Ann. Inst. Statist. Math.
Vol. 46, No. 4, 649–666 (1994)

THE GEOMETRIC STRUCTURE OF THE EXPECTED/OBSERVED LIKELIHOOD EXPANSIONS*

LUIGI PACE** AND ALESSANDRA SALVAN

Department of Statistics, University of Padova, Via S.Francesco 33, I-35121 Padova, Italy

(Received June 21, 1993; revised March 22, 1994)

Abstract. Stochastic expansions of likelihood quantities are a basic tool for asymptotic inference. The traditional derivation is through ordinary Taylor expansions, rearranging terms according to their asymptotic order. The resulting expansions are called here *expected/observed*, being expressed in terms of the score vector, the expected information matrix, log likelihood derivatives and their joint moments. Though very convenient for many statistical purposes, expected/observed expansions are not usually written in tensorial form. Recently, within a differential geometric approach to asymptotic statistical calculations, invariant Taylor expansions based on likelihood yokes have been introduced. The resulting formulae are invariant, but the quantities involved are in some respects less convenient for statistical purposes. The aim of this paper is to show that, through an invariant Taylor expansion of the coordinates related to the expected likelihood yoke, expected/observed expansions up to the fourth asymptotic order may be re-obtained from invariant Taylor expansions. This derivation produces *invariant expected/observed* expansions.

Key words and phrases: Asymptotic expansions, index notation, invariant Taylor series expansions, likelihood, tensors, yokes.

1. Introduction

Let $\mathcal{F} = \{P_\omega : \omega \in \Omega \subseteq \mathbb{R}^d\}$ be a parametric family of probability distributions defined on a sample space $\mathcal{X}$ and dominated by a σ-finite measure μ. The parameter space Ω is assumed to be an open non-empty subset of $\mathbb{R}^d$. Let us denote by $p(x;\omega)$, $x \in \mathcal{X}$, the density of P_ω with respect to μ and by $l(\omega) = \log p(x;\omega)$ the log likelihood function based on the sample data x. We assume that the log likelihood function is a smooth function of the parameter and that the usual additional regularity conditions hold ensuring, in particular, that the maximum likelihood

* This research was partially supported by the Italian National Research Council grant n.93.00824.CT10.

** Now at Department of Economics, University of Udine, Via Tomadini 30/A, I-33100 Udine, Italy.

estimator (MLE) $\hat{\omega}$ exists and that it is uniquely defined as the solution of the likelihood equation $\nabla l(\omega) = 0$.

Let the statistic $f(\hat{\omega})$ be a parametrization invariant function (a scalar function in the language of differential geometry) defined on a copy of Ω that represents the range space of the MLE. The maximized log likelihood is the most notable instance. Stochastic expansions for $f(\hat{\omega})$ are basic for higher-order asymptotic inference. However, the ordinary Taylor formula for $f(\hat{\omega})$ is not parametrization invariant, depending on the coordinate system adopted for the statistical manifold $\mathcal{F}$. Let us assume that $f(\hat{\omega})$ is of order $O_P(n^{\beta})$ under repeated sampling of size n. An important aim of a 'geometric' stochastic calculus is to obtain an asymptotic expansion for $f(\hat{\omega})$ typically of the form

$$f(\hat{\omega}) = f(\omega) + b_1 + b_2 + b_3 + b_4 + O_P(n^{\beta-5/2}), \tag{1.1}$$

where each term b_m is a scalar function of order $O_P(n^{\beta-m/2})$, $m = 1, 2, 3, 4$; if $f(\hat{\omega})$ is a more general geometric object, for instance a tensor, it is desirable that each term b_m should follow the same transformation law as $f(\hat{\omega})$. A first possibility for obtaining the geometric expansion (1.1) is through the ordinary Taylor formula and *a posteriori* elicitation of geometric ingredients; a second possibility is to use an invariant Taylor series expansion (Barndorff-Nielsen (1987, 1989), see also Blæsild (1990)), namely a Taylor expansion defined on the manifold $\mathcal{F}$ (Murray (1988), Murray and Rice ((1993), Chapter 9)).

Invariant Taylor series expansions are intrinsically geometric and rely upon the definition of an appropriate *yoke* (Barndorff-Nielsen (1987)). Two instances of a yoke are relevant for asymptotic statistical calculations: the *observed likelihood yoke* and the *expected likelihood yoke*, which are both defined in terms of the log likelihood function. The former yoke gives rise to *observed* likelihood expansions, where the coefficients depend on mixed derivatives of the log likelihood function and the coordinates are products of elements of the score vector. The latter yoke gives rise to *expected* likelihood expansions. Their coefficients consist of moments of log likelihood derivatives together with derivatives of $f(\omega)$; the coordinates are products of expected values of elements of the score vector, where the expectation is taken under $\hat{\omega}$. Examples of expansions of both kinds are given in Barndorff-Nielsen *et al.* (1991*a*). A possible difficulty connected with observed expansions is that the calculation of their coefficients requires the specification of an auxiliary statistic. On the other hand, in expected expansions, coordinates are defined through an integration and the resulting expressions are not easy to handle nor to interprete; moreover, the summands in each invariant term b_m are not in general of the same asymptotic order in n.

For statistical applications, the most useful structure of a stochastic expansion would have coefficients analogous to those of the expected expansions and coordinates like those of the observed expansions. These requirements are met by the expansions considered in Lawley (1956), McCullagh and Cox (1986), McCullagh ((1987), Chapter 7), Barndorff-Nielsen *et al.* ((1991*a*), Section 8). Following the last reference, here this kind of expansions is referred to as *expected/observed* likelihood expansions. In this case, the built-in invariance of geometrically obtained expansions is lost and further manipulation is required for a geometric structure to become visible. In McCullagh and Cox (1986) a technique is suggested for recovering invariant terms in the expansion of log likelihood ratio statistics. Their contribution inspired much of subsequent work on invariant Taylor series in statistics.

The purpose of this paper is to produce a geometric formulation for the expected/observed expansions, relating them to the expected likelihood expansions. In particular, it is shown that they may be alternatively derived as expected likelihood expansions with coordinates substituted by a suitable geometric expansion. Our derivation is bound to the first four asymptotic orders (which is, however, all that one is likely to need in practice).

Section 2 contains some notation and background material. In Section 3 expected/observed likelihood expansions are briefly reviewed following Barndorff-Nielsen *et al.* (1991*a*), while Section 4 deals with expected likelihood expansions. Section 5 introduces the expansion for the coordinates of the expected likelihood expansions that permits us to obtain the required geometric formulation of the expected/observed expansions.

2. Basic notation and preliminaries

Throughout the paper we use index notation and the Einstein summation convention (whenever an index appears twice or more in a product of symbols summation over that index is understood). Let us denote by $\omega^r, \omega^s, \dots$ $(r, s = 1, \dots, d)$ generic components of ω. For a smooth real function $f(\omega)$, defined on $\mathcal{F}$ and expressed in coordinate form, we write

$$f_r = f_r(\omega) = \frac{\partial}{\partial \omega^r} f(\omega), \qquad f_{rs} = f_{rs}(\omega) = \frac{\partial^2}{\partial \omega^r \partial \omega^s} f(\omega),$$

and so on. More generally, denoting by R_m the set of indices $r_1 \cdots r_m$ we write

$$f_{R_m} = f_{R_m}(\omega) = \frac{\partial^m}{\partial \omega^{r_1} \cdots \partial \omega^{r_m}} f(\omega).$$

With reference to the log likelihood function $l(\omega)$, l_r is a component of the score vector and $i_{rs} = E_\omega(-l_{rs})$ indicates an element of the expected information matrix $[i_{rs}]$. We denote by i^{rs} an element of the matrix inverse of $[i_{rs}]$ and define $l^r = i^{rs} l_s$, $l^{rs} = l^r l^s$, $l^{rst} = l^r l^s l^t$ and so on. For moments of log likelihood derivatives we use the symbols $\mu_r = E_\omega(l_r), \mu_{rs} = E_\omega(l_{rs}), \dots, \mu_{R_m} = E_\omega(l_{R_m})$ and $\mu_{r,s} = E_\omega(l_r l_s), \mu_{r,st} = E_\omega(l_r l_{st}), \dots, \mu_{R_m, S_n, \dots, U_q} = E_\omega(l_{R_m} l_{S_n} \cdots l_{U_q})$. Observe that $\mu_{rs} = -i_{rs}$. Differentiating the identity $\mu_r = 0$ one obtains a sequence of balance relations, known as the Bartlett relations, the first instances of which are $\mu_{rs} + \mu_{r,s} = 0$, $\mu_{rst} + \mu_{r,st}[3] + \mu_{r,s,t} = 0$. The symbol $[k]$ indicates a sum over k similar terms obtained by suitable permutation of indices.

Let ψ be an alternative parametrization of $\mathcal{F}$, namely a one-to-one smooth function of ω with smooth inverse. We will denote the generic components of ψ by $\psi^a, \psi^b, \dots$, in contrast with the use of letters $r, s, \dots$ to denote components of ω. Accordingly, a likelihood quantity denoted by indices $a, b, \dots$ is understood as referred to the ψ parametrization. A re-parametrization does not alter the log likelihood function itself while it affects log likelihood derivatives and their moments. In particular, it is easy to see that

$$
\begin{aligned}
l_a &= l_r \omega^r_a, \\
l_{ab} &= l_{rs}\omega^r_a\omega^s_b + l_r\omega^r_{ab}, \\
l_{abc} &= l_{rst}\omega^r_a\omega^s_b\omega^t_c + l_{rs}\omega^r_{ab}\omega^s_c[3] + l_r\omega^r_{abc},
\end{aligned}
\tag{2.1}
$$

where $\omega^r_a = \partial\omega^r/\partial\psi^a$, $\omega^r_{ab} = \partial^2\omega^r/\partial\psi^a\partial\psi^b$ and so on. Evaluating joint moments up to the fourth order transformation rules of one of the following forms are obtained:

$$
\begin{aligned}
i_{ab} &= \mu_{a,b} = \mu_{r,s}\omega^r_a\omega^s_b = i_{rs}\omega^r_a\omega^s_b, \\
\mu_{a,b,c} &= \mu_{r,s,t}\omega^r_a\omega^s_b\omega^t_c,
\end{aligned}
\tag{2.2}
$$

and

$$
\begin{aligned}
\mu_{abc} &= \mu_{rst}\omega^r_a\omega^s_b\omega^t_c + \mu_{rs}\omega^r_{ab}\omega^s_c[3], \\
\mu_{a,bc} &= (\mu_{r,st}\omega^s_b\omega^t_c + \mu_{r,s}\omega^s_{bc})\omega^r_a, \\
\mu_{a,b,cd} &= (\mu_{r,s,tu}\omega^t_c\omega^u_d + \mu_{r,s,t}\omega^t_{cd})\omega^r_a\omega^s_b,
\end{aligned}
\tag{2.3}
$$

and, finally,

$$
\begin{aligned}
\mu_{ab,cd} = {} & \mu_{rs,tu}\omega^r_a\omega^s_b\omega^t_c\omega^u_d + \mu_{rs,t}\omega^r_a\omega^s_b\omega^t_{cd} \\
& + \mu_{r,st}\omega^r_{ab}\omega^s_c\omega^t_d + \mu_{r,s}\omega^r_{ab}\omega^s_{cd}.
\end{aligned}
\tag{2.4}
$$

Moreover we have

$$
\begin{aligned}
i^{ab} &= i^{rs}\psi^a_r\psi^b_s, \\
l^a &= l^r\psi^a_r,
\end{aligned}
\tag{2.5}
$$

where $\psi^a_r = \partial\psi^a/\partial\omega^r$.

Under re-parametrization, log likelihood derivatives and their moments transform according to fairly regular patterns. The simplest one is that displayed in (2.2), (2.5) and that obeyed by l_r, where only multilinear combinations of the quantity itself together with the Jacobian matrices $[\omega^r_a]$, $[\psi^a_r]$ are involved. Quantities that follow a transformation rule of this kind are called *tensors*. To be specific, a collection of smooth real functions $T^{R_m}_{S_n} = T^{R_m}_{S_n}(\omega) = T^{r_1\cdots r_m}_{s_1\cdots s_n}(\omega)$ is called a (m, n) tensor on $\mathcal{F}$ if under re-parametrization it obeys the transformation rule

$$
T^{A_m}_{B_n} = T^{R_m}_{S_n}\psi^{a_1}_{r_1}\cdots\psi^{a_m}_{r_m}\omega^{s_1}_{b_1}\cdots\omega^{s_n}_{b_n}.
\tag{2.6}
$$

According to this definition, l_r is a $(0, 1)$ tensor, i_{rs} is a $(0, 2)$ tensor, i^{rs} is a $(2, 0)$ tensor, l^r is a $(1, 0)$ tensor. A $(0, 0)$ tensor is a scalar. Since $\omega^r_a\psi^a_s = \delta^r_s$, where the symbol δ^r_s denotes the Kronecker delta ($\delta^r_s = 1$ if $r = s$, $\delta^r_s = 0$ if $r \neq s$), it follows immediately from (2.6) that scalars formed by a suitable product of tensors (contraction) are invariants. For instance, $i^{ab}l_a l_b = i^{rs}l_r l_s$; in general, if T^{R_m} is a $(m, 0)$ tensor and U_{S_m} is a $(0, m)$ tensor, their contraction $T^{R_m}U_{R_m}$ is invariant. Tensors are therefore particularly appropriate tools in writing geometric expansions: if we are able to write an expansion in terms of tensors, invariance may be easily established.

It is apparent from (2.1), (2.3) and (2.4) that higher-order log likelihood derivatives and their moments do not usually transform as tensors. Since these quantities constitute the basic ingredients of the usual expansions considered in statistics, the study of their geometric structure represents a core problem for obtaining invariant Taylor series. Two main approaches have been followed. The first is coordinate-bound and relies upon the theory of *strings* developed in Barndorff-Nielsen (1986) and Barndorff-Nielsen and Blæsild (1987); tensors are recovered from strings, namely from quantities obeying transformation laws like (2.1), (2.3) and (2.4), by a process referred to as *intertwining*, which generalizes ideas in McCullagh and Cox (1986). The second approach, stemming also from McCullagh and Cox (1986), is pursued in Murray (1988), where a coordinate-free theory of geometric Taylor series is presented; the basic concept is that of *coordinate strings*, namely of local coordinate systems such that coefficients in Taylor series for scalar quantities transform as tensors. The close relation between this approach and the theory of intertwining is elucidated in the same paper. In this work we follow the former approach, which is of algorithmic nature and closer to the usual view in doing likelihood calculations.

Sequences of likelihood quantities such as $\{l_{R_m}\} = \{l_{R_m}(\omega), m = 1, 2, \ldots\}$, $\{\mu_{R_m}\} = \{\mu_{R_m}(\omega), m = 1, 2, \ldots\}$, $\{\mu_{r,S_m}\} = \{\mu_{r,S_m}(\omega), m = 1, 2, \ldots\}$ and $\{\mu_{r,s,T_m}\} = \{\mu_{r,s,T_m}(\omega), m = 1, 2, \ldots\}$ which follow transformation rules that generalize (2.1), (2.3) and (2.4) are all instances of *co-strings*. A co-string of *covariant degree k* is a (possibly finite) collection of smooth real functions $\{C_{R_k S_m}\} = \{C_{r_1 \cdots r_k s_1 \cdots s_m}(\omega), m = 1, 2, \ldots\}$, symmetric in the indices $s_1, \ldots, s_m$ and obeying, under re-parametrization, the transformation rule

$$C_{A_k B_m} = \left\{ \sum_{h=1}^{m} \sum_{B_m/h} C_{R_k s_1 \cdots s_h} \omega^{s_1}_{B_{m_1}} \cdots \omega^{s_h}_{B_{m_h}} \right\} \omega^{r_1}_{a_1} \cdots \omega^{r_k}_{a_k}. \tag{2.7}$$

The symbol B_m/h indicates summation over all *ordered* partitions of B_m into h non-empty subsequences $B_{m_1}, \ldots, B_{m_h}$, such that the order of indices in each of these subsequences is the same as their order in B_m and, for $j = 1, \ldots, h-1$, the first index in B_{m_j} comes before the first index in $B_{m_{j+1}}$ as compared with the ordering within B_m; the sum is defined as equal to zero when $h > m$. The sequences $\{l_{R_m}\}$, $\{\mu_{R_m}\}$ as well as $\{H_{R_m} = l_{R_m} - \mu_{R_m}\}$ are co-strings of covariant degree zero, while the sequences $\{\mu_{r,S_m}\}$ and $\{\mu_{r,s,T_m}\}$ constitute co-strings of covariant degree one and two, respectively. Notice that the first element $C_{R_k s}$ of a co-string of covariant degree k is a $(0, k+1)$ tensor.

The transformation rule of $\mu_{rs,tu}$ given by (2.4) is not of the form (2.7). Since $\mu_{R_m,S_n} = E_\omega(l_{R_m} l_{S_n})$, its transformation rule is obtained from the product of the transformation rules for l_{R_m} and l_{S_n}. The sequence $\{\mu_{R_m,S_n}\} = \{\mu_{R_m,S_n}(\omega), m, n = 1, 2, \ldots\}$ is an instance of a *double co-string* (of *covariant degree zero*) whose general definition may be found in Barndorff-Nielsen (1986).

The process of intertwining operates on co-strings and double co-strings (and, more generally, on *strings*) to generate tensors. For instance, starting from a co-string $\{C_{R_m}\}$ of covariant degree zero a sequence of $(0, m)$ tensors $\{t_{R_m}, m = 1, 2, \ldots\}$ may be obtained from the recursive relations

$$
\begin{aligned}
&C_r = t_r,\\
&C_{rs} = t_{rs} + \beta^v_{rs} t_v,\\
&C_{rst} = t_{rst} + \beta^v_{rs} t_{vt}[3] + \beta^v_{rst} t_v,\\
&C_{rstu} = t_{rstu} + \beta^v_{rs} t_{vtu}[6] + \beta^v_{rs}\beta^w_{tu} t_{vw}[3] + \beta^v_{rst} t_{vu}[4] + \beta^v_{rstu} t_v,
\end{aligned} \tag{2.8}
$$

and so on. The sequence $\{t_{R_m}\}$ obeys the transformation law for tensors, (2.6), provided that the sequence of coefficients $\{\beta^v_{rs}, \beta^v_{rst}, \beta^v_{rstu}, \ldots\}$, symmetric in the lower indices, transforms as a *connection string*, namely according to the relations

$$
\begin{aligned}
\beta^a_{bc} &= (\beta^r_{st}\omega^s_b\omega^t_c + \omega^r_{bc})\psi^a_r,\\
\beta^a_{bcd} &= (\beta^r_{stu}\omega^s_b\omega^t_c\omega^u_d + \beta^r_{st}\omega^s_{bc}\omega^t_d[3] + \omega^r_{bcd})\psi^a_r,\\
\beta^a_{bcde} &= (\beta^r_{stuv}\omega^s_b\omega^t_c\omega^u_d\omega^v_e + \beta^r_{stu}\omega^s_b\omega^t_c\omega^u_{de}[6]\\
&\quad + \beta^r_{st}\omega^s_b\omega^t_{cde}[4] + \beta^r_{st}\omega^s_{bc}\omega^t_{de}[3] + \omega^r_{bcde})\psi^a_r,
\end{aligned} \tag{2.9}
$$

and so on. It is convenient to put $\beta^v_r = \delta^v_r$, where δ^v_r is the Kronecker delta.

An instance of a connection string is $\{\beta^v_{R_m}\} = \{i^{vs}\mu_{s,R_m}\}$; it is used in McCullagh and Cox (1986) (see also McCullagh (1987), Section 7.2.3) to define a local parametrization such that log likelihood derivatives transform as tensors (McCullagh and Cox refer to log likelihood derivatives in such a parametrization as Möbius derivatives).

For co-strings of positive covariant degree and for double co-strings intertwining operates essentially in analogy with (2.8). Examples of tensors based on likelihood quantities obtained via intertwining are collected in Tables 1–3; their detailed description is deferred to Section 5.

Table 1. Tensors obtained by intertwining the co-string $\{H_{R_m}\}$ with respect to the 1 connection string $\{\overset{1}{\mathcal{G}}{}^r_{T_m}\} = \{i^{rs}\mu_{s,T_m}\}$.

$$
\begin{aligned}
T_r &= H_r = l_r\\
T_{rs} &= H_{rs} - \mu_{t,rs} l^t\\
T_{rst} &= H_{rst} - i^{vw}\mu_{v,rs}H_{wt}[3] - (\mu_{u,rst} - i^{vw}\mu_{v,rs}\mu_{u,tw}[3])l^u\\
T_{rstu} &= H_{rstu} - i^{vw}\mu_{v,rs}H_{tuw}[6] - i^{vw}\mu_{v,rst}H_{uw}[4]\\
&\quad + i^{vw}i^{xy}(\mu_{v,rs}\mu_{x,tu}H_{wy}[3] + \mu_{v,rs}\mu_{x,tw}H_{uy}[12])\\
&\quad + i^{vw}\{\mu_{v,rs}\mu_{z,tuw}[6] + \mu_{v,rst}\mu_{z,uw}[4] - \mu_{z,rst}\\
&\quad - i^{xy}(\mu_{v,rs}\mu_{x,tu}\mu_{z,wy}[3] + \mu_{v,rs}\mu_{x,tw}\mu_{z,uy}[12])\}l^z
\end{aligned}
$$

Note. The tensors $T^{-}_{R_m}$ with respect to the -1 connection string $\{\overset{-1}{\mathcal{G}}{}^r_{T_m}\} = \{i^{rs}\mu_{s;T_m}\}$ are obtained from the tensors T_{R_m} replacing μ_{u,T_m} with $\mu_{u;T_m}$.

Table 2. Tensors obtained by intertwining the co-string $\{\mu_{T_m}\}$ with respect to the -1 connection string $\{\overset{-1}{\mathcal{G}}{}^r_{T_m}\} = \{i^{rs}\mu_{s;T_m}\}$.

$\tau_{rs} = -i_{rs}$
$\tau_{rst} = \mu_{rst} + \mu_{r;st}[3] = 2\mu_{r,s,t}$
$\tau_{rstu} = \mu_{rstu} + \mu_{r;stu}[4] + i^{vw}\mu_{v;rs}\mu_{w;tu}[3] - i^{vw}\tau_{rsv}\mu_{w;tu}[6]$

Table 3. Tensors obtained by intertwining the double co-string $\{\mu_{R_m;S_n}\}$ with respect to the -1 connection string $\{\overset{-1}{\mathcal{G}}{}^r_{T_m}\} = \{i^{rs}\mu_{s;T_m}\}$.

$\tau_{t,u} = i_{tu}$
$\tau_{T_2;u} = \mu_{T_2;u} - \mu_{u;T_2}$
$\tau_{T_3;u} = \mu_{T_3;u} - \mu_{u;T_3} - i^{vw}\mu_{v;t_1t_2}(\mu_{wt_3;u} - \mu_{u;wt_3})[3]$
$\tau_{T_2;U_2} = \mu_{T_2;U_2} - i^{vw}\mu_{v;U_2}\mu_{T_2;w}$
$\tau_{T_3;U_2} = \mu_{T_3;U_2} - i^{vw}(\mu_{T_3;v}\mu_{w;U_2} + \mu_{v;t_1t_2}\mu_{t_3w;U_2}[3]) + i^{vw}i^{xy}\mu_{v;t_1t_2}\mu_{x;U_2}\mu_{t_3w;y}[3]$
$\tau_{T_2;U_3} = \mu_{T_2;U_3} - i^{vw}(\mu_{T_2;v}\mu_{w;U_3} + \mu_{v;u_1u_2}\mu_{T_2;u_3w}[3]) + i^{vw}i^{xy}\mu_{v;u_1u_2}\mu_{T_2;x}\mu_{y;u_3w}[3]$

3. Expected/observed likelihood expansions

Let f denote a real-valued function defined in a neighbourhood around ω. Starting from Lawley (1956), the traditional way to find expansions of $f(\hat\omega)$ around ω is to consider first the ordinary Taylor expansion

$$f(\hat\omega) = f(\omega) + f_r(\hat\omega-\omega)^r + \frac{1}{2}f_{rs}(\hat\omega-\omega)^{rs} + \frac{1}{6}f_{rst}(\hat\omega-\omega)^{rst} + \frac{1}{24}f_{rstu}(\hat\omega-\omega)^{rstu} + \cdots, \tag{3.1}$$

where $(\hat\omega-\omega)^{R_m} = (\hat\omega-\omega)^{r_1}\cdots(\hat\omega-\omega)^{r_m}$. An expansion for $(\hat\omega-\omega)^r$ is then obtained by inversion of the Taylor expansion around ω of the left-hand side of the likelihood equation $l_r(\hat\omega)=0$, namely of

$$l_r + \sum_{k=1}^{\infty}\frac{1}{k!}l_{rS_k}(\hat\omega-\omega)^{S_k} = 0. \tag{3.2}$$

Let us consider the quantities

$$H_{R_m} = l_{R_m} - \mu_{R_m}, \tag{3.3}$$

$$\mu^r_{S_k} = i^{rs}\mu_{sS_k}, \tag{3.4}$$

$$H^r_{S_k} = i^{rs}H_{sS_k}. \tag{3.5}$$

Under repeated sampling, i_{rs} and μ_{R_m} $(m>2)$ are of order $O(n)$, l_r and H_{R_m} are of order $O_P(n^{1/2})$ and by inversion of (3.2) we obtain

$$(\hat{\omega}-\omega)^r = A_1^r + A_2^r + A_3^r + A_4^r + O_P(n^{-5/2}), \tag{3.6}$$

where the term A_m^r is of order $O_P(n^{-m/2})$, $m=1,2,3,4$. These terms are given by

$$\begin{aligned}
A_1^r &= l^r,\\
A_2^r &= H_s^r l^s + \frac{1}{2}\mu_{st}^r l^{st},\\
A_3^r &= H_v^r H_s^v l^s + \frac{1}{2}H_{st}^r l^{st} + \mu_{tv}^r H_s^v l^{st} + \frac{1}{2}\mu_{st}^v H_v^r l^{st} + \frac{1}{6}\mu_{stu}^r l^{stu} + \frac{1}{2}\mu_{sv}^r \mu_{tu}^v l^{stu},\\
A_4^r &= H_v^r H_w^v H_s^w l^s + H_{vs}^r H_t^v l^{st} + \frac{1}{2}H_v^r H_{st}^v l^{st} + \mu_{vs}^r H_w^v H_t^w l^{st} + \frac{1}{2}\mu_{st}^v H_v^w H_w^r l^{st}\\
&\quad + \frac{1}{2}\mu_{vw}^r H_s^v H_t^w l^{st} + \mu_{ws}^v H_v^r H_t^w l^{st} + \frac{1}{6}H_{stu}^r l^{stu} + \frac{1}{2}\mu_{stv}^r H_u^v l^{stu}\\
&\quad + \frac{1}{6}\mu_{stu}^v H_v^r l^{stu} + \frac{1}{2}\mu_{st}^v H_{uv}^r l^{stu} + \frac{1}{2}\mu_{sv}^r H_{tu}^v l^{stu} + \mu_{sv}^r \mu_{tw}^v H_u^w l^{stu}\\
&\quad + \frac{1}{2}\mu_{vw}^r \mu_{st}^v H_u^w l^{stu} + \frac{1}{2}\mu_{sv}^r \mu_{tu}^w H_w^v l^{stu} + \frac{1}{2}\mu_{sw}^v \mu_{tu}^w H_v^r l^{stu}\\
&\quad + \frac{1}{24}\mu_{stuu'}^r l^{stuu'} + \frac{1}{6}\mu_{sv}^r \mu_{tuu'}^v l^{stuu'} + \frac{1}{4}\mu_{stv}^r \mu_{uu'}^v l^{stuu'}\\
&\quad + \frac{1}{8}\mu_{vw}^r \mu_{st}^v \mu_{uu'}^w l^{stuu'} + \frac{1}{2}\mu_{sv}^r \mu_{wt}^v \mu_{uu'}^w l^{stuu'}.
\end{aligned}$$

Let $f(\omega)$ be such that higher-order derivatives with respect to ω may be decomposed as

$$f_{R_m}(\omega) = \phi_{R_m} + F_{R_m} \qquad m=1,2,\ldots, \tag{3.7}$$

where $\phi_{R_m} = E_\omega(f_{R_m}(\omega))$ is of order $O(n^\beta)$ and the deviation $F_{R_m} = f_{R_m}(\omega) - \phi_{R_m}$ is of order $O_P(n^{\beta-1/2})$ or zero. The general form of the expected/observed likelihood expansion is obtained by substituting (3.6) into the Taylor expansion (3.1) of $f(\hat{\omega})$ taking into account the decomposition (3.7). The resulting formula is

$$f(\hat{\omega}) = f(\omega) + B_1 + B_2 + B_3 + B_4 + O_P(n^{\beta-5/2}), \tag{3.8}$$

where the terms B_m are of order $O_P(n^{\beta-m/2})$, $m=1,2,3,4$ and are given by

$$\begin{aligned}
B_1 &= \phi_r A_1^r,\\
B_2 &= F_r A_1^r + \phi_r A_2^r + \frac{1}{2}\phi_{rs} A_1^r A_1^s,\\
B_3 &= F_r A_2^r + \frac{1}{2}F_{rs} A_1^r A_1^s + \phi_r A_3^r + \phi_{rs} A_1^r A_2^s + \frac{1}{6}\phi_{rst} A_1^r A_1^s A_1^t,\\
B_4 &= F_r A_3^r + F_{rs} A_1^r A_2^s + \frac{1}{6}F_{rst} A_1^r A_1^s A_1^t + \phi_r A_4^r + \frac{1}{2}\phi_{rs} A_2^r A_2^s + \phi_{rs} A_1^r A_3^s\\
&\quad + \frac{1}{2}\phi_{rst} A_1^r A_1^s A_2^t + \frac{1}{24}\phi_{rstu} A_1^r A_1^s A_1^t A_1^u.
\end{aligned}$$

Note that the pattern displayed by the quantities B_m is simple and easy to extrapolate to higher-order terms.

Example. Expansion of the log likelihood ratio. If $f(\hat{\omega}) = l(\hat{\omega})$, the term B_1 in expansion (3.8) is zero and we obtain

$$l(\hat{\omega}) - l(\omega) = B_2 + B_3 + B_4 + O_P(n^{-3/2}), \tag{3.9}$$

with

$$B_2 = \frac{1}{2} i_{rs} l^{rs},$$
$$B_3 = \frac{1}{6}(\mu_{rst} l^{rst} + 3H_{rs} l^{rs}),$$
$$B_4 = \frac{1}{24}(\mu_{rstu} + 3\mu_{rsv}\mu^{v}_{tu}) l^{rstu} + \frac{1}{2}\mu_{rsv} H^{v}_{t} l^{rst} + \frac{1}{2} H_{rv} H^{v}_{s} l^{rs} + \frac{1}{6} H_{rst} l^{rst}.$$

It is not difficult to check that B_2, B_3 and B_4 are scalar functions.

In general, if the function f is parametrization invariant it is desirable that each term B_m in (3.8), $m = 1, 2, 3, 4$, turns out to be invariant. This fact is not guaranteed by the derivation of (3.8) outlined above, so that the transformation law of each term B_m has to be investigated. Invariance of B_1 and B_2 in (3.8) may be established after some algebra. After extremely longwinded calculations the same result is seen to hold for B_3 (Blæsild, personal communication). A direct check of invariance of B_4 seems to be out of reach by hand calculations or by to date available computer algebra packages such as that described in Kendall (1992). Moreover, by this route invariance of terms in (3.8) appears to be a somehow fortuitous fact. In the following sections the elicitation of the geometric structure of the expected/observed expansion will be carried out by an indirect route, namely by comparison with the invariant Taylor expansion based on the expected likelihood yoke.

4. Invariant Taylor series expansions; the expected likelihood expansions

In Section 2 we recalled how the log likelihood function induces in a natural way a series of geometric objects defined on $\mathcal{F}$. The identity $\mu_r = 0$ together with the Bartlett relations, the metric tensor i_{rs} and the connection string $\{i^{vs}\mu_{s,R_m}\}$ are the most notable instances. Assuming a more abstract point of view, analogous geometrical structures on $\mathcal{F}$ may be induced by the general class of scalar functions termed yokes (Barndorff-Nielsen (1987), Blæsild (1991)).

Let $\tilde{\Omega}$ be a copy of Ω (typically $\tilde{\Omega}$ represents the range space of the MLE). Consider a smooth scalar function $g(\omega, \tilde{\omega})$ defined on $\Omega \times \tilde{\Omega}$ and let us write

$$g_{R_m;S_n}(\omega, \tilde{\omega}) = \partial_{r_1} \cdots \partial_{r_m} \tilde{\partial}_{s_1} \cdots \tilde{\partial}_{s_n} g(\omega, \tilde{\omega}),$$

where $\partial_r = \partial/\partial\omega^r$, $\tilde{\partial}_s = \partial/\partial\tilde{\omega}^s$, and in addition let us define

$$\mathcal{G}_{R_m;S_n} = g_{R_m;S_n}(\omega, \omega).$$

The function g is said to be a yoke if for every ω

(i) $\mathcal{G}_{r;} = 0$,

(ii) $[\mathcal{G}_{r;s}]$ is non-singular.

For a recent review of differential geometries derived from yokes see Section 3 of Barndorff-Nielsen *et al.* (1991*b*). In particular, a yoke gives rise to a family of strings and connection strings; consequently it induces a family of tensors (Blæsild (1991)). If g is a yoke, $\mathcal{G}_{;} = \{\mathcal{G}_{R_m;S_n} : m, n = 1, 2, \ldots\}$ is a double co-string and a family of tensors based on g can be obtained via intertwining. We indicate with $\{\overset{1}{\mathcal{G}}{}^r_{T_m}\}$ and $\{\overset{-1}{\mathcal{G}}{}^r_{T_m}\}$ the 1 and -1 connection strings associated with the yoke g; they have, respectively, generic element

$$\overset{1}{\mathcal{G}}{}^r_{T_m} = \mathcal{G}^{rs}\mathcal{G}_{T_m;s} \qquad m = 1, 2, \ldots \tag{4.1}$$

and

$$\overset{-1}{\mathcal{G}}{}^r_{T_m} = \mathcal{G}^{rs}\mathcal{G}_{s;T_m} \qquad m = 1, 2, \ldots \tag{4.2}$$

where $[\mathcal{G}^{rs}] = [\mathcal{G}_{r;s}]^{-1}$.

The quantity $\tilde{\gamma}^r = \mathcal{G}^{rs} g_{s;}(\omega, \tilde{\omega})$ is a $(1, 0)$ tensor; products of like quantities, $\tilde{\gamma}^{R_m} = \tilde{\gamma}^{r_1} \cdots \tilde{\gamma}^{r_m}$, play the role of 'tensorial coordinates' in invariant Taylor series expansions. The coefficients f_{R_m} appearing in ordinary Taylor expansions are replaced in invariant Taylor series expansions by 'tensorial coefficients' $\overset{-1}{f}{}_{R_m}$. They are the tensorial components of f with respect to the -1 connection string $\{\overset{-1}{\mathcal{G}}{}^r_{T_m}\}$ and are defined recursively by the equations

$$f_{T_m} = \sum_{h=1}^{m} \overset{-1}{f}{}_{R_h} \overset{-1}{\mathcal{G}}{}^{R_h}_{T_m}, \tag{4.3}$$

where

$$\overset{-1}{\mathcal{G}}{}^{rs}_{T_m} = \sum_{T_m/2} \overset{-1}{\mathcal{G}}{}^r_{T_{m_1}} \overset{-1}{\mathcal{G}}{}^s_{T_{m_2}}, \qquad \overset{-1}{\mathcal{G}}{}^{rst}_{T_m} = \sum_{T_m/3} \overset{-1}{\mathcal{G}}{}^r_{T_{m_1}} \overset{-1}{\mathcal{G}}{}^s_{T_{m_2}} \overset{-1}{\mathcal{G}}{}^t_{T_{m_3}}, \tag{4.4}$$

and so on. If f is a scalar function, $\overset{-1}{f}{}_{T_m}$ is a $(0, m)$ tensor. The invariant Taylor series expansion of $f(\tilde{\omega})$ has thus the form

$$f(\tilde{\omega}) = f(\omega) + \sum_{m=1}^{\infty} \frac{1}{m!} \overset{-1}{f}{}_{R_m} \tilde{\gamma}^{R_m}. \tag{4.5}$$

From (4.3), the first four instances of $\overset{-1}{f}{}_{R_m}$ are

$$\begin{aligned}
\overset{-1}{f}{}_r &= f_r, \\
\overset{-1}{f}{}_{rs} &= f_{rs} - f_v \overset{-1}{\mathcal{G}}{}^v_{rs}, \\
\overset{-1}{f}{}_{rst} &= f_{rst} - f_v \overset{-1}{\mathcal{G}}{}^v_{rst} - (f_{rv} - f_x \overset{-1}{\mathcal{G}}{}^x_{rv}) \overset{-1}{\mathcal{G}}{}^v_{st}[3], \\
\overset{-1}{f}{}_{rstu} &= f_{rstu} - f_v \overset{-1}{\mathcal{G}}{}^v_{rstu} - \overset{-1}{f}{}_{rv} \overset{-1}{\mathcal{G}}{}^v_{stu}[4] \\
&\quad - \overset{-1}{f}{}_{vw} \overset{-1}{\mathcal{G}}{}^v_{rs} \overset{-1}{\mathcal{G}}{}^w_{tu}[3] - \overset{-1}{f}{}_{rsv} \overset{-1}{\mathcal{G}}{}^v_{tu}[6].
\end{aligned} \tag{4.6}$$

Expansion (4.5) does not depend on the choice of the parametrization of $\mathcal{F}$; of course it depends on the choice of the yoke. For statistical applications, two main instances of a yoke have been introduced: the expected likelihood yoke and the observed likelihood yoke. Expansions based on the former yoke turn out to be the most convenient starting point for the study of the geometric features of the traditional expected/observed expansions.

The expected likelihood yoke is defined, in coordinate form, as

$$g(\omega,\tilde{\omega}) = E_{\tilde{\omega}}(l(\omega) - l(\tilde{\omega})). \tag{4.7}$$

This yoke gives rise to the double co-string $\mu_{;} = \{\mu_{R_m;S_n}, m, n = 1, 2, \ldots\}$, where $\mu_{R_m;S_n} = \mathcal{G}_{R_m;S_n}$. The quantities $\mu_{R_m;S_n}$ are related to the joint moments $\mu_{R_m,S_n,\ldots,U_q}$ via the relations

$$\mu_{R_m;S_n} = \sum_{h=1}^{n}\sum_{S_n/h} \mu_{R_m,S_{n_1},\ldots,S_{n_h}}, \qquad \mu_{R_m;} = \mu_{R_m}. \tag{4.8}$$

The 1 and -1 connection strings $\{\overset{1}{\mathcal{G}}{}^r_{T_m}\}$ and $\{\overset{-1}{\mathcal{G}}{}^r_{T_m}\}$ based on the expected yoke have, respectively, generic element

$$\overset{1}{\mathcal{G}}{}^r_{T_m} = i^{rs}\mu_{s,T_m} \qquad m = 1, 2, \ldots \tag{4.9}$$

and

$$\overset{-1}{\mathcal{G}}{}^r_{T_m} = i^{rs}\mu_{s;T_m} \qquad m = 1, 2, \ldots. \tag{4.10}$$

The extended normal coordinates are products of the quantities $\tilde{\gamma}^r = \tilde{\mu}^r$, where

$$\tilde{\mu}^r = i^{rs}\tilde{\mu}_s = i^{rs}E_{\tilde{\omega}}(l_s(\omega)). \tag{4.11}$$

When $\tilde{\omega} = \hat{\omega}$ we write $\hat{\mu}^r = i^{rs}\hat{\mu}_s = i^{rs}E_{\hat{\omega}}(l_s(\omega))$ and the invariant Taylor series expansion of $f(\hat{\omega})$ based on the expected yoke may be expressed as

$$f(\hat{\omega}) = f(\omega) + \sum_{m=1}^{\infty}\frac{1}{m!}\overset{-1}{f}_{R_m}\hat{\mu}^{R_m}, \tag{4.12}$$

where $\hat{\mu}^{R_m} = \hat{\mu}^{r_1}\cdots\hat{\mu}^{r_m}$ and the quantities $\overset{-1}{f}_{R_m}$ denote the tensorial components of f with respect to the -1 connection string $\{i^{rs}\mu_{s;T_m}\}$. The first four instances of $\overset{-1}{f}_{R_m}$ are obtained from (4.6) using (4.10).

Two comments are useful in order to get a deeper insight into the impact in statistical applications of the expansion (4.12).

First, the coordinates $\hat{\mu}^r$ are rather unnatural from a statistical point of view: they are simply l^r in natural exponential families but otherwise they are not susceptible to a clear-cut interpretation and to closed form expression.

The second point is that (4.12) does not guarantee that the summands in each invariant term $\overset{-1}{f}_{R_m}\hat{\mu}^{R_m}$ are exactly of the same asymptotic order in n. In view of (4.6), and more generally of (4.3), a derivative f_{R_m} with zero expectation causes to the corresponding summand a collapse of the order in n, from $O_P(n^{\beta})$ to $O_P(n^{\beta-1/2})$. Consider for instance the expansion of the log likelihood ratio $l(\hat{\omega}) - l(\omega)$; $f_r = l_r$ is of order $O_P(n^{1/2})$, while $f_{rs} = l_{rs}$ is of order $O_P(n)$, so that $\overset{-1}{f}_{rs}\hat{\mu}^{rs}$ is the sum of a quantity of order $O_P(1)$ and of a quantity of order $O_P(n^{-1/2})$. The implication of this second remark is that if, for some $k > 1$, contributions of order $o_P(n^{\beta-k/2})$ are neglected, the term of order $O_P(n^{\beta-k/2})$ may lose invariance.

The latter point raised above is simpler to treat and will be addressed first; a rearrangement which preserves invariance may be obtained in analogy to the reasoning leading to (3.8). The answer to the former point is less straightforward and is deferred to Section 5. Let us assume, in addition to (3.7), that the function $f(\omega)$ itself satisfies the decomposition

$$f(\omega) = \phi(\omega) + F(\omega),$$

where

$$\phi(\omega) = E_{\omega}(f(\omega)),$$

with $\phi(\omega)$ of order $O(n^{\beta})$ and $F(\omega)$ of order $O_P(n^{\beta-1/2})$ or zero. Under reparametrization, the transformation laws for f, ϕ and F are the same. In particular, if f is a scalar function, ϕ and F are scalar functions as well. It follows that the expected likelihood expansion for $f(\hat{\omega})$ may be decomposed as the sum of the expected likelihood expansions of $\phi(\hat{\omega})$ and of $F(\hat{\omega})$. From (4.12) we have

$$\phi(\hat{\omega}) = \phi(\omega) + \sum_{m=1}^{\infty} \frac{1}{m!} \overset{-1}{\phi}_{R_m} \hat{\mu}^{R_m}, \tag{4.13}$$

$$F(\hat{\omega}) = F(\omega) + \sum_{m=1}^{\infty} \frac{1}{m!} \overset{-1}{F}_{R_m} \hat{\mu}^{R_m}, \tag{4.14}$$

where $\overset{-1}{\phi}_{R_m}$ is a sum of quantities of order $O(n^{\beta})$, while $\overset{-1}{F}_{R_m}$ is a sum of quantities of order $O_P(n^{\beta-1/2})$ or zero.

Adding (4.13) and (4.14) and collecting terms of like order, we have

$$f(\hat{\omega}) = f(\omega) + \phi_r\hat{\mu}^r + \sum_{m=2}^{\infty} \frac{1}{m!}(m\overset{-1}{F}_{R_{m-1}} + \overset{-1}{\phi}_{R_m}\hat{\mu}^{r_m})\hat{\mu}^{R_{m-1}}, \tag{4.15}$$

where $m\overset{-1}{F}_{R_{m-1}} + \overset{-1}{\phi}_{R_m}\hat{\mu}^{r_m}$ is of order $O_P(n^{\beta-1/2})$ and is itself a $(0, m-1)$ tensor.

Example (continued). *Expansion of the log likelihood ratio.* Let $f(\hat{\omega}) = l(\hat{\omega})$. Under repeated sampling $\beta = 1$ and the basic quantities needed to compute (4.15) with error of order $O_P(n^{-3/2})$ are $\phi_r = 0$, $\phi_{rs} = -i_{rs}$, $\phi_{R_m} = \mu_{R_m}$; $F_r = l_r$, $F_{rs} = H_{rs}$, $F_{R_m} = H_{R_m}$, $m \geq 3$. Expansion (4.15) gives

$$l(\hat{\omega}) - l(\omega) = b_2 + b_3 + b_4 + O_P(n^{-3/2}), \tag{4.16}$$

where

$$\begin{aligned}
b_2 &= \frac{1}{2}\{2l_r - i_{rs}\hat{\mu}^s\}\hat{\mu}^r,\\
b_3 &= \frac{1}{6}\{3(H_{rs} - l^t\mu_{t;rs}) + (\mu_{rst} + \mu_{r;st}[3])\hat{\mu}^t\}\hat{\mu}^{rs},\\
b_4 &= \frac{1}{24}\{4(H_{rst} - l^u\mu_{u;rst} - (H_{rv} - l^u\mu_{u;rv})i^{vw}\mu_{w;st}[3])\\
&\quad + \{\mu_{rstu} + \mu_{r;stu}[4] + i^{vw}\mu_{v;rs}\mu_{w;tu}[3] - i^{vw}\mu_{rsv}\mu_{w;tu}[6]\\
&\quad - i^{vw}\mu_{r;sv}\mu_{w;tu}[12] - i^{vw}\mu_{v;rs}\mu_{w;tu}[6]\}\hat{\mu}^u\}\hat{\mu}^{rst}.
\end{aligned}$$

The terms in expansion (4.16) may be written in a compact, explicitly tensorial form as

$$\begin{aligned}
b_2 &= \frac{1}{2}(2T_r^- + \tau_{rs}\hat{\mu}^s)\hat{\mu}^r,\\
b_3 &= \frac{1}{6}(3T_{rs}^- + \tau_{rst}\hat{\mu}^t)\hat{\mu}^{rs},\\
b_4 &= \frac{1}{24}(4T_{rst}^- + \tau_{rstu}\hat{\mu}^u)\hat{\mu}^{rst},
\end{aligned}$$

where $T_{R_m}^-$ and τ_{R_m} are $(0, m)$ tensors obtained by intertwining the co-strings $\{H_{R_m}\}$ and $\{\mu_{R_m}\}$ with respect to the -1 connection string $\{i^{rs}\mu_{s;T_m}\}$. The expressions of these tensors are collected in Tables 1 and 2.

5. Expected/observed likelihood expansions as expected likelihood expansions with expanded coordinates

Expansion (4.15) is still unsuitable for statistical applications, due to the presence of the coordinates $\hat{\mu}^r$. The first objective of this section is to show that $\hat{\mu}^r$ can be given an expansion of the form

$$\hat{\mu}^r = \hat{\mu}_1^r + \hat{\mu}_2^r + \hat{\mu}_3^r + \hat{\mu}_4^r + O_P(n^{-5/2}), \tag{5.1}$$

where each term $\hat{\mu}_m^r$ $(m = 1, 2, 3, 4)$:

(i) is of order $O_P(n^{-m/2})$;

(ii) is such that $\hat{\mu}_m^r = \hat{\mu}_m^a\omega_a^r$, i.e., like $\hat{\mu}^r$, it behaves as a $(1, 0)$ tensor under re-parametrization;

(iii) depends on moments of log likelihood derivatives, with expectation taken under ω, and on the quantities H_{R_k}, $k \leq m$, defined by (3.3).

Consider the expected likelihood expansion around ω, given by (4.15), of the left-hand side of the likelihood equation $f(\hat{\omega}) = l_z(\hat{\omega}) = 0$ $(z = 1, \ldots, d)$. We have

$\phi_r = -i_{zr}$, $\phi_{rs} = \mu_{zrs}$, $\phi_{rst} = \mu_{zrst}$; $F_r = H_{zr}$, $F_{rs} = H_{zrs}$, $F_{rst} = H_{zrst}$ and we obtain

$$0 = l_z + {}_1b_z + {}_2b_z + {}_3b_z + {}_4b_z + O_P(n^{-3/2}), \tag{5.2}$$

where

$$\begin{aligned}
{}_1b_z &= -\hat{\mu}_z, \\
{}_2b_z &= T_{r\cdot z}\hat{\mu}^r + \tau_{rs\cdot z}\hat{\mu}^{rs}, \\
{}_3b_z &= T_{rs\cdot z}\hat{\mu}^{rs} + \tau_{rst\cdot z}\hat{\mu}^{rst}, \\
{}_4b_z &= T_{rst\cdot z}\hat{\mu}^{rst} + \tau_{rstu\cdot z}\hat{\mu}^{rstu},
\end{aligned}$$

with

$$\begin{aligned}
T_{r\cdot z} &= H_{zr}, \\
\tau_{rs\cdot z} &= \frac{1}{2}(\mu_{zrs} + \mu_{z;rs}), \\
T_{rs\cdot z} &= \frac{1}{2}(H_{zrs} - H_{zv}i^{vw}\mu_{w;rs}), \\
\tau_{rst\cdot z} &= \frac{1}{6}(\mu_{zrst} + \mu_{z;rst} - (\mu_{zrv} + \mu_{z;rv})i^{vw}\mu_{w;st}[3]), \\
T_{rst\cdot z} &= \frac{1}{6}(H_{zrst} - H_{zv}i^{vw}\mu_{w;rst} - (H_{zrv} - H_{zx}i^{xy}\mu_{y;rv})i^{vw}\mu_{w;st}[3]), \\
\tau_{rstu\cdot z} &= \frac{1}{24}\{\mu_{zrstu} + \mu_{z;rstu} - (\mu_{zrv} + \mu_{z;rv})i^{vw}\mu_{w;stu}[4] \\
&\quad - (\mu_{zvx} + \mu_{z;vx})i^{vw}i^{xy}\mu_{w;rs}\mu_{y;tu}[3] - (\mu_{zrsv} + \mu_{z;rsv})i^{vw}\mu_{w;tu}[6] \\
&\quad + (\mu_{zrx} + \mu_{z;rx})i^{vw}i^{xy}\mu_{y;sv}\mu_{w;tu}[12] \\
&\quad + (\mu_{zvx} + \mu_{z;vx})i^{vw}i^{xy}\mu_{y;rs}\mu_{w;tu}[6]\}.
\end{aligned}$$

Expansion (5.2) relates components l_z of the score vector to the coordinates $\hat{\mu}^z$. Apart from the coefficients $1/2, 1/6, \ldots$, the quantities $T_{R_m\cdot z}$ and $\tau_{R_m\cdot z}$ may be obtained by treating formally the sequences $\{H_{zR_m}, m = 1, 2, \ldots\}$ and $\{\mu_{zR_m}, m = 1, 2, \ldots\}$ as co-strings of covariant degree zero and intertwining them with respect to the -1 connection string (4.10). However, $T_{R_m\cdot z}$ and $\tau_{R_m\cdot z}$ do not behave as tensors.

We may express $\hat{\mu}_z$ as a function of l_z in the form

$$\hat{\mu}_z = l_z + {}_2b_z + {}_3b_z + {}_4b_z + O_P(n^{-3/2}), \tag{5.3}$$

and inverting the above asymptotic expansion we obtain

$$\begin{aligned}
{}_2b_z &= {}_{21}b_z + {}_{22}b_z + {}_{23}b_z + O_P(n^{-3/2}), \\
{}_3b_z &= {}_{31}b_z + {}_{32}b_z + O_P(n^{-3/2}), \\
{}_4b_z &= {}_{41}b_z + O_P(n^{-3/2}),
\end{aligned}$$

where

$$
\begin{aligned}
{}_{21}b_z &= T_{r\cdot z}l^r + \tau_{rs\cdot z}l^{rs},\\
{}_{22}b_z &= i^{vw}\{T_{v\cdot z}T_{r\cdot w}l^r + T_{v\cdot z}\tau_{rs\cdot w}l^{rs} + 2T_{r\cdot v}\tau_{sw\cdot z}l^{rs}\\
&\qquad + 2\tau_{rv\cdot z}\tau_{st\cdot w}l^{rst}\},\\
{}_{23}b_z &= i^{vw}\{T_{v\cdot z}T_{rs\cdot w}l^{rs} + T_{v\cdot z}\tau_{rst\cdot w}l^{rst}\\
&\qquad + 2T_{rs\cdot v}\tau_{tw\cdot z}l^{rst} + 2\tau_{rv\cdot z}\tau_{stu\cdot w}l^{rstu}\}\\
&\quad + i^{vw}i^{xy}\{T_{v\cdot z}T_{x\cdot w}T_{r\cdot y}l^r + T_{v\cdot z}T_{x\cdot w}\tau_{rs\cdot y}l^{rs}\\
&\qquad + T_{r\cdot v}T_{s\cdot x}\tau_{wy\cdot z}l^{rs} + 2T_{v\cdot z}T_{r\cdot x}\tau_{sy\cdot w}l^{rs}\\
&\qquad + 2T_{x\cdot v}T_{r\cdot y}\tau_{sw\cdot z}l^{rs} + 2T_{v\cdot z}\tau_{rx\cdot w}\tau_{st\cdot y}l^{rst}\\
&\qquad + 2T_{r\cdot v}\tau_{wx\cdot z}\tau_{st\cdot y}l^{rst} + 2T_{x\cdot v}\tau_{rw\cdot z}\tau_{st\cdot y}l^{rst}\\
&\qquad + 4T_{r\cdot v}\tau_{sx\cdot z}\tau_{tw\cdot y}l^{rst} + \tau_{vx\cdot z}\tau_{rs\cdot w}\tau_{tu\cdot y}l^{rstu}\\
&\qquad + 4\tau_{rv\cdot z}\tau_{sx\cdot w}\tau_{tu\cdot y}l^{rstu}\},\\
{}_{31}b_z &= T_{rs\cdot z}l^{rs} + \tau_{rst\cdot z}l^{rst},\\
{}_{32}b_z &= 2i^{vw}T_{rv\cdot z}T_{s\cdot w}l^{rs} + 2i^{vw}T_{rv\cdot z}\tau_{st\cdot w}l^{rst}\\
&\quad + 3i^{vw}\tau_{rsv\cdot z}T_{t\cdot w}l^{rst} + 3i^{vw}\tau_{rsv\cdot z}\tau_{tu\cdot w}l^{rstu},\\
{}_{41}b_z &= T_{rst\cdot z}l^{rst} + \tau_{rstu\cdot z}l^{rstu}.
\end{aligned}
$$

Rearranging the above expansions of the summands in (5.3) according to their order in n and recalling that $\hat{\mu}^r = i^{rz}\hat{\mu}_z$, we obtain an expansion for $\hat{\mu}^r$ of the desired form (5.1), with

$$
\begin{aligned}
&\hat{\mu}^r_1 = l^r, \qquad \hat{\mu}^r_2 = i^{rs}{}_{21}b_s,\\
&\hat{\mu}^r_3 = i^{rs}({}_{31}b_s + {}_{22}b_s), \qquad \hat{\mu}^r_4 = i^{rs}({}_{41}b_s + {}_{32}b_s + {}_{23}b_s).
\end{aligned}
$$

It is not difficult to see that ${}_{21}b_s = H_{st}l^t + \frac{1}{2}(\mu_{stu} + \mu_{s;tu})l^{tu}$ behaves as a $(0,1)$ tensor under re-parametrizations. This is not the case for ${}_{31}b_s$ or ${}_{22}b_s$ individually, while their sum is again a $(0,1)$ tensor. Indeed, after some algebra, $\hat{\mu}^r_2$, $\hat{\mu}^r_3$ and $\hat{\mu}^r_4$ may be written in the following tensorial form which demonstrates that each of them is a $(1,0)$ tensor:

$$\hat{\mu}^r_2 = T^r_s l^s, \tag{5.4}$$

$$\hat{\mu}^r_3 = T^r_v T^v_s l^s + \frac{1}{2}(T^r_{st} - i^{vw}\mu_{s,t,w}T^r_v)l^{st} - \frac{1}{2}i^{rv}\tau_{sv;tu}l^{stu}, \tag{5.5}$$

$$
\begin{aligned}
\hat{\mu}^r_4 &= T^r_v T^v_w T^w_s l^s\\
&\quad + \frac{1}{2}\{2T^v_s(T^r_{tv} - i^{xy}\mu_{t,v,x}T^r_y) + T^r_v(T^v_{st} - i^{xy}\mu_{s,t,x}T^v_y)\}l^{st}\\
&\quad + \frac{1}{6}\{T^r_{stu} - i^{vw}\mu_{s,t,u,v}T^r_w - 3i^{rv}(\mu_{s,t,w}T^w_{uv}\\
&\qquad + (\tau_{vw;st} + \tau_{sv;tw})T^w_u + (\tau_{sw;tu} + \tau_{s,w,tu})T^w_v)\}l^{stu}\\
&\quad - \frac{1}{24}i^{rv}(4\tau_{sv;tuu'} + 6\tau_{stv;uu'}\\
&\qquad - 12i^{xy}\tau_{sv;x}\tau_{ty;uu'} + \mu_{s,t,u,u',v})l^{stuu'},
\end{aligned}
\tag{5.6}
$$

where $T^r_{R_k} = i^{rs} T_{sR_k}$ and the symbols T_{R_m} denote $(0, m)$ tensors obtained by intertwining the co-string $\{H_{R_m}\}$ with respect to the 1 connection string $\{i^{rs}\mu_{s,T_m}\}$. In addition, the symbols $\tau_{T_m;U_n}$ denote the $(0, m+n)$ tensors obtained by intertwining the double co-string $\{\mu_{T_m;U_n}\}$ with respect to the -1 connection string $\{i^{rs}\mu_{s;T_m}\}$. The relevant instances of the tensors T_{R_m} and $\tau_{T_m;U_n}$ are given in Tables 1 and 3, respectively. The additional quantity appearing in $\hat{\mu}^r_4$, namely $\tau_{s,w,tu}$, is defined as

$$\tau_{s,w,tu} = \mu_{s,w,tu} - i^{xy}\mu_{y;tu}\mu_{s,w,x}$$

and it is a $(0, 4)$ tensor obtained by intertwining the co-string $\{\mu_{r,s,T_m}\}$ with respect to the -1 connection string $\{i^{rs}\mu_{s;T_m}\}$.

Substituting the expansion of the form (5.1) obtained above into the first four terms of (4.15) and rearranging according to the order in n, we may write

$$f(\hat{\omega}) = f(\omega) + b_1 + b_2 + b_3 + b_4 + O_P(n^{\beta-5/2}), \tag{5.7}$$

where b_m is of order $O_P(n^{\beta-m/2})$, $m = 1, 2, 3, 4$; in particular:

$$\begin{aligned}
b_1 &= \phi_r l^r, \\
b_2 &= F_r l^r + \phi_r \hat{\mu}^r_2 + \frac{1}{2}\overset{-1}{\phi}_{rs} l^{rs}, \\
b_3 &= F_r \hat{\mu}^r_2 + \frac{1}{2}\overset{-1}{F}_{rs} l^{rs} + \phi_r \hat{\mu}^r_3 + \overset{-1}{\phi}_{rs} l^r \hat{\mu}^s_2 + \frac{1}{6}\overset{-1}{\phi}_{rst} l^{rst}, \\
b_4 &= F_r \hat{\mu}^r_3 + \overset{-1}{F}_{rs} l^r \hat{\mu}^s_2 + \frac{1}{6}\overset{-1}{F}_{rst} l^{rst} + \phi_r \hat{\mu}^r_4 \\
&\quad + \frac{1}{2}\overset{-1}{\phi}_{rs} \hat{\mu}^r_2 \hat{\mu}^s_2 + \overset{-1}{\phi}_{rs} l^r \hat{\mu}^s_3 + \frac{1}{2}\overset{-1}{\phi}_{rst} l^{rs} \hat{\mu}^t_2 + \frac{1}{24}\overset{-1}{\phi}_{rstu} l^{rstu},
\end{aligned}$$

with $\hat{\mu}^r_2$, $\hat{\mu}^r_3$, $\hat{\mu}^r_4$ given by (5.4), (5.5), (5.6).

If f is a scalar function each b_m is a sum of scalar contributions of the same order in n. There is a strict analogy between expansion (5.7) and expansion (3.8). Our objective now is to show that they actually coincide. This fact seems to be quite natural since the same basic ingredients (namely l^r, μ_{R_m}, H_{R_m}) are present in both expansions.

Example (continued). *Expansion for the log likelihood ratio.* After a little algebra, expansion (5.7) for the log likelihood ratio statistic is seen to be

$$l(\hat{\omega}) - l(\omega) = b_2 + b_3 + b_4 + O_P(n^{-3/2}), \tag{5.8}$$

where $b_2 = \frac{1}{2} i_{rs} l^{rs}$, $b_3 = \frac{1}{2} T^{-}_{rs} l^{rs} + \frac{1}{6}\tau_{rst} l^{rst}$, and $b_4 = \frac{1}{2} T_{rv} T^v_s l^{rs} + \frac{1}{6} T^{-}_{rst} l^{rst} + \frac{1}{24}\tau_{rstu} l^{rstu}$. It is very easy to check that b_2, b_3 and b_4 coincide respectively with B_2, B_3 and B_4 of formula (3.9), so that expansion (5.8) simply rewrites in tensor form the expected/observed expansion (3.9).

As a first step for establishing the desired coincidence, consider the expansion of $f(\hat{\omega}) = \hat{\omega}^r$ according to (5.7). Note that $f_s = \delta^r_s$ and that, consequently, $\phi_s = f_s$ and $F_s = 0$. It follows that, for $n > 1$, $f_{S_n} = \phi_{S_n} = F_{S_n} = 0$. We obtain

$$(\hat{\omega} - \omega)^r = A^r_1 + A^r_2 + A^r_3 + A^r_4 + O_P(n^{-5/2}), \tag{5.9}$$

where

$$A_1^r = l^r, \tag{5.10}$$

$$A_2^r = \hat{\mu}_2^r - \frac{1}{2}\overset{-1}{\mathcal{G}}{}^r_{st} l^{st}, \tag{5.11}$$

$$A_3^r = \hat{\mu}_3^r - \overset{-1}{\mathcal{G}}{}^r_{st} l^s \hat{\mu}_2^t - \frac{1}{6}\big(\overset{-1}{\mathcal{G}}{}^r_{stu} - \overset{-1}{\mathcal{G}}{}^r_{vw}\overset{-1}{\mathcal{G}}{}^{vw}_{stu}\big) l^{stu}, \tag{5.12}$$

$$\begin{aligned} A_4^r = {} & \hat{\mu}_4^r - \frac{1}{2}\overset{-1}{\mathcal{G}}{}^r_{st}\hat{\mu}_2^s\hat{\mu}_2^t - \overset{-1}{\mathcal{G}}{}^r_{st} l^s \hat{\mu}_3^t \\ & - \frac{1}{2}\big(\overset{-1}{\mathcal{G}}{}^r_{stu} - \overset{-1}{\mathcal{G}}{}^r_{vw}\overset{-1}{\mathcal{G}}{}^{vw}_{stu}\big) l^{st}\hat{\mu}_2^u \\ & - \frac{1}{24}\big(\overset{-1}{\mathcal{G}}{}^r_{stuu'} - \overset{-1}{\mathcal{G}}{}^r_{vw}\overset{-1}{\mathcal{G}}{}^{vw}_{stuu'} \\ & \qquad + \overset{-1}{\mathcal{G}}{}^r_{xy}\overset{-1}{\mathcal{G}}{}^{xy}_{vwz}\overset{-1}{\mathcal{G}}{}^{vwz}_{stuu'} - \overset{-1}{\mathcal{G}}{}^r_{vwz}\overset{-1}{\mathcal{G}}{}^{vwz}_{stuu'}\big) l^{stuu'}, \end{aligned} \tag{5.13}$$

with $\overset{-1}{\mathcal{G}}{}^r_{T_m}$ given by (4.10) and $\overset{-1}{\mathcal{G}}{}^{rs}_{T_m}$, $\overset{-1}{\mathcal{G}}{}^{rst}_{T_m}$ given by (4.4).

After some algebra, it is seen that each term in expansion (5.9) coincides with the corresponding term appearing in (3.6). For instance, (5.11) may be written as

$$\begin{aligned} \hat{\mu}_2^r - \frac{1}{2} i^{rv}\mu_{v;st} l^{st} &= T_s^r l^s - \frac{1}{2} i^{rv}\mu_{v;st} l^{st} \\ &= i^{rv}(H_{vs} - \mu_{t,sv} l^t) l^s - \frac{1}{2} i^{rv}(\mu_{v,st} + \mu_{v,s,t}) l^{st} \\ &= H_s^r l^s - \frac{1}{2} i^{rv}(\mu_{t,sv}[2] + \mu_{v,st} + \mu_{v,s,t}) l^{st} \\ &= H_s^r l^s + \frac{1}{2} i^{rv}\mu_{vst} l^{st} \\ &= H_s^r l^s + \frac{1}{2}\mu^r_{st} l^{st}, \end{aligned}$$

which coincides with the term A_2^r in (3.6). In the above calculations, the balance relation $\mu_{rst} + \mu_{r,st}[3] + \mu_{r,s,t} = 0$ is used.

The final step consists in rewriting (3.8) according to the expressions (5.10)–(5.13) of A_1^r, A_2^r, A_3^r, A_4^r and in comparing the expressions of B_1, B_2, B_3, B_4 thus obtained with the corresponding terms b_1, b_2, b_3, b_4 in (5.7). Again, calculations that are lengthy only for B_4 give the desired identities $B_1 = b_1$, $B_2 = b_2$, $B_3 = b_3$, $B_4 = b_4$. The first identity is trivial; we limit ourselves to giving the detailed calculations leading to the second one. Taking into account (5.10) and (5.11) we have

$$\begin{aligned} B_2 &= F_r l^r + \phi_r\left(\hat{\mu}_2^r - \frac{1}{2}\overset{-1}{\mathcal{G}}{}^r_{st} l^{st}\right) + \frac{1}{2}\phi_{rs} l^{rs} \\ &= F_r l^r + \phi_r\hat{\mu}_2^r + \frac{1}{2}\big(\phi_{rs} - \phi_v\overset{-1}{\mathcal{G}}{}^v_{rs}\big) l^{rs} \\ &= F_r l^r + \phi_r\hat{\mu}_2^r + \frac{1}{2}\overset{-1}{\phi}{}_{rs} l^{rs} = b_2. \end{aligned}$$

The further identities $B_3 = b_3$ and $B_4 = b_4$ are established following essentially the same pattern. When written according to (5.10)–(5.13), A^r_2, A^r_3 and A^r_4 are seen to contain the combinations of elements of the -1 connection string $\{\overset{-1}{\mathcal{G}}{}^{r}_{T_m}\}$ needed to reconstruct the tensorial derivatives $\overset{-1}{F}_{R_m}$, $\overset{-1}{\phi}_{R_m}$ appearing in b_3 and b_4.

Acknowledgements

The authors would like to thank Professor Ole E. Barndorff-Nielsen, Dr. Peter Jupp and Dr. Preben Blæsild for helpful discussions and comments on earlier versions of the paper. Thanks are also expressed to the anonymous referees whose comments led to improvements in the presentation.

References

Barndorff-Nielsen, O. E. (1986). Strings, tensorial combinants and Bartlett adjustments, *Proc. Roy. Soc. London Ser. A*, **406**, 127–137.

Barndorff-Nielsen, O. E. (1987). Differential geometry and statistics: some mathematical aspects, *Indian J. Math.*, **29**, 335–350.

Barndorff-Nielsen, O. E. (1989). Comment to Kass, R. E.: the geometry of asymptotic inference, *Statist. Sci.*, **4**, 188–234.

Barndorff-Nielsen, O. E. and Blæsild, P. (1987). Strings: mathematical theory and statistical examples, *Proc. Roy. Soc. London Ser. A*, **411**, 155–176.

Barndorff-Nielsen, O. E., Blæsild, P., Pace, L. and Salvan, A. (1991*a*). Formulas for asymptotic statistical calculations, Research Rep., No. 207, Department of Theoretical Statistics, Aarhus University, Denmark.

Barndorff-Nielsen, O. E., Jupp, P. E. and Kendall, W. S. (1991*b*). Stochastic calculus, statistical asymptotics, Taylor strings and phyla, *Annales de Toulouse* (to appear).

Blæsild, P. (1990). Yokes: orthogonal and extended normal coordinates, Research Rep., No. 205, Department of Theoretical Statistics, Aarhus University, Denmark.

Blæsild, P. (1991). Yokes and tensors derived from yokes, *Ann. Inst. Statist. Math.*, **43**, 95–113.

Kendall, W. S. (1992). Computer algebra and yoke geometry I: when is an expression a tensor?, Research Rep., No. 237, Department of Statistics, University of Warwick, U.K.

Lawley, D. M. (1956). A general method for approximating to the distribution of the likelihood-ratio criteria, *Biometrika*, **43**, 295–303.

McCullagh, P. (1987). *Tensor Methods in Statistics*, Chapman and Hall, London.

McCullagh, P. and Cox, D. R. (1986). Invariants and likelihood ratio statistics, *Ann. Statist.*, **14**, 1419–1430.

Murray, M. K. (1988). Coordinate systems and Taylor series in statistics, *Proc. Roy. Soc. London Ser. A*, **415**, 445–452.

Murray, M. K. and Rice, J. W. (1993). *Differential Geometry and Statistics*, Chapman and Hall, London.