{in}(\theta), \qquad j = 1, 2, 3,$$

and we call these coefficients the *Bartlett functions* of the statistic T_n. Note that an inversion of (3.11) yields

(3.12) $$\begin{aligned} a_{0n} &= -B_{1n}, \quad a_{1n} = B_{1n} - B_{2n}, \\ a_{2n} &= B_{2n} - B_{3n}, \ a_{3n} = B_{3n}, \end{aligned}$$

which shows that the functions a_{jn} in the asymptotic expansion (3.7) can be expressed in terms of the Bartlett functions.

k/j	0	1	2	3
1	$\frac{1}{8}$	$-\frac{1}{4}$	$\frac{1}{8}$	
2	$-\frac{1}{8}$	$\frac{3}{8}$	$-\frac{3}{8}$	$\frac{1}{8}$
3	$-\frac{1}{12}$	$\frac{1}{4}$	$-\frac{1}{4}$	$\frac{1}{12}$
4	$\frac{1}{2}$	-1	$\frac{1}{2}$	
5		-1	1	
6	$-\frac{1}{2}$	$\frac{1}{4}(c_1+3)$	$-\frac{1}{2}c_1$	$\frac{1}{4}(c_1-1)$
7		1	$-\frac{1}{2}(c_1+1)$	$\frac{1}{2}(c_1-1)$
8		$\frac{1}{2}(c_1-1)$	$-c_1+1$	$\frac{1}{2}(c_1-1)$
9		$\frac{1}{2}(c_1^2-c_1)$	$-c_1^2+\frac{1}{2}(3c_1-1)$	$\frac{1}{2}(c_1-1)^2$
10		$\frac{1}{2}(c_1^2-1)$	$-c_1^2+c_1$	$\frac{1}{2}(c_1-1)^2$
11		$\frac{1}{2}(c_1-1)^2$	$-(c_1-1)^2$	$\frac{1}{2}(c_1-1)^2$
12	$-\frac{1}{2}$	$\frac{1}{8}(c_1^2+4-4c_4)$	$-\frac{1}{4}(c_1^2-c_1+\frac{1}{2}-2c_4)$	$\frac{1}{8}(c_1-1)^2$
13	$\frac{1}{2}$	$\frac{1}{4}(c_1^2-2-4c_4)$	$-\frac{1}{2}(c_1^2-c_1+\frac{1}{2}-2c_4)$	$\frac{1}{4}(c_1-1)^2$
14	$\frac{1}{2}$	$-\frac{1}{2}(c_2+1)$	$\frac{1}{2}c_2$	
15	$-\frac{1}{2}$	$-c_2+\frac{1}{2}$	c_2	
16		$-\frac{1}{2}(3c_3-1)$	$\frac{1}{2}(3c_3-1)$	

Table 3.2: Coefficients λ_{jk} in the representation (3.4) for u_n-representable statistics.

Finally, if we recall the definition of the Bartlett adjustment $\Delta_n = \Delta_n(\theta)$ by the equations (1.7) and (1.8), we are able to express the adjustment Δ_n in terms of the Bartlett functions B_{1n}, B_{2n} and B_{3n}. To this end define $a^2 = \chi^2_{1-\alpha}(m)$ as the $(1-\alpha)-$ quantile of the central χ^2 distribution with m degrees of freedom and
$$t = t_n = a^2(1 + \Delta_n n^{-1}),$$

then a Taylor expansion yields

$$\bar{G}_m(t) \;=\; \bar{G}_m(a^2) - \frac{g_m(a^2)}{1!} \cdot \frac{a^2\Delta_n}{n} + o(n^{-1})$$

$$g_{m+2i}(t) \;=\; g_{m+2i}(a^2) + o(1).$$

Observing (3.10) this gives

$$P_\theta^n\Big\{T_n \geq t\Big\} = \bar{G}_m(a^2) - \frac{a^2 g_m(a^2)}{n}\Big\{\Delta_n - \frac{2}{a^2}\sum_{i=1}^{3} B_{in}(\theta)\frac{g_{m+2i}(a^2)}{g_m(a^2)}\Big\} + o(n^{-1})$$

and equating coefficients yields for the Bartlett adjustment

(3.13) $$\Delta_n(\theta) = \frac{2}{a^2}\sum_{i=1}^{3}\frac{g_{m+2i}(a^2)}{g_m(a^2)}B_{ni}(\theta) = \frac{2}{a^2}\sum_{i=1}^{3}\Big(\frac{a^2}{2}\Big)^i\frac{\Gamma(\frac{m}{2})}{\Gamma(\frac{m+2i}{2})}B_{ni}(\theta)\ ,$$

where $\Gamma(\cdot)$ denotes the Gamma-function and the last equality follows from the representation of the density of the χ_r^2-distribution g_r.

4 Statistical invariants

Recall the definition of the expectation surface E^m in (1.4). An invariant of E^m is any object, which is not changed by a local transformation of the coordinates $\theta = (\theta^1, \ldots, \theta^m)^T$ [see Veblen (1933)]. For example every point $A = A(\theta^1, \ldots, \theta^m) \in E^m$ is an invariant because a new parametrization obtained by a diffeomorphism $\bar{\theta} : \Theta \to \Theta$ does not change the point A but only its coordinates. Consequently, any system of points in E^m and any function of these points define invariants of the expectation surface E^m. If such a function has the representation $I(\theta)$ in the original coordinate system, it has the form $\bar{I}(\bar{\theta})$ in the new coordinate system. Throughout this paper we call a reparametrization $\bar{\theta} : \Theta \to \Theta$ of the expectation surface regular if the reparametrizised model

(4.1) $$y = \bar{\eta}(\bar{\theta}) + e$$

satisfies the regularity assumptions I - VI of Grigoriev and Ivanov (1993). Consider the statistical model $\{\mathbb{R}^n, B^n, \bar{P}^n_{\bar{\theta}}, \bar{\theta} \in \Theta\}$ generated by the new model (4.1) and let $\bar{\Psi}_n$ denote the analogue of the class Ψ_n defined in Section 1. It is easy to see that there is a one-to-one correspondence between the sets of statistics Ψ_n and $\bar{\Psi}_n$: a statistic $T_n \in \Psi_n$ with structural coefficients $(\alpha_1, \alpha_2, \beta_1, \ldots, \beta_9)$ is simply associated with a statistic $\bar{T}_n \in \bar{\Psi}_n$ with the same structural coefficients.

Definition 4.1: *A statistic $T_n \in \Psi_n$ is called invariant with respect to a regular reparametrization $\bar{\theta} : \Theta \to \Theta$ of the expectation surface E^m if for every $\theta_0 \in \Theta$ and $\bar{\theta}_0 = \bar{\theta}(\theta_0)$*

(4.2) $$\sup_{x\in\mathbb{R}}\Big|P^n_{\theta_0}\Big\{T_n < x\Big\} - \bar{P}_{\bar{\theta}_0}\Big\{\bar{T}_n < x\Big\}\Big| = o(n^{-1}).$$

Some invariant statistics will be considered in Section 6. Roughly speaking Definition 4.1 means that the coefficients $a_{jn}(\theta)$ in the asymptotic expansions (3.2) and (3.7) can be represented in terms of differential geometric invariants. To be precise let $\omega_i = \omega_i(\theta)$ denote a covariant tensor field on E^m and consider a function $I(\omega, \theta)$ depending on a finite number of variables of the form

$$\omega^a, \frac{\partial \omega^a}{\partial \theta^i}, \frac{\partial^2 \omega^a}{\partial \theta^i \partial \theta^j} \ldots$$

(note that $I(\omega, \theta)$ is a function of the parameter θ). Following Grigoriev (1994) and Ivanov (1997) [see page 267, formula (18.69)] we call I a (scalar) differential invariant of the expectation surface E^m, if for any reparametrization $\bar{\theta} : \Theta \to \Theta$

$$I(\omega, \theta) = I(\omega, \bar{\theta}) \tag{4.3}$$

(note that $I(\omega, \theta)$ and $I(\omega, \bar{\theta})$ are the values of the function I in the different coordinate systems).

Example 4.2: Consider the inner product $\langle \cdot, \cdot \rangle$ with respect to the matrix $n^{-1}\sigma^{-2}\delta_{ab}$ on $\mathbb{R}^n$ and the projection

$$P = P^{ab} = F_i^a F_j^b \Lambda^{ij} n^{-1} \tag{4.4}$$

onto the tangent space $T^m(\theta)$ defined in (1.5). The field

$$\nabla_{ij} = \langle e, (I - P) F_{ij} \rangle$$

is called a covariant derivative of the tensor field ∂_i along of the tensor field ∂_j [see Rashevsky (1967)], where

$$\partial_i = \langle F_i, e \rangle = n^{-1/2} \sigma^{-2} b_i \tag{4.5}$$

is a stochastic basis of the tangent space $T^m(\theta)$. The matrix ∇_{ij} is a differential invariant of the expectation surface E^m and the quantities

$$\nabla_{ij} \Lambda^{ij}, \quad \nabla^{ij} \partial_i \partial_j \tag{4.6}$$

are scalar invariants of the expectation surface E^m.

Example 4.3: The Christoffel symbols of the first and second kind

$$\Gamma_{i,jk} = \frac{1}{\sigma^2} \Pi_{(i)(jk)}, \quad \Gamma^i_{jk} = \Lambda^{i\alpha} \Pi_{(\alpha)(jk)} \tag{4.7}$$

are differential invariants of the expectation surface E^m. With the notation

$$\omega_{i_1 \ldots i_k} = \frac{\gamma_k}{\sigma^{2k}} \Pi_{(i_1) \ldots (i_k)} \quad (k \geq 2)$$

the tensors of covariance, skewness and excess of the expectation surface E^m can be written as

$$
\begin{aligned}
\omega_{ij} &= \frac{1}{\sigma^2}\Pi_{(i)(j)}, \\
\omega_{ijk} &= \frac{\gamma_3}{\sigma^6}\Pi_{(i)(j)(k)}, \\
\omega_{ijk\ell} &= \frac{\gamma_4}{\sigma^8}\Pi_{(i)(j)(k)(\ell)}.
\end{aligned}
\tag{4.8}
$$

Because the quantities P_1, P_2, P_3 in (3.1) can be represented as

$$
\begin{aligned}
P_1 &= \omega^{is}\omega^{jr}\omega_{ijrs}, \\
P_2 &= \omega^{is}\omega^{jr}\omega^{kt}\omega_{ikt}\omega_{sjr}, \\
P_3 &= \omega^{is}\omega^{jr}\omega^{kt}\omega_{ijk}\omega_{srt},
\end{aligned}
\tag{4.9}
$$

it follows that P_1, P_2 and P_3 are scalar invariants of the expectation surface E^m. For the subsequent discussion in Section 5 and 6 we also introduce the scalar invariant

$$X(\alpha, \beta, \gamma) = \alpha P_1 + \beta P_2 + \gamma P_3, \tag{4.10}$$

where $\alpha, \beta, \gamma \in \mathbb{R}$ are arbitrary constants.

Example 4.4: A couple of further scalar invariants, which can be expressed in terms of the constants $P_4 - P_{16}$ in (3.1) have been discussed in the literature and will be used in following sections. The invariants

$$Q_1 = P_4 - P_6, \tag{4.11}$$

$$Q_2 = P_5 - P_7, \tag{4.12}$$

have been found by Grigoriev and Ivanov (1993). The quantity

$$H = P_{14} - P_{12} \tag{4.13}$$

is called Efron's statistical curvature [see Efron (1975)]. Finally, the Ricci curvature is defined as

$$R = (P_{14} - P_{12}) - (P_{15} - P_{13}) \tag{4.14}$$

For $m = \dim E^m = 1$ Ricci's curvature R is always zero. In contrast to Ricci's curvature Efron's curvature H is not necessarily zero in the one-dimensional case. By reason of its definitions Efron's curvature is non-negative, i.e. $H \geq 0$, but Ricci's curvature can have both signs. Efron's and Ricci's curvatures of the surface E^m are related by the inequality [see Grigoriev (1994) and Ivanov (1997)]

$$H \geq R. \tag{4.15}$$

The quantity

$$B = 3H - 2R \tag{4.16}$$

is called Beale's measure of intrinsic nonlinearity of the expectation surface E^m [see Beale (1960)]. From (4.15) and (4.16) we obtain $B \geq 0$. Finally, recall the definition of the normal curvature of McCullagh and Cox (1986)

$$Y = H - 2R. \tag{4.17}$$

It will be demonstrated in the following section that the invariants X, Q_1, Q_2, H and R (or equivalently B and Y) describe the coefficients in the asymptotic expansion (3.7) for the distribution of an invariant statistic $T_n \in \Psi_n$ completely.

5 Invariant and u_n-representable statistics

In this section we characterize the properties of invariance and u_n-representability of a statistic $T_n \in \Psi_n$ in terms of its structural coefficients $\alpha_1, \alpha_2, \beta_1, \ldots, \beta_9$.

Theorem 5.1: *A statistic $T_n \in \Psi_n$ is invariant with respect to a regular reparametrization if its structural coefficients in the representation (1.9) satisfy*

$$\begin{aligned} \alpha_1 + \alpha_2 &= 0, \\ \beta_1 + \beta_9 &= 0, \\ \beta_2 + \beta_4 + \beta_6 &= 0, \\ \beta_3 + \beta_5 &= 0, \\ \beta_7 + \beta_8 &= 0. \end{aligned} \tag{5.1}$$

Proof: If the equations in (5.1) are satisfied, then we obtain from Table 3.1 and the representations (4.11), (4.13) and (4.14) for the coefficients $a_{jn} = a_{jn}(\theta)$ in the expansion (3.7)

$$\begin{aligned} a_{0n} &= \frac{1}{8}(\alpha_1^2 - \beta_2)(3H - 2R) - \frac{1}{8}\beta_2(H - 2R) + \frac{1}{4}\alpha_1 Q_1 \\ &+ X(\frac{1}{8}, -\frac{1}{8}, -\frac{1}{12}), \\ a_{1n} &= (-\frac{1}{4}\alpha_1^2 + \frac{1}{8}\beta_2 + \frac{1}{2}\beta_7)(3H - 2R) + \frac{1}{8}\beta_2(H - 2R) \\ &- \frac{1}{2}\alpha_1(Q_1 + Q_2) + X(-\frac{1}{4}, \frac{3}{8}, \frac{1}{4}), \\ a_{2n} &= (\frac{1}{8}\alpha_1^2 - \frac{1}{2}\beta_7)(3H - 2R) + \frac{1}{4}\alpha_1(Q_1 + 2Q_2) + X(\frac{1}{8}, -\frac{3}{8}, -\frac{1}{4}), \\ a_{3n} &= X(0, \frac{1}{8}, \frac{1}{12}). \end{aligned} \tag{5.2}$$

This means that the statistic T_n is invariant. □

Remark: Roughly speaking a converse of Theorem 5.1 is available. More precisely if the representation of the coefficients in the expansion (3.7) is of the form (5.2) it follows that the system of equations in (5.1) must be satisfied. To prove this inclusion, we note that in general Table 3.1 gives the same representation for the coefficient a_{0n} and for the remaining coefficients

$$\begin{aligned}
a_{1n} &= -\frac{1}{4}\alpha_1^2(3H-2R)+\frac{1}{2}\beta_2(H-R)-\frac{1}{2}\alpha_1(Q_1+Q_2)+X(-\frac{1}{4},\frac{3}{8},\frac{1}{4}) \\
&+\frac{1}{4}(\alpha_1+\alpha_2)P_6+\frac{1}{2}(\alpha_1+\alpha_2)P_8+\frac{1}{2}[(\alpha_1+\alpha_2)^2+(\beta_3+\beta_5)]P_9 \\
&+[\frac{1}{2}(\alpha_1+\alpha_2)^2-(\beta_3+\beta_5)-\frac{1}{2}(\beta_2+\beta_4+\beta_6)]P_{10} \\
&+[\frac{1}{2}(\alpha_1+\alpha_2)^2-(\beta_2+\beta_4+\beta_6)]P_{11}+\frac{1}{8}(\alpha_1+\alpha_2)^2P_{12} \\
&-\frac{1}{2}(\beta_7P_{12}+\beta_8P_{14})-(\beta_7P_{13}+\beta_8P_{15}) \\
&+[\frac{1}{4}(\alpha_1+\alpha_2)^2-\frac{1}{2}(\beta_2+\beta_4+\beta_6)]P_{13}-\frac{3}{2}(\beta_1+\beta_9)P_{16}, \\
a_{2n} &= \frac{1}{8}\alpha_1^2(3H-2R)+\frac{1}{4}\alpha_1(Q_1+2Q_2)+X(\frac{1}{8},-\frac{3}{8},-\frac{1}{4}) \\
&-\frac{1}{2}(\alpha_1+\alpha_2)(P_6+P_7)-(\alpha_1+\alpha_2)P_8 \\
&+[\frac{1}{2}(\beta_3+\beta_5)-(\alpha_1+\alpha_2)^2]P_9 \\
&+[-(\alpha_1+\alpha_2)^2+\frac{1}{2}(\beta_2+\beta_4+\beta_6)+(\beta_3+\beta_5)]P_{10} \\
&+[-(\alpha_1+\alpha_2)^2+\frac{1}{2}(\beta_2+\beta_4+\beta_6)]P_{11}+\frac{3}{2}(\beta_1+\beta_9)P_{16} \\
&-\frac{1}{4}(\alpha_1+\alpha_2)^2P_{12}+[-\frac{1}{2}(\alpha_1+\alpha_2)^2+\frac{1}{2}(\beta_2+\beta_4+\beta_6)]P_{13} \\
&+(\beta_7P_{13}+\beta_8P_{15})+\frac{1}{2}(\beta_7P_{12}+\beta_8P_{14}), \\
a_{3n} &= X(0,\frac{1}{8},\frac{1}{12})+\frac{1}{4}(\alpha_1+\alpha_2)P_6+\frac{1}{2}(\alpha_1+\alpha_2)(P_7+P_8) \\
&+\frac{1}{2}(\alpha_1+\alpha_2)^2(P_9+P_{10}+P_{11}+\frac{1}{4}P_{12}+\frac{1}{2}P_{13}).
\end{aligned}$$

Comparing these representations with (5.2) yields the system of equations in (5.1).

Definition 5.2: *A u_n-representable statistic is said to have the c-property if and only if the c-vector in the representation (2.7) satisfies*

$$\frac{c_1^2}{4}=c_2+c_4. \tag{5.3}$$

The class of u_n-representable statistics with c-property is quite rich [see Dette and Grigoriev (2000)] and a couple of examples will be presented in Section 6. The following result gives an alternative representation of the coefficients in the asymptotic expansion (3.7) for u_n-representable statistics with c-property. The proof is an immediate consequence of Definition 5.2 and the representation of the coefficients in Table 3.2.

Theorem 5.3: *Let $T_n \in \Psi_n$ denote a u_n-representable statistic with c-property, then the coefficients in the asymptotic representation (3.7) can be represented as follows*

$$
\begin{aligned}
(5.4) \qquad a_{0n} &= \frac{1}{8}(3H-2R) - \frac{3}{8}(H-2R) + \frac{1}{2}Q_1 + X(\frac{1}{8}, -\frac{1}{8}, -\frac{1}{12}) \ , \\
a_{1n} &= -\frac{1}{2}(c_2+\frac{1}{4})(3H-2R) + \frac{3}{8}(H-2R) - (Q_1+Q_2) + X(-\frac{1}{4}, \frac{3}{8}, \frac{1}{4}) \\
&+ \frac{1}{2}(c_1-1)\{\frac{1}{2}P_6 + P_8 + c_1P_9 + (c_1+1)P_{10} + (c_1-1)P_{11}\} - \frac{3}{2}(c_3-\frac{1}{3})P_{16}, \\
a_{2n} &= \frac{1}{2}c_2(3H-2R) + \frac{1}{2}(Q_1+2Q_2) + X(\frac{1}{8}, -\frac{3}{8}, -\frac{1}{4}) \\
&- \frac{1}{2}(c_1-1)\Big\{P_6 + P_7 + 2P_8 + (2c_1-1)P_9 + 2c_1P_{10} \\
&+ (c_1-1)(2P_{11} + \frac{1}{4}P_{12} + \frac{1}{2}P_{13}) + \frac{3}{2}(c_3-\frac{1}{3})P_{16}\Big\}, \\
a_{3n} &= X(0, \frac{1}{8}, \frac{1}{12}) \\
&+ \frac{1}{2}(c_1-1)\Big\{\frac{1}{2}P_6 + P_7 + P_8 + (c_1-1)(P_9 + P_{10} + P_{11} + \frac{1}{4}P_{12} + \frac{1}{2}P_{13})\Big\}.
\end{aligned}
$$

It is worthwhile to mention that for a u_n-representable statistic the coefficient a_{0n} is always invariant [see equation (5.4)]. In the next definition we introduce a further concept to classify quadratic statistics of the form $T_n = T_n(\theta, \hat{\theta}_n)$ [see for example the statistic of Kullback Leibler or Wald]. To this end we call two statistics $V_n, W_n \in \Psi_n$ equivalent if the coefficients $(\alpha_1, \alpha_2, \beta_1, \dots, \beta_9)$ in the asymptotic expansion (1.9) and (1.10) coincide and denote this property by $W_n \sim V_n$ as $n \to \infty$.

Definition 5.4: *Let $T_n = T_n(\theta, \hat{\theta}_n)$ denote a u_n-representable statistic, then the statistic $T_n^* = T_n(\hat{\theta}_n, \theta)$ is called conjugate statistic of T_n. If $T_n \sim T_n^*$, then the statistic T_n is called self-conjugate.*

Theorem 5.5: *Let T_n, T_n^* denote u_n-representable and conjugate statistics with structural coef-*

ficients $(c_i)_{i=1}^4$ *and* $(c_i^*)_{i=1}^4$*, respectively, then the following properties are satisfied*

(i)

$$c_1^* = 2 - c_1 \quad (5.5)$$
$$c_2^* = 1 - c_1 + c_2,$$
$$c_3^* = 1 - c_1 + c_3,$$
$$c_4^* = c_4,$$

(ii) T_n *is an invariant statistic if and only if*

$$\begin{aligned} c_1 &= 1, \\ c_2 + c_4 &= \frac{1}{4}, \\ c_3 &= \frac{1}{3}. \end{aligned} \quad (5.6)$$

(iii) T_n *is self-conjugate if and only if* $c_1 = 1$.

(iv) If T_n *is invariant, then* T_n *is self-conjugate and has the c-property.*

Proof: The proof of (i) is obtained from the expansion (2.7) observing that $u_n^* = -u_n$. The equation (5.6) is necessary and sufficient for the invariance of the statistic T_n by Theorem 5.1 and the equations (3.6) which express the structural coefficients $(\alpha_1, \alpha_2, \beta_1, \ldots, \beta_9)$ in terms of the vector $(c_i)_{i=1}^4$ of a u_n-representable statistic. Finally, part (iii) and (iv) are immediate consequences of Definition 5.2 and 5.4.

□

6 Applications

In this section we discuss a couple of quadratic statistics in the class Ψ_n, which are well known from the literature. We start with several statistics in the subset $\Psi_{1n} \subset \Psi_n$ of invariant statistics. To this end recall the definition of the projection P onto the tangent space $T^m(\theta)$ of the expectation surface at the point $\theta \in \Theta$ in (4.4) and let $\hat{P}$ denote the corresponding projection onto the tangent space $T^m(\hat{\theta}_n)$. We introduce the matrices

$$H_{ij} = M_{ij} - \langle (I - P)e, F_{ij} \rangle \quad (6.1)$$
$$\hat{Q}_{ij} = \hat{M}_{ij} - \langle (I - \hat{P})(\hat{\eta} - \eta), F_{ij} \rangle \quad (6.2)$$

and define five quadratic statistics by

$$\sigma^2 T_n^{(1)} = \|e\|^2 - \|\hat{e}\|^2, \tag{6.3}$$

$$\sigma^2 T_n^{(2)} = \Lambda^{ij} b_i b_j = \|Pe\|^2, \tag{6.4}$$

$$\sigma^2 T_n^{(3)} = H^{ij} b_i b_j, \tag{6.5}$$

$$\sigma^2 T_n^{(4)} = \|\hat{\eta} - \eta\|^2, \tag{6.6}$$

$$\sigma^2 T_n^{(5)} = \|P(\hat{\eta} - \eta)\|^2. \tag{6.7}$$

These statistics have been discussed by several authors in the case of a Gaussian error. The statistics $T_n^{(1)}$ and $T_n^{(2)}$ appear in the likelihood ratio test and in Rao's test [see Rao (1965), Section 6e.2] for testing the hypothesis $H_0 : \theta = \theta_0$, respectively. Similary, $T_n^{(4)}$ is proportional to the Kullback-Leibler distance between the Gaussian measures $P^n_{\hat{\theta}_n}$ and P^n_θ and the statistic $T_n^{(3)}$ has been proposed by Hamilton, Watts and Bates (1982) in the context of inference regions in nonlinear regression models. Finally, Grigoriev (1994) introduced the statistic $T_n^{(5)}$. The structural coefficients of the corresponding statistics (1.9) and (1.10) are presented in Table 6.1 and yield the second order asymptotic expansion (3.2). From the discussion in Section 2 and this table it follows that the asymptotic expansion used by Grigoriev and Ivanov (1993) is not applicable for the statistic of Hamilton, Watts and Bates ($T_n^{(3)}$) and Grigoriev ($T_n^{(5)}$) because $\beta_5 + \beta_6 \neq 4\beta_7$. Morover, a straightforward application of Theorem 5.1 yields the following result.

Corollary 6.1: *The statistics $T_n^{(1)} - T_n^{(5)}$ defined by (6.3) - (6.7), respectively, are invariant, i.e.*

$$T_n^{(r)} \in \Psi_{1n} \quad r = 1, 2, 3, 4, 5.$$

For a supplementary discussion of the invariance properties of these statistic we use a geometric point of view and present in Figure 6.1 the points

$$A = \eta(\theta),\ B = \hat{\eta} = \eta(\hat{\theta}_n),\ Y = y,$$
$$C = Py,\ D = P(\hat{\eta} - \eta).$$

Recalling the definitions (6.3) - (6.7) we obtain

$$\begin{aligned} T_n^{(1)} &= \|A - Y\|^2 - \|B - Y\|^2, \\ T_n^{(2)} &= \|A - C\|^2, \\ T_n^{(4)} &= \|A - B\|^2, \\ T_n^{(5)} &= \|A - D\|^2, \end{aligned}$$

and the discussion in Section 4 shows that the $T_n^{(1)}, T_n^{(2)}, T_n^{(4)}$ and $T_n^{(5)}$ are invariant with respect to regular reparametrizations, because they are functions of specific points of the expectation surface (and its tangent spaces). The invariance of the statistic $T_n^{(3)}$ cannot be obtained by such a simple geometric argument but follows from the representation

$$T_n^{(3)} = nH^{ij}\partial_i\partial_j \ ,$$

where ∂_i is defined in (4.5). Now (6.1) shows $H_{ij} = M_{ij} - \nabla_{ij}$ and M_{ij} and $\nabla_{ij} = \langle (I-P)e, F_{ij}\rangle$ are differential invariants. Therefore it follows that $T_n^{(3)}$ is invariant.

$T_n^{(r)}$	statistic	$n^{-1/2}$		n^{-1}								
r		α_1	α_2	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9
1	Neyman-Pearson	1	-1	$\frac{1}{3}$	1	-1	-2	1	1	$\frac{1}{4}$	$-\frac{1}{4}$	$-\frac{1}{3}$
2	Rao											
3	Hamilton-Watts-Bates	1	-1		1		-2		1			
4	Kullback-Leibler	2	-2	1	3	-3	-6	3	3	$\frac{3}{4}$	$-\frac{3}{4}$	-1
5	Grigoriev	2	-2	1	3	-3	-6	3	3	1	-1	-1
6	Wald, modif.	2	-3	1	3	-4	-8	5	5	$\frac{5}{4}$	-1	$-\frac{4}{3}$
7	Wald	2	-1	1	3	-2	-4	1	1	$\frac{1}{4}$		$-\frac{1}{3}$
8	Pazman, modif.	2	-3	1	3	-4	-8	5	5	$\frac{3}{4}$	$-\frac{1}{2}$	$-\frac{4}{3}$
9	Pazman	2	-1	1	3	-2	-4	1	1	$-\frac{1}{4}$	$\frac{1}{2}$	$\frac{2}{3}$

Table 6.1: Structural coefficients in the expansion (6.1) for various quadratic statistics proposed in the literature.

Corollary 6.2: *The statistics $T_n^{(4)}, T_n^{(5)}$ are u_n-representable, while $T_n^{(1)} - T_n^{(3)}$ are not u_n-representable.*

Proof: The proof follows by a careful comparison of the system (3.6) with the structural coefficients in the expansion (1.9) given in Table 6.1 and is left to the reader.

□

The second subset Ψ_{2n} of Ψ_n is the class of u_n -representable statistics. By the above Corollary $T_n^{(4)}$ and $T_n^{(5)}$ are u_n-representable and additionally we introduce the statistics

(6.8) $$\sigma^2 T_n^{(6)} = \Pi_{(i)(j)}(\theta) u_n^i u_n^j,$$

(6.9) $$\sigma^2 T_n^{(7)} = \Pi_{(i)(j)}(\hat{\theta}_n) u_n^i u_n^j,$$

$$\sigma^2 T_n^{(8)} = Q_{ij}(\theta) u_n^i u_n^j, \tag{6.10}$$

$$\sigma^2 T_n^{(9)} = Q_{ij}(\hat{\theta}_n) u_n^i u_n^j, \tag{6.11}$$

where the matrix Q_{ij} is defined in (6.2). Note that $T_n^{(7)}$ is the statistic introduced by Wald [see Rao (1965)] and that $T_n^{(9)}$ Pazman's statistic [see Pazman (1992)]. We will call $T_n^{(6)}$ and $T_n^{(8)}$ modified Wald and Pazman statistic, respectively. The structural coefficients of the corresponding statistics (1.9) and (1.10) are presented in Table 6.1 and yield the second order asymptotic expansion (3.2). From the discussion in Section 2 and Table 6.1 it follows that the asymptotic expansion used by Grigoriev and Ivanov (1993) is not applicable for the statistics of Pazman ($T_n^{(8)}$, $T_n^{(9)}$) because $\beta_5 + \beta_6 \neq 4\beta_7$. From Table 6.1 and the equations (3.6) it also follows that the statistics $T_n^{(6)} - T_n^{(9)}$ are u_n-representable and the corresponding c-vectors are given in Table 6.2.

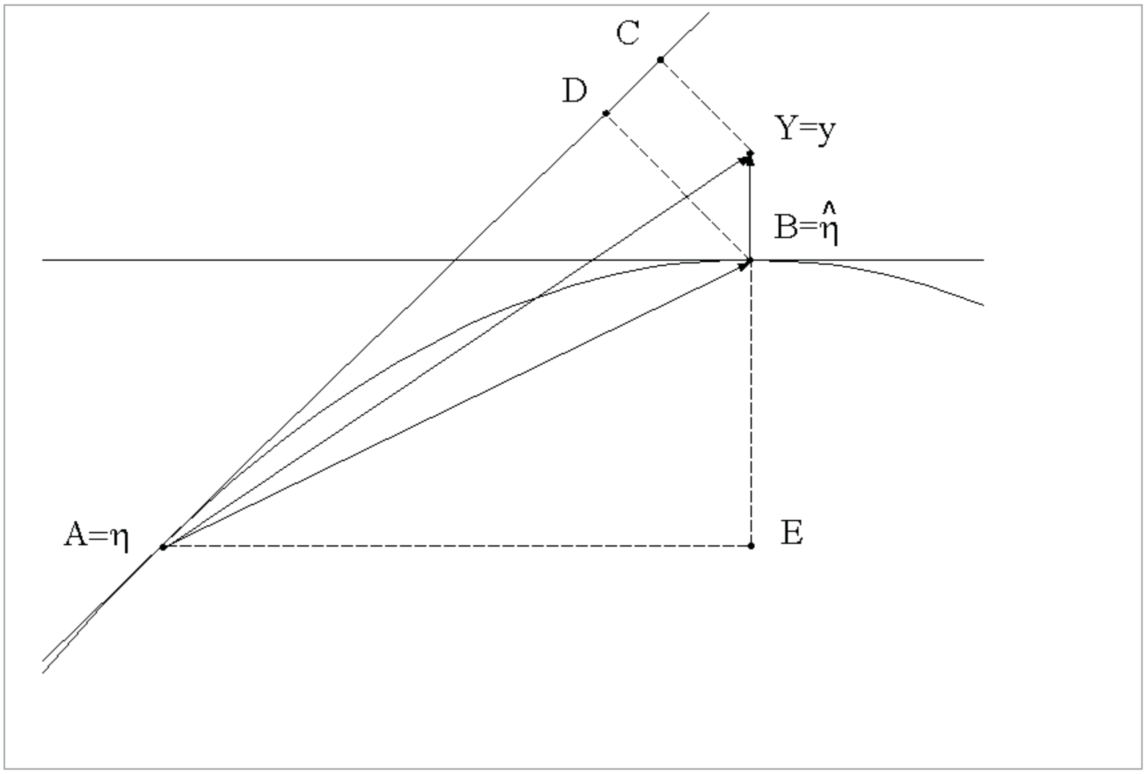

Figure 6.1. *The geometry of the statistics* $T_n^{(r)}$, $1, 2, 4, 5$

Statistic	$n^{-1/2}$	n^{-1}		
	c_1	c_2	c_3	c_4
$T_n^{(4)}$, Kullback-Leibler	1	$\frac{1}{4}$	$\frac{1}{3}$	0
$T_n^{(5)}$, Grigoriev	1	0	$\frac{1}{3}$	$\frac{1}{4}$
$T_n^{(6)}$, modified Wald	0	0	0	0
$T_n^{(7)}$, Wald	2	1	1	0
$T_n^{(8)}$, modified Pazman	0	$\frac{1}{2}$	0	$-\frac{1}{2}$
$T_n^{(9)}$, Pazman	2	$\frac{3}{2}$	1	$-\frac{1}{2}$

Table 6.2: c-vectors of u_n-representable statistics.

Obviously the statistics $T_n^{(4)}$ and $T_n^{(5)}$ are self-conjugate (which follows either by definition or by Theorem 5.5 (iii) and Table 6.2). Similary, $T_n^{(6)}$ and $T_n^{(7)}$ and $T_n^{(8)}$ and $T_n^{(9)}$ are conjugate statistics. Note that $T_n^{(4)}$ and $T_n^{(5)}$ are invariant and u_n-representable and therefore self-conjugate (see part (iv) of Theorem 5.5). Moreover,

$$T_n^{(4)} = T_n^{(4)*} \ ,$$

(see Figure 6.1) but

$$T_n^{(5)} \neq T_n^{(5)*} \ ,$$

because $T_n^{(5)} = \|A - D\|^2$ and $T_n^{(5)*} = \|A - E\|^2$.

We conclude this section presenting the Bartlett functions (see Section 3) for the statistics under consideration. For the invariant statistics $T_n^{(1)}, \ldots, T_n^{(5)}$ we obtain

$$\begin{aligned}
B_{1n} &= -\frac{1}{8}(\alpha_1^2 - \beta_2)B + \frac{1}{8}\beta_2 Y - \frac{1}{4}\alpha_1 Q_1 + X(-\frac{1}{8}, \frac{1}{8}, \frac{1}{12}), \\
B_{2n} &= \frac{1}{8}(\alpha_1^2 - 4\beta_7)B + \frac{1}{4}\alpha_1(Q_1 + 2Q_2) + X(\frac{1}{8}, -\frac{1}{4}, -\frac{1}{6}), \\
B_{3n} &= X(0, \frac{1}{8}, \frac{1}{12}),
\end{aligned}$$

where the differential invariants B, Y, Q_1, Q_2 and X have been defined in (4.16), (4.17), (4.11) , (4.12) and (4.10), respectively. The statistics $T_n^{(4)}$ and $T_n^{(5)}$ are u_n -representable with c-property and we obtain a further simplification

$$\begin{aligned}
B_{1n} &= -\frac{1}{2}R - \frac{1}{2}Q_1 + X(-\frac{1}{8}, \frac{1}{8}, \frac{1}{12}), \\
B_{2n} &= \frac{1}{2}c_2 B + \frac{1}{2}(Q_1 + 2Q_2) + X(\frac{1}{8}, -\frac{1}{4}, -\frac{1}{6}), \\
B_{3n} &= X(0, \frac{1}{8}, \frac{1}{12}),
\end{aligned}$$

where R has been defined in (4.14). Finally, the Bartlett functions for the non-invariant but u_n-representable (with c-property) statistics $T_n^{(6)} - T_n^{(9)}$ are given by

$$\begin{aligned}
B_{1n} &= -\frac{1}{2}R - \frac{1}{2}Q_1 + X(-\frac{1}{8}, \frac{1}{8}, \frac{1}{12}), \\
B_{2n} &= \frac{1}{2}c_2 B + \frac{1}{2}(Q_1 + 2Q_2) + X(\frac{1}{8}, -\frac{1}{4}, -\frac{1}{6}) + \frac{3}{2}(c_3 - \frac{1}{3})P_{16} \\
&- \frac{1}{2}(c_1 - 1)\Big\{\frac{1}{2}P_6 + P_8 + c_1 P_9 + (c_1+1)P_{10} + (c_1 - 1)P_{11}\Big\}, \\
B_{3n} &= X(0, \frac{1}{8}, \frac{1}{12}) + \frac{1}{2}(c_1 - 1)\Big\{\frac{1}{2}P_6 + P_7 + P_8 + (c_1 - 1)[P_9 + P_{10} + P_{11} + \frac{1}{4}P_{12} + \frac{1}{2}P_{13}]\Big\}.
\end{aligned}$$

Acknowledgements. This work was done during a visit of the second author at the Department of Mathematics, Ruhr-Universität Bochum, Germany. The authors would like to thank the Heinrich Hertz Stiftung for the financial support which made this visit possible and I. Gottschlich who typed this manuscript with considerable technical expertise. The work of H. Dette was supported by the Deutsche Forschungsgemeinschaft (SFB 475: Komplexitätsreduktion in multivariaten Datenstrukturen). The authors are also grateful to C. von Lieres and Wilkau for the preparation of Figure 6.1.

7 Appendix

The polynomials $M_{1n}(\cdot), M_{2n}(\cdot)$ in the asymptotic expansion (2.14) are given by:

$$\begin{aligned}
M_{1n}(\theta; y) &= \Big\{\frac{\gamma_3}{6\sigma^6}\Pi_{(\alpha)(\beta)(\gamma)} + \frac{1}{\sigma^2}(\pi_1 + 6\pi_2)\Pi_{(\alpha\beta)(\gamma)}\Big\}y^\alpha y^\beta y^\gamma \\
&- \Big\{(\pi_1 + 8\pi_2)\Pi_{(\alpha)(\beta\gamma)} + (\pi_1 + 4\pi_2))\Pi_{(\alpha\beta)(\gamma)} + \frac{\gamma_3}{2\sigma^4}\Pi_{(\alpha)(\beta)(\gamma)}\Big\}\Lambda^{\alpha\beta}y^\gamma, \\
M_{2n}(\theta; y) &= \sum_{i=1}^{9} m_i M_i + \sum_{i=10}^{37} m_i M_{i,\alpha\beta}y^\alpha y^\beta + \sum_{i=38}^{53} m_i M_{i,\alpha\beta\gamma\delta}y^\alpha y^\beta y^\gamma y^\delta \\
&+ \sum_{i=54}^{56} m_i M_{i,\alpha\beta\gamma\delta\varepsilon\nu}y^\alpha y^\beta y^\gamma y^\delta y^\varepsilon y^\nu.
\end{aligned}$$

Here the coefficients m_i and M_i are defined by:

$$\begin{aligned}
m_1 &= \frac{1}{8}, \quad M_1 = \sigma^{-4}\gamma_4\Lambda^{ir}\Lambda^{js}\Pi_{(i)(j)(r)(s)}, \\
m_2 &= -\frac{1}{8}, \quad M_2 = \sigma^{-6}\gamma_3^2\Lambda^{ij}\Lambda^{kl}\Lambda^{rs}\Pi_{(i)(j)(k)}\Pi_{(\ell)(r)(s)}, \\
m_3 &= -\frac{1}{12}, \quad M_3 = \sigma^{-6}\gamma_3^2\Lambda^{ij}\Lambda^{k\ell}\Lambda^{rs}\Pi_{(i)(k)(r)}\Pi_{(j)(\ell)(s)},
\end{aligned}$$

$$
\begin{aligned}
m_4 &= \frac{1}{2}\pi_1, \quad M_4 = \sigma^2\gamma_3\Lambda^{ir}\Lambda^{js}\Pi_{(ir)(j)(s)},\\
m_5 &= -\frac{1}{2}\pi_1, \quad M_5 = \sigma^{-2}\gamma_3\Lambda^{ir}\Lambda^{js}\Lambda^{k\ell}\Pi_{(ir)(j)}\Pi_{(k)(\ell)(s)},\\
m_6 &= \frac{1}{2}\pi_1^2, \quad M_6 = \sigma^2\Lambda_n^{ir}\Lambda_n^{js}\Pi_{(ir)(js)},\\
m_7 &= -\frac{1}{2}\pi_1^2, \quad M_7 = \sigma^2\Lambda^{ir}\Lambda^{js}\Lambda^{k\ell}\Pi_{(ir)(k)}\Pi_{(js)(\ell)},\\
m_8 &= \frac{1}{2}\pi_1^2-\rho_2, \quad M_8 = \sigma^2\Lambda^{ir}\Lambda^{js}\Pi_{(jr)(is)},\\
m_9 &= -\Big(\frac{1}{2}\pi_1^2-\rho_2\Big), \quad M_9 = \sigma^2\Lambda^{ir}\Lambda^{js}\Lambda^{k\ell}\Pi_{(jr)(k)}\Pi_{(is)(\ell)},\\
m_{10} &= -\frac{1}{4}, \quad M_{10,\alpha\beta} = \sigma^{-6}\gamma_4\Lambda^{ij}\Pi_{(i)(j)(\alpha)(\beta)},\\
m_{11} &= \frac{1}{4}, \quad M_{11,\alpha\beta} = \sigma^{-8}\gamma_3^2\Lambda^{ij}\Lambda^{k\ell}\Pi_{(i)(j)(k)}\Pi_{(\ell)(\alpha)(\beta)},\\
m_{12} &= \frac{1}{4}, \quad M_{12,\alpha\beta} = \sigma^{-8}\gamma_3^2\Lambda^{ij}\Lambda^{k\ell}\Pi_{(i)(k)(\alpha)}\Pi_{(j)(\ell)(\beta)},\\
m_{13} &= \frac{1}{8}, \quad M_{13,\alpha\beta} = \sigma^{-8}\gamma_3^2\Lambda^{ij}\Lambda^{k\ell}\Pi_{(i)(j)(\alpha)}\Pi_{(k)(\ell)(\beta)},\\
m_{14} &= \frac{1}{2}(\pi_1+8\pi_2), \quad M_{14,\alpha\beta} = \sigma^{-4}\gamma_3\Lambda^{ir}\Lambda^{js}\Pi_{(j\alpha)(s)}\Pi_{(i)(r)(\beta)},\\
m_{15} &= \frac{1}{2}(\pi_1+2\pi_2), \quad M_{15,\alpha\beta} = \sigma^{-4}\gamma_3\Lambda^{ir}\Lambda^{js}\Pi_{(\alpha\beta)(j)}\Pi_{(r)(s)(i)},\\
m_{16} &= \pi_1, \quad M_{16,\alpha\beta} = \sigma^{-4}\gamma_3\Lambda^{ir}\Lambda^{js}\Pi_{(\alpha j)(i)}\Pi_{(r)(s)(\beta)},\\
m_{17} &= \frac{1}{2}\pi_1, \quad M_{17,\alpha\beta} = \sigma^{-4}\gamma_3\Lambda^{ir}\Lambda^{js}\Pi_{(js)(i)}\Pi_{(r)(\alpha)(\beta)},\\
m_{18} &= -\pi_1, \quad M_{18,\alpha\beta} = \sigma^{-4}\gamma_3\Lambda^{ij}\Pi_{(ia)(j)(\beta)},\\
m_{19} &= -\frac{1}{2}\pi_1, \quad M_{19,\alpha\beta} = \sigma^{-4}\gamma_3\Lambda^{ij}\Pi_{(\alpha\beta)(i)(j)},\\
m_{20} &= -\frac{1}{2}\pi_1, \quad M_{20,\alpha\beta} = \sigma^{-4}\gamma_3\Lambda^{ij}\Pi_{(ij)(\alpha)(\beta)},\\
m_{21} &= \frac{1}{2}(\pi_1+4\pi_2), \quad M_{21,\alpha\beta} = \sigma^{-4}\gamma_3\Lambda^{ir}\Lambda^{js}\Pi_{(js)(\alpha)}\Pi_{(i)(r)(\beta)},\\
m_{22} &= \frac{1}{2}(\pi_1+4\pi_2), \quad M_{22,\alpha\beta} = \sigma^{-4}\gamma_3\Lambda^{ir}\Lambda^{js}\Pi_{(\alpha)(\beta j)}\Pi_{(s)(i)(r)},\\
m_{23} &= \frac{1}{2}(\pi_1+8\pi_2)^2, \quad M_{23,\alpha\beta} = \Lambda^{ir}\Lambda^{js}\Pi_{(r)(i\alpha)}\Pi_{(s)(j\beta)},\\
m_{24} &= (\pi_1+4\pi_2)(\pi_1+8\pi_2), \quad M_{24,\alpha\beta} = \Lambda^{ir}\Lambda^{js}\Pi_{(s)(j\alpha)}\Pi_{(ir)(\beta)},\\
m_{25} &= \frac{1}{2}(\pi_1+4\pi_2)^2, \quad M_{25,\alpha\beta} = \Lambda^{ir}\Lambda^{js}\Pi_{(ir)(\alpha)}\Pi_{(js)(\beta)},\\
m_{26} &= 2\pi_1\pi_2+16\pi_2^2-(\rho_3+\rho_8+\rho_{10}), \quad M_{26,\alpha\beta} = \Lambda^{ir}\Lambda^{js}\Pi_{(r)(ij)}\Pi_{(s)(\alpha\beta)},\\
m_{27} &= (\pi_1+4\pi_2)(\pi_1+8\pi_2)-(\rho_2+2\rho_3+\rho_4+\rho_5+\rho_6+\rho_7+\rho_9+2\rho_{10}+\rho_{11}+\rho_{12})\\
M_{27,\alpha\beta} &= \Lambda^{ir}\Lambda^{js}\Pi_{(s)(i\alpha)}\Pi_{(rj)(\beta)}
\end{aligned}
$$

$$
\begin{aligned}
m_{28} &= (\pi_1+4\pi_2)(\pi_1+8\pi_2) - (\rho_2+\rho_5+\rho_6+\rho_7+\rho_{11}+\rho_{12}) \\
M_{28,\alpha\beta} &= \Lambda^{ir}\Lambda^{js}\Pi_{(s)(ij)}\Pi_{(r\alpha)(\beta)} \\
m_{29} &= (\pi_1+4\pi_2)^2 - (\rho_4+\rho_9), \quad M_{29,\alpha\beta} = \Lambda^{ir}\Lambda^{js}\Pi_{(ir)(j)}\Pi_{(s\alpha)(\beta)}, \\
m_{30} &= \frac{1}{2}(\pi_1+4\pi_2)^2 - (\rho_2+\rho_6+\rho_7+\rho_{12}), \quad M_{30,\alpha\beta} = \Lambda^{ir}\Lambda^{js}\Pi_{(rj)(\alpha)}\Pi_{(si)(\beta)}, \\
m_{31} &= -\frac{3}{2}\pi_1^2 + \rho_2 - 2\rho_{14}, \quad M_{31,\alpha\beta} = \Lambda^{ir}\Pi_{(r\alpha)(i\beta)}, \\
m_{32} &= -(\pi_1^2+\rho_{14}), \quad M_{32,\alpha\beta} = \Lambda^{ir}\Pi_{(ir)(\alpha\beta)}, \\
m_{33} &= \frac{1}{2}(3\pi_1^2+8\pi_1\pi_2+32\pi_2^2) - (\rho_2+\rho_5+\rho_{11}+2\rho_{13}), \quad M_{33,\alpha\beta} = \Lambda^{ir}\Lambda^{js}\Pi_{(s)(i\alpha)}\Pi_{(r\beta)(j)}, \\
m_{34} &= (\pi_1^2+2\pi_1\pi_2+8\pi_2^2) - \rho_{13}, \quad M_{34,\alpha\beta} = \Lambda^{ir}\Lambda^{js}\Pi_{(ir)(j)}\Pi_{(s)(\alpha\beta)}, \\
m_{35} &= \frac{1}{2}(\pi_1^2+8\pi_1\pi_2+32\pi_2^2) - (\rho_4+2\rho_8+\rho_9), \quad M_{35,\alpha\beta} = \Lambda^{ir}\Lambda^{js}\Pi_{(r)(j\alpha)}\Pi_{(s)(i\beta)}, \\
m_{36} &= -(\rho_1+3\rho_{15}+\rho_{16}), \quad M_{36,\alpha\beta} = \Lambda^{ir}\Pi_{(i)(r\alpha\beta)}, \\
m_{37} &= -2(\rho_1+\rho_{16}), \quad M_{37,\alpha\beta} = \Lambda^{ir}\Pi_{(\alpha)(ir\beta)}, \\
m_{38} &= \frac{1}{24}, \quad M_{38,\alpha\beta\gamma\delta} = \sigma^{-8}\gamma_4\Pi_{(\alpha)(\beta)(\gamma)(\delta)}, \\
m_{39} &= -\frac{1}{12}, M_{39,\alpha\beta\gamma\delta} = \sigma^{-10}\gamma_3^2\Lambda^{ij}\Pi_{(i)(j)(\alpha)}\Pi_{(\beta)(\gamma)(\delta)} \\
m_{40} &= -\frac{1}{8}, \quad M_{40,\alpha\beta\gamma\delta} = \sigma^{-10}\gamma_3^2\Lambda^{ij}\Pi_{(i)(\alpha)(\beta)}\Pi_{(j)(\gamma)(\delta)} \\
m_{41} &= \frac{1}{2}\pi_1, \quad M_{41,\alpha\beta\gamma\delta} = \sigma^{-6}\gamma_3\Pi_{(\alpha)(\beta)(\gamma\delta)} \\
m_{42} &= -\frac{1}{6}(\pi_1+8\pi_2), \quad M_{42,\alpha\beta\gamma\delta} = \sigma^{-6}\gamma_3\Lambda^{ij}\Pi_{(i\alpha)(j)}\Pi_{(\beta)(\gamma)(\delta)} \\
m_{43} &= -\frac{1}{2}(\pi_1+6\pi_2), \quad M_{43,\alpha\beta\gamma\delta} = \sigma^{-6}\gamma_3\Lambda^{ij}\Pi_{(\alpha\beta)(\gamma)}\Pi_{(i)(j)(\delta)} \\
m_{44} &= -\frac{1}{2}(\pi_1+2\pi_2), \quad M_{44,\alpha\beta\gamma\delta} = \sigma^{-6}\gamma_3\Lambda^{ij}\Pi_{(\alpha\beta)(i)}\Pi_{(j)(\gamma)(\delta)} \\
m_{45} &= -\frac{1}{6}(\pi_1+4\pi_2), \quad M_{45,\alpha\beta\gamma\delta} = \sigma^{-6}\gamma_3\Lambda^{ij}\Pi_{(ij)(\alpha)}\Pi_{(\beta)(\gamma)(\delta)} \\
m_{46} &= -\frac{1}{2}(\pi_1+4\pi_2), \quad M_{46,\alpha\beta\gamma\delta} = \sigma^{-6}\gamma_3\Lambda^{ij}\Pi_{(\alpha)(\beta i)}\Pi_{(\gamma)(\delta)(j)} \\
m_{47} &= -(\pi_1^2+12\pi_1\pi_2+40\pi_2^2), \quad M_{47,\alpha\beta\gamma\delta} = \sigma^{-2}\Lambda^{rs}\Pi_{(r)(s\alpha)}\Pi_{(\beta)(\gamma\delta)} \\
m_{48} &= -(\pi_1+4\pi_2)(\pi_1+8\pi_2), \quad M_{48,\alpha\beta\gamma\delta} = \sigma^{-2}\Lambda^{rs}\Pi_{(\alpha)(rs)}\Pi_{(\beta\gamma)(\delta)} \\
m_{49} &= -(\pi_1^2+10\pi_1\pi_2+32\pi_2^2) + (\rho_3+\rho_4+\rho_8+\rho_9+\rho_{10}), \quad M_{49,\alpha\beta\gamma\delta} = \sigma^{-2}\Lambda^{rs}\Pi_{(r)(\alpha\beta)}\Pi_{(s\gamma)(\delta)} \\
m_{50} &= -\frac{1}{2}(\pi_1+4\pi_2)(3\pi_1+28\pi_2) + (\rho_2+\rho_5+\rho_6+\rho_7+\rho_{11}+\rho_{12}) \\
M_{50,\alpha\beta\gamma\delta} &= \sigma^{-2}\Lambda^{rs}\Pi_{(r\alpha)(\beta)}\Pi_{(sj)(\delta)} \\
m_{51} &= \frac{1}{2}\pi_1^2 + \rho_{14}, \quad M_{51,\alpha\beta\gamma\delta} = \sigma^{-2}\Pi_{(\alpha\beta)(\gamma\delta)} \\
m_{52} &= -\frac{1}{2}(\pi_1^2+4\pi_2^2) + \rho_{13}, \quad M_{52,\alpha\beta\gamma\delta} = \sigma^{-2}\Lambda^{rs}\Pi_{(r)(\alpha\beta)}\Pi_{(s)(\gamma\delta)}
\end{aligned}
$$

$$
\begin{aligned}
m_{53} &= \rho_1 + \rho_{15} + \rho_{16}, \quad M_{53,\alpha\beta\gamma\delta} = \sigma^{-2}\Pi_{(\alpha)(\beta\gamma\delta)} \\
m_{54} &= \frac{1}{72}, \quad M_{54,\alpha\beta\gamma\delta\varepsilon\nu} = \sigma^{-12}\gamma_3^2\Pi_{(\alpha)(\beta)(\gamma)}\Pi_{(\delta)(\varepsilon)(\nu)} \\
m_{55} &= \frac{1}{6}(\pi_1 + 6\pi_2), \quad M_{55,\alpha\beta\gamma\delta\varepsilon\nu} = \sigma^{-8}\gamma_3\Pi_{(\alpha)(\beta\gamma)}\Pi_{(\delta)(\varepsilon)(\nu)} \\
m_{56} &= \frac{1}{2}(\pi_1 + 6\pi_2)^2, \quad M_{56,\alpha\beta\gamma\delta\varepsilon\nu} = \sigma^{-4}\Pi_{(\alpha)(\beta\gamma)}\Pi_{(\delta)(\varepsilon\nu)}.
\end{aligned}
$$

References

Abramowitz, M., Stegun, I.A. (1965). Handbook of Mathematical Functions, Dover, N.Y.

Amari, S.I. (1985). Differential Geometric Methods in Statistics. Springer, N.Y.

Bardadym, T.A., Ivanov, A.V. (1985). An asymptotic expansion related to an empirical regression function. Theor. Math. Stat., 30, 7-13.

Bartlett, M.S. (1954). A note on multiplying factors in various χ^2-approximations. J. Roy. Statist. Soc. B, 16, 296-298.

Bates, D. , Watts, W. (1981). Parameter transformation for improved approximate confidence regions in nonlinear least squares. Ann. Statist. 9, 1151-1167.

Bates, D. , Watts, W. (1988). Nonlinear Regression Analysis and its Applications. Wiley, N.Y.

Borovkov, A.A. (1998). Mathematical Statistics. Gordon and Breach, London.

Beale, E.M.L. (1960). Confidence regions in non-linearity estimation. (with Discussion), J. Roy. Statist. Soc. B, 22, 41-88.

Chandra, T.K. , Ghosh, J.K. (1979). Valid asymptotic expansions for the likelihood ration statistic and other perturbed chi-square variables. Sankhya, Ser. A, 41, 22-47.

Chibisov, D.M. (1973). An asymptotic expansion for the distribution of sums of a special form with an application to minimum contrast estimates. Theory Prob. Appl., 18, 649-661.

Dette, H., Grigoriev, Yu.D. (2000). A unified approach to second order optimality criteria in nonlinear design of experiments. Ann. Inst. Stat. Math. 52, 574-597.

Efron, B. (1975). Defining the curvature of a statistical problem with applications to second order efficiency (with discussion). Ann. Statist., 3, 1199-1242.

Grigoriev, Yu.D. (1994). Development and Investigation of Algorithms in the Analysis of Nonlinear Regression Models. Doctor of Science Thesis. State Technical University, Novosibirsk (in Russian).

Grigoriev, Yu.D. (1995). Asymptotic expansions in the theory of nonlinear regression. Izvestiya vischikh uchebnikh zavedenij (Physik), No. 9, 20-25 (in Russian).

Grigoriev, Yu. D., Ivanov, A.V. (1987a). Asymptotic expansions in nonlinear regression analysis. Zavod. Labor., 53, 48-51. (in Russian)

Grigoriev, Yu. D. , Ivanov A.V. (1992). Asymptotic expansion of the distribution of some functional of least squares estimator. Dokl. Akad. Nauk Ukrain., 7, 26-30 (in Russian).

Grigoriev, Yu.D. , Ivanov, A.V. (1993). Asymptotic expansions for quadratic functionals of the least squares estimator of a nonlinear regression parameter. Math. Math. Statist., 2, 269-294.

Hamilton, D., Watts, D., Bates, D. (1982). Accounting for intrinsic nonlinearity in nonlinear regression parameter inference regions. Ann. Statist., 10, 386-393.

Ivanov, A.V. (1997). Asymptotic Theory of Nonlinear Regression. Kluwer Academic Publishers.

Ivanov, A.V., Zwanzig, S. (1983). An asymptotic expansion of the distribution of least squares estimators in the nonlinear regression model. Math. Operationsforsch. und Statist., Ser. Statist., 14, 7-27.

Linnik, J.V., Mitrofanova, N.M. (1965). Some asymptotic expansions for the distribution of the maximum likelihood estimate. Sankhya, 27, 73-82.

McCullagh, P. and Cox, D.R. (1986). Invariants and likelihood ratio Statistics. Ann. Statist., 14, 1419-1430.

Michel, R. (1975). An asymptotic expansion for the distribution of asymptotic maximum likelihood estimators of vector parameters. J. Multivar. Anal., 5, 67-82.

Pazman, A. (1990). Small-sample distribution of properties of nonlinear regression estimators (a geometric approach). Statist., 21, 323-367.

Pazman, A. (1992). A classification of nonlinear regression models and parameter confidence regions. Kybernetika, 28, no. 6, 444-453.

Pazman, A. (1993). Nonlinear Statistical Models. Kluwer Academic Publishers.

Pfanzagl, J. (1973). Asymptotically optimum estimation and test procedures. Proc. Prague Symp. and asymptotic Statistics, Vol. 1, Charles University Prague, 201-272.

Rao, C.R. (1965). Linear Statistical Inference and its Applications. Wiley.

Rashevsky, P.C. (1967). Riemannian Geometry and Tensor Analysis, Nauk, Moscow (in Russian).

Veblen, O. (1933). Invariants of Quadratic Differential Forms. Cambridge, University Press.