Mathematica Slovaca, **49**(1999), pages 577–598

ASYMPTOTIC DISTRIBUTION OF THE LIKELIHOOD RATIO TEST STATISTIC IN THE MULTISAMPLE CASE

František Rublík

ABSTRACT. Classical results on asymptotic distribution of the likelihood ratio test statistic are extended to multipopulation setting. The assertions include a statement on asymptotic distribution in the case of linear hypotheses and a statement on asymptotic distribution for the hypotheses approximable by cones. The later framework includes usual smooth hypotheses and is dealt with under validity of local alternatives.

1. Introduction and the main results.

Suppose that probabilities $\{\,\overline{P}_\gamma;\ \gamma\in\Xi\,\}$ are defined by means of densities $\{\,f(x,\gamma);\ \gamma\in\Xi\,\}$ with respect to a measure ν on $(X,\mathcal{S})$. Let

$$L(x_1,\dots,x_n,\Omega)=\sup\Big\{\,\prod_{i=1}^{n} f(x_i,\gamma);\ \gamma\in\Omega\,\Big\}\,. \tag{1.1}$$

According to the classical Wilks' result

$$\mathcal{L}\left[2\log\frac{L(x_1,\dots,x_n,\Xi)}{L(x_1,\dots,x_n,\Omega)}\ \Big|\ \overline{P}_\gamma\right]\longrightarrow\chi^2_k \tag{1.2}$$

as $n\to\infty$, provided that $\Xi\subset R^m$, γ belongs to $\Omega=\{\,\gamma\in\Xi;\gamma_1=0,\dots,\dots,\gamma_k=0\,\}$, log denotes the logarithm to the base e and certain regularity conditions are fulfilled. It is also well-known that (1.2) holds with a more general hypothesis $\Omega=\{\gamma\in\Xi;\ g_1(\gamma)=0,\dots,g_k(\gamma)=0\,\}$ provided that the underlying

1991 Mathematics Subject Classification: Primary 62 F 05, 62 A 10.

Key words: Asymptotic distribution, hypotheses approximable by cones, likelihood ratio test, linear hypotheses, local alternatives, local asymptotic normality, multipopulation case.

The research was supported by the grant no. 95/5305/468 from VEGA.

functions possess continuous partial derivatives which form a full rank matrix. The proof can be found e.g. in Section 6e.3 of [9], on pp. 240 – 242 of [11] or on pp. 156 – 160 of [12].

However, the mentioned results do not cover the variety of the testing problems when sampling is made from several populations and a hypothesis on the overall parameter is tested. These multipopulation hypotheses have to be handled from case to case, because in typical situations sample sizes from individual populations are not mutually equal and therefore the i.i.d. scheme cannot be employed for finding the limiting distribution.

The aim of this paper is to provide general assertions of the type (1.2) in the multipopulation case. Throughout the paper we assume that $q \geq 1$ is an arbitrary but fixed positive integer denoting the number of underlying statistical populations. The parameter space of overall parameters is the q-fold Cartesian product

$$\Theta = \Xi^q ,$$

where in $\theta = (\theta_1^T, \ldots, \theta_q^T)^T \in \Theta$ the symbol θ_j stands for parameter of the jth population and the superscript ^{T} denotes the transpose of the vector. The outcome of the sampling from the jth population will be denoted by

$$x(j, n_j) = (x_1^{(j)}, \ldots, x_{n_j}^{(j)}) . \tag{1.3}$$

Thus

$$x_{(n_1,\ldots,n_q)} = \Big(x(1, n_1), \ldots, x(q, n_q)\Big) \tag{1.4}$$

is the pooled sample and its distribution is the product measure

$$P_\theta^{(n_1,\ldots,n_q)} = \overline{P}_{\theta_1}^{(n_1)} \times \ldots \times \overline{P}_{\theta_q}^{(n_q)} , \tag{1.5}$$

where $\overline{P}_{\theta_j}^{(n_j)}$ is the product measure of n_j copies of $\overline{P}_{\theta_j}$.

The asymptotic results of the paper are based on the assumption that

$$n_j \longrightarrow +\infty , \;\; j = 1, \ldots, q. \tag{1.6}$$

Here it is tacitly assumed that $n_j = n_j^{(u)}$ denotes sample size from the jth population in the u-th experiment, $u = 1, 2, \ldots$, and the limits in (1.6) are related to u tending to infinity, but to avoid abundant indexing the index of the order of the experiment is omitted.

For $\Omega \subset \Theta$ let

$$L(x_{(n_1,\ldots,n_q)}, \Omega) = \sup \left\{ \prod_{j=1}^{q} \prod_{i=1}^{n_j} f(x_i^{(j)}, \theta_j);\; (\theta_1^T, \ldots, \theta_q^T)^T \in \Omega \right\} . \tag{1.7}$$

It will be shown in various settings that under validity of (1.6) the weak convergence

$$\mathcal{L} \left[2 \log \frac{L(x_{(n_1,\ldots,n_q)}, \Theta)}{L(x_{(n_1,\ldots,n_q)}, \Omega_0)} \,\Big|\, P_\theta^{(n_1,\ldots,n_q)} \right] \longrightarrow \chi_s^2 \tag{1.8}$$

or

$$\mathcal{L}\left[2\log\frac{L(x_{(n_1,\ldots,n_q)},\Omega_1)}{L(x_{(n_1,\ldots,n_q)},\Omega_0)}\,\Big|\,P_\theta^{(n_1,\ldots,n_q)}\right] \longrightarrow \chi_s^2 \tag{1.9}$$

holds, where χ_s^2 denotes the chi-square distribution with s degrees of freedom.

Methods of the proofs used in this paper are based on the fact, that the logarithm of the likelihood ratio is asymptotically equivalent to the difference of distances of the MLE from the hypotheses, which was in the one sample case for hypotheses sequentially approximable by disjoint cones established in the proof of Theorem 1 in [4]. One of the tools which we use in the proofs is a multipopulation variant of the Chernoff lemma, presented in Lemma 2.3. Asymptotic distribution of the LR statistics in the case of linear hypotheses is the topic of Theorem 1.1 and is derived by means of the weak convergence result presented in Lemma 2.5. A general class of hypotheses which includes also the hypotheses of order restrictions studied in [13], is handled in Theorem 1.2. The asymptotic distribution of the LR statistics in the case of the smooth hypotheses is the topic of Corrolary 1.2 and is derived by means of the sequential approximation result of Lemma 2.9.

The probability densities will be subjected to the following regularity conditions.

(C 1) Ξ is an open subset of R^m, for each $x \in X$ there exist partial derivatives

$$\frac{\partial^2 f(x,\gamma)}{\partial\gamma_i\partial\gamma_j}, \qquad i,j = 1,\ldots,m$$

and they are continuous on Ξ.

(C 2) The equalities

$$\int \frac{\partial^2 f(x,\gamma)}{\partial\gamma_i\partial\gamma_j}\,d\nu(x) \;=\; 0$$

hold for all $\gamma \in \Xi$ and $i,j = 1,\ldots,m$.

(C 3) The function $f(.,.)$ is positive on $X \times \Xi$ and for each parameter $\gamma \in \Xi$ there exist a $\overline{P}_\gamma$ integrable function h_γ and a neighbourhood $U_\gamma \subset \Xi$ of the point γ such that the inequality

$$\left|\frac{\partial^2 \log f(x,\gamma^*)}{\partial\gamma_i^*\partial\gamma_j^*}\right| \;\le\; h_\gamma(x)$$

holds for all $\gamma^* \in U_\gamma$, $x \in X$ and $i,j = 1,\ldots m$.

(C 4) For every $\gamma \in \Xi$ the function

$$\frac{\partial\,\log f(x,\gamma)}{\partial\gamma} = \left(\frac{\partial\,\log f(x,\gamma)}{\partial\gamma_1},\ldots,\frac{\partial\,\log f(x,\gamma)}{\partial\gamma_m}\right)^T$$

belongs to $\mathcal{L}_2(\overline{P}_\gamma)$ and the matrix

$$\mathbf{J}(\gamma) \;=\; \left(E_\gamma(\frac{\partial\,\log f(x,\gamma)}{\partial\gamma_i}\,\frac{\partial\,\log f(x,\gamma)}{\partial\gamma_j})\right)_{i,j=1}^m \tag{1.10}$$

is positive definite and continuous on Ξ.

(C 5) There exist measurable mappings $\hat{\gamma}_n : X^n \to \Xi$ *such that for each parameter* $\gamma \in \Xi$ *and every real number* $\varepsilon > 0$

$$\lim_{n\to\infty} \overline{P}_{\gamma}^{(n)}[\ L(x_1,\ldots,x_n,\Xi) = L(x_1,\ldots,x_n,\hat{\gamma}_n(x_1,\ldots,x_n))\ ] = 1\,, \tag{1.11}$$

$$\lim_{n\to\infty} \overline{P}_{\gamma}^{(n)}[\,\|\hat{\gamma}_n(x_1,\ldots,x_n) - \gamma\| \ \geq \varepsilon\,] = 0\,.$$

We remark that the symbol E_γ in (1.10) relates to the measure $\overline{P}_\gamma$.

Theorem 1.1 *Suppose that (C 1) - (C 5) and (1.6) hold, and*

$$\Omega_i = \{\,\theta \in \Theta;\ \mathbf{A_i}\theta = \mathbf{b_i}\,\}\,,$$

where $\mathbf{A_i}$ *is a* $k_i \times mq$ *matrix, rank*$(\mathbf{A_i}) = k_i$ *and* $\mathbf{b_i}$ *is a vector from* R^{k_i}. *Let* $\theta \in \Omega_0$ *and for* $i = 0,1$ *there exist measurable mappings* $\tilde{\theta}^{(i)}_{n_1,\ldots,n_q} = \tilde{\theta}^{(i)}_{n_1,\ldots,n_q}(x_{(n_1,\ldots,n_q)})$ *of the argument* $x_{(n_1,\ldots,n_q)}$ *taking values in* Ω_i *such that*

$$P_\theta^{(n_1,\ldots,n_q)}[\ L(x_{(n_1,\ldots,n_q)},\Omega_i) = L(x_{(n_1,\ldots,n_q)},\tilde{\theta}^{(i)}_{n_1,\ldots,n_q}(x_{(n_1,\ldots,n_q)})\,)\ ] \longrightarrow 1 \tag{1.12}$$

and for every $\varepsilon > 0$

$$P_\theta^{(n_1,\ldots,n_q)}[\,\|\tilde{\theta}^{(i)}_{n_1,\ldots,n_q} - \theta\| > \varepsilon\,] \longrightarrow 0\,. \tag{1.13}$$

(I) The weak convergence (1.8) of distributions holds with $s = k_0$.

(II) If $\Omega_0 \subset \Omega_1$ *and* $k_0 > k_1$, *then the weak convergence (1.9) of distributions holds with* $s = k_0 - k_1$.

An immediate application of the previous theorem yields the following assertion.

Corollary 1.1 *Let the homogeneity hypotheses*

$$\Omega_0 = \{\,\theta \in \Theta;\ \mu_1 = \ldots = \mu_q\,\}\,, \qquad \Omega_1 = \{\,\theta \in \Theta;\ \mu_1^{(1)} = \ldots = \mu_q^{(1)}\,\}\,,$$

where for the overall parameter $\theta = (\theta_1^T,\ldots,\theta_q^T)^T \in \Theta$ *either for all* $j = 1,\ldots,q$ *the equality* $\mu_j = \theta_j$ *holds, or* $\theta_j = (\mu_j^T,\sigma_j^T)^T$ *denotes partition of the* j*th population parameter into the subvectors* $\mu_j \in R^p$ *and* $\sigma_j \in R^{m-p}$ *(thus in the first case* $dim(\mu_j) = m$ *and in the second case* $dim(\mu_j) = p$*). Let* $\mu_j = (\mu_j^{(1)\,T},\mu_j^{(2)\,T})^T$ *denotes partition of the subvector* μ_j *into the subsubvectors* $\mu_j^{(1)} \in R^{p_1}$ *and* $\mu_j^{(2)} \in R^{dim(\mu_j)-p_1}$.

Suppose that (C 1) - (C 5) and (1.6) hold and the assumptions of the previous theorem concerning $\tilde{\theta}^{(i)}_{n_1,\dots,n_q}$ *are fulfilled.*

(I) The convergence (1.8) holds with $s = (q-1)dim(\mu_j)$*.*

(II) The convergence (1.9) holds with $s = (q-1)(dim(\mu_j) - p_1)$*.*

We remark that if in *(C 4)* the assumption of continuity of $\mathbf{J}(\gamma)$ on Ξ is omitted and in *(C 2)* also the validity of

$$\int \frac{\partial f(x,\gamma)}{\partial \gamma_i} \, \mathrm{d}\nu(x) \;=\; 0\,, \qquad i = 1, \dots, m \tag{1.14}$$

is assumed, then all assertions of this paper remain true. However, the present form of the conditions makes possible to use the local asymptotic normality theory of Le Cam, Ibragimov and Hasminskii in the form expounded in [2], which simplifies the proofs of contiguity assertions. We remark that in comparison with [6], the conditions *(C 2)*, *(C 3)* are less stringent than their counterparts *(R 2)*, *(R 3)* ibidem, and no assumption on the Kullback-Leibler information quantity is here included. In difference from various sets of classical regularity conditions, used for example in Section 4.4.2 of [12], in Section 6e of [9] or on pp. 88 of [1], the present conditions *(C 1) – (C 5)* do not require existence of the third partial derivatives of the densities. Also, they do not include the integrability of a higher power of partial derivatives of logarithm of the density, imposed in the condition C in [5].

A deeper insight into limiting behaviour of the test statistic can provide its asymptotic distribution when the true parameter tends to the null hypothesis, usually the rate proportional to the square root of the sample size is considered. Such an approach is for the likelihood ratio test statistics in the one sample case used in [6] for the hypotheses approximable by disjoint cones, and in [5] for the hypothesis of nulity of a part of the parameter. Multipopulation versions of these results are presented in Theorem 1.2 and in Corollary 1.2, the local alternatives are in their proofs handled by means of contiguity properties.

Following [6] and [4] we shall say that a set $\Omega \subset \Theta$ is at $\theta \in \Omega$ *sequentially approximable by the cone* C, if for every sequence $\{a_n\}_{n=1}^{\infty}$ of positive numbers which converges to zero

$$\sup \{\, \rho(\theta^*, \theta + C);\ \theta^* \in \Omega\,,\ \|\theta^* - \theta\| \le a_n \,\} = o(a_n)\,,$$

$$\sup \{\, \rho(\theta + y, \Omega);\ y \in C\,,\ \|y\| \le a_n \,\} = o(a_n)\,. \tag{1.15}$$

Here

$$\|x\| = (\sum_i x_i^2)^{1/2}\,, \quad \rho(z, S) = \inf \{\, \|z - y\|;\, y \in S \,\} \tag{1.16}$$

is the Euclidean distance from the set S, and by the cone C we mean any closed convex set such that $\alpha x \in C$ whenever $x \in C$ and the real number α is nonnegative.

For $\theta = (\theta_1^T, \ldots, \theta_q^T)^T \in \Theta$ the block diagonal $mq \times mq$ matrix

$$\mathbf{J}(\theta) = \mathrm{diag}\Big(\mathbf{J}(\theta_1), \ldots, \mathbf{J}(\theta_q) \Big) \tag{1.17}$$

denotes the overall Fisher information matrix whose blocks are defined by (1.10),

$$\pi_j(\theta) = \pi_j(\, (\theta_1^T \ldots, \theta_q^T)^T \,) = \theta_j \tag{1.18}$$

is projection onto the jth coordinate space and $h = (\pi_1(h)^T, \ldots, \pi_q(h)^T)^T$ describes decomposition of the vector $h \in R^{mq}$ into the subvectors from R^m. Finally, the number

$$n = n_1 + \ldots + n_q$$

denotes the total sample size, i.e., $n = n^{(u)}$ where u is the order number of the experiment.

Theorem 1.2 *Let (C 1) - (C 5) be fulfilled, (1.6) hold and*

$$\frac{n_j}{n} \longrightarrow p_j \in (0, 1\rangle\,, \qquad j = 1, \ldots, q\,. \tag{1.19}$$

Suppose that $\theta \in \Theta$ is a fixed parameter, for $i = 0, 1$ the set $\Omega_i \subset \Theta$ contains θ, is at θ sequentially approximable by a cone C_i and there exist measurable mappings $\tilde{\theta}^{(i)}_{n_1,\ldots,n_q} = \tilde{\theta}^{(i)}_{n_1,\ldots,n_q}(x_{(n_1,\ldots,n_q)})$ of the argument $x_{(n_1,\ldots,n_q)}$ taking values in Ω_i such that both (1.12) holds and (1.13) is true for every $\varepsilon > 0$. If

$$\lim_{u \to \infty} h_u = h \in R^{mq} \tag{1.20}$$

and the product measure corresponding to the u-th experiment

$$P^* = P_u^* = \overline{P}^{(n_1^{(u)})}_{\gamma(1,u)} \times \ldots \times \overline{P}^{(n_q^{(u)})}_{\gamma(q,u)}\,, \qquad \gamma(j, u) = \pi_j(\theta) + \frac{\pi_j(h_u)}{\sqrt{n_j^{(u)}}}\,, \tag{1.21}$$

then

$$\mathcal{L}\left[2 \log \frac{L(x_{(n_1,\ldots,n_q)}, \Omega_1)}{L(x_{(n_1,\ldots,n_q)}, \Omega_0)} \,\Big|\, P^* \right] \longrightarrow \mathcal{L}\left[\rho^2(x, G_0) - \rho^2(x, G_1) \,\Big|\, N\Big(\mathbf{J}(\theta)^{1/2} h\,,\, \mathbf{I_{mq}}\Big) \right]. \tag{1.22}$$

Here ρ is the distance (1.16) from the set $G_i = \mathbf{J}(\theta)^{1/2}\mathbf{D}(\mathbf{p})^{1/2} C_i$ and

$$\mathbf{D}(\mathbf{p})^{1/2} = diag(\, p_1^{1/2}, \ldots, p_1^{1/2}, p_2^{1/2}, \ldots, p_2^{1/2},\ \ldots\ldots\ , p_q^{1/2}, \ldots, p_q^{1/2}\,) \tag{1.23}$$

denotes the diagonal $mq \times mq$ matrix with this diagonal. Especially, if $C_0 \subset C_1$ are linear spaces, then

$$\mathcal{L}\left[2 \log \frac{L(x_{(n_1,\ldots,n_q)}, \Omega_1)}{L(x_{(n_1,\ldots,n_q)}, \Omega_0)} \,\Big|\, P^* \right] \longrightarrow \chi_s^2(\lambda)\,, \tag{1.24}$$

where $s = dim(C_1) - dim(C_0)$ *and the noncentrality parameter of the chi-square distribution*

$$\lambda = \rho^2(\mathbf{J}(\theta)^{1/2}h, G_0) - \rho^2(\mathbf{J}(\theta)^{1/2}h, G_1)\,. \tag{1.25}$$

The previous theorem implies the following assertion, in which $\mathcal{C}_1$ denotes the class of mappings whose components have on their domain all partial derivatives of the first order continuous. In (1.27) the symbol θ_t stands for the t-th coordinate of the vector $\theta \in R^{mq}$.

Corollary 1.2 *Suppose that (C 1) - (C 5), (1.6) and* (1.19) *hold,*

$$\Omega_i = \{\, \theta^* \in \Theta;\; g_1(\theta^*) = 0, \ldots, g_{k_i}(\theta^*) = 0 \,\}\,, \tag{1.26}$$

and the functions $g_j : \Theta \to R^1$ *belong to* $\mathcal{C}_1$. *Further, assume that for* $i = 0,\, 1$ *the parameter* θ *belongs to* Ω_i, *the matrix*

$$\boldsymbol{\partial_i}(\theta) = \begin{pmatrix} \dfrac{\partial g_1(\theta)}{\partial \theta_1}, & \cdots & , \dfrac{\partial g_1(\theta)}{\partial \theta_{mq}} \\ \vdots & & \vdots \\ \dfrac{\partial g_{k_i}(\theta)}{\partial \theta_1}, & \cdots & , \dfrac{\partial g_{k_i}(\theta)}{\partial \theta_{mq}} \end{pmatrix} \tag{1.27}$$

is of rank k_i *and the assumptions of the previous theorem concerning* $\tilde{\theta}^{(i)}_{n_1,\ldots,n_q}$ *are fulfilled. Let (1.20) hold and* P^* *be the probability defined in (1.21).*

(I) The convergence

$$\boldsymbol{\mathcal{L}}\left[2\log \frac{L(x_{(n_1,\ldots,n_q)},\Theta)}{L(x_{(n_1,\ldots,n_q)},\Omega_0)} \,\Big|\, P^* \right] \longrightarrow \chi^2_{k_0}(\lambda)$$

holds with

$$\lambda = h^T \mathbf{F}_0^T (\mathbf{F}_0 \mathbf{J}(\theta)^{-1} \mathbf{F}_0^T)^{-1} \mathbf{F}_0 h\,, \qquad \mathbf{F}_0 = \boldsymbol{\partial_0}(\theta)\mathbf{D}(\mathbf{p})^{-1/2}\,. \tag{1.28}$$

(II) If $k_1 < k_0$ *(and therefore* $\Omega_0 \subset \Omega_1$*), then (1.24) holds with* $s = k_0 - k_1$ *and*

$$\lambda = h^T \left[\mathbf{F}_0^T (\mathbf{F}_0 \mathbf{J}(\theta)^{-1} \mathbf{F}_0^T)^{-1} \mathbf{F}_0 - \mathbf{F}_1^T (\mathbf{F}_1 \mathbf{J}(\theta)^{-1} \mathbf{F}_1^T)^{-1} \mathbf{F}_1 \right] h\,, \quad \mathbf{F}_1 = \boldsymbol{\partial_1}(\theta)\mathbf{D}(\mathbf{p})^{-1/2}.$$

(III) If for the relative sample sizes the inequality $\liminf_{u\to\infty} n_j^{(u)}/n^{(u)} > 0$ *holds for* $j = 1, \ldots, q$ *and if the vector* $h = 0$, *then the results on limiting distributions in the assertions (I) and (II) are valid with* $\lambda = 0$ *provided that their assumptions with the exception of (1.19) remain true.*

2. Proofs.

Lemma 2.1 *Let (C 1) - (C 4) be fulfilled. Then (1.14) holds and for* i , $j = 1, \dots, m$

$$\mathbf{J}(\gamma)_{ij} = -\, E_\gamma \left(\frac{\partial^2 \log L(x, \gamma)}{\partial \gamma_i \partial \gamma_j} \right) . \tag{2.1}$$

Proof. (I) Validity of (1.14) follows from Proposition 1 and the relation (8) on pp. 13 – 16 in [2], validity of (2.1) immediately follows from (1.14) and *(C 2)*. □

Since sampling with the sample sizes $n_j = n_j^{(u)}$ is carried out in the sequence of experiments whose ordering is denoted by $u = 1, 2, \dots$, for the sake of simplicity the pooled sample will be denoted by the symbol (cf. (1.4), (1.3))

$$x^{(u)} = x_{(n_1^{(u)}, \dots, n_q^{(u)})} . \tag{2.2}$$

In accordance with this and (1.5) let

$$P_\theta^{(u)} = P_\theta^{(n_1^{(u)}, \dots, n_q^{(u)})} . \tag{2.3}$$

A basic tool for finding stochastic order of the remainder term in the concerned Taylor expansion will be in this paper the next assertion.

Lemma 2.2 *Suppose that (C 1) - (C 4) hold and using the notation (1.1), (1.10) for* $\gamma \in \Xi$ *put*

$$\begin{gathered} d(x_1, \dots, x_n, \gamma, \delta) \\ = \sup \left\{ \left| \frac{1}{n} \frac{\partial^2 \log L(x_1, \dots, x_n, \gamma^*)}{\partial \gamma_i^* \partial \gamma_j^*} + \mathbf{J}(\gamma)_{ij} \right| ; \; \|\gamma^* - \gamma\| \leq \delta, \; i, j = 1, \dots, m \right\} . \end{gathered} \tag{2.4}$$

(I) The function $d(., \gamma, \delta)$ *is measurable and for every* $\varepsilon > 0$ *there exists a real number* $\delta > 0$ *such that*

$$\lim_{n \to \infty} \overline{P}_\gamma^{(n)} \left[d(x_1, \dots, x_n, \gamma, \delta) \geq \varepsilon \right] = 0.$$

(II) If (1.6) holds, $\theta \in \Theta$ *and measurable real-valued functions* $\psi_u = \psi_u(x^{(u)})$ *converge to zero in the probabilities (2.3), then (cf. (1.18))*

$$\lim_{u \to \infty} P_\theta^{(u)} \left[d(\, x(j, n_j^{(u)}), \theta_j, \psi_u(x^{(u)}) \,) \geq \varepsilon \right] = 0$$

for all $j = 1, \dots, q$ *and every* ε *positive.*

Proof.

(I) The measurability follows from continuity of the partial derivatives and separability of R^m. Let

$$g(x,\gamma,\delta)=\sup\left\{\left|\frac{\partial^2\log L(x,\gamma^*)}{\partial\gamma_i^*\partial\gamma_j^*}-\frac{\partial^2\log L(x,\gamma)}{\partial\gamma_i\partial\gamma_j}\right|;\ \|\gamma^*-\gamma\|\le\delta,\ i,j=1,\ldots,m\right\}.$$

By *(C 1)*, *(C 3)* and the Lebesgue theorem

$$\lim_{\delta\to 0^+}\int g(x,\gamma,\delta)\,\mathrm{d}\overline{P}_\gamma(x)=0\,,$$

and given $\varepsilon>0$ there is a positive number δ such that $\mathrm{E}_\gamma(\,g(.,\gamma,\delta)\,)<\frac{\varepsilon}{2}$. Employing the law of large numbers we obtain that for such a number δ

$$\overline{P}_\gamma^{(n)}\,[\,d(x_1,\ldots,x_n,\gamma,\delta)\ge\varepsilon\,]$$

$$\le\overline{P}_\gamma^{(n)}\left[\frac{1}{n}\sum_{i=1}^n g(x_i,\gamma,\delta)\ge\frac{\varepsilon}{2}\right]+\overline{P}_\gamma^{(n)}\left[d(x_1,\ldots,x_n,\gamma,0)\ge\frac{\varepsilon}{2}\right]\longrightarrow 0$$

as $n\to\infty$, because (2.1) holds.

(II) Let ε be a fixed positive number. According to (I) there exist positive real numbers δ_j and $N(j,t)$ such that

$$\overline{P}_{\theta_j}^{(n)}[\,d(x_1,\ldots,x_n,\theta_j,\delta_j)\ge\varepsilon\,]\ \le\ \frac{1}{t}$$

for all $n\ge N(j,t)$. Further, since (1.6) holds, given sequences $\{n_j^{(u)}\}_{u=1}^\infty$, $j=1,\ldots,q$, there exists an increasing sequence $\{u_t\}_{t=1}^\infty$ of positive integers such that for all $u\ge u_t$

$$n_j=n_j^{(u)}\ge N(j,t)\,,\qquad P_\theta^{(u)}[\,\psi_u(x^{(u)})\ge\delta_j\,]\le\frac{1}{t}\,.$$

Hence $P_\theta^{(u)}[\,d(x(j,n_j^{(u)}),\theta_j,\psi_u(x^{(u)}))\ge\varepsilon\,]\le 2/t$ for all $u\ge u_t$. □

In the next considerations we shall use the notation

$$\mathbf{B}((x_1,\ldots,x_n),\gamma)=\left(\frac{\partial^2\log L(x_1,\ldots,x_n,\gamma)}{\partial\gamma_i\partial\gamma_j}\right)_{i,j=1}^m\,. \tag{2.5}$$

The matrix

$$\mathbf{B}(x^{(u)},\theta)=\ \mathrm{diag}\Big(\mathbf{B}(x(1,n_1^{(u)}),\pi_1(\theta)),\ldots,\mathbf{B}(x(q,n_q^{(u)}),\pi_q(\theta))\Big) \tag{2.6}$$

is the $mq\times mq$ block diagonal matrix whose blocks are defined by means of (2.5) and (1.18). Finally, let

$$\mathbf{D_u}=\mathrm{diag}(\,\sqrt{n_1^{(u)}},\ldots,\sqrt{n_1^{(u)}},\sqrt{n_2^{(u)}},\ldots,\sqrt{n_2^{(u)}},\ \ldots\ldots\ ,\sqrt{n_q^{(u)}},\ldots,\sqrt{n_q^{(u)}}\,) \tag{2.7}$$

denote the diagonal $mq\times mq$ matrix with this diagonal.

The following lemma is a multisample version of Lemma 1 in [4].

Lemma 2.3 *Suppose that (C 1) - (C 4) and (1.6) hold, $\theta \in \Omega \subset \Theta$ and measurable mappings $\tilde\theta_u = \tilde\theta_u(x^{(u)})$ taking values in Ω are such that (cf. (1.7))*

$$\lim_{u\to\infty} P_\theta^{(u)}[\ L(x^{(u)},\Omega) = L(\ x^{(u)}, \tilde\theta_u(x^{(u)})\ )\ ] = 1\,.$$

Let $\tilde\theta_u \to \theta$ in probabilities (2.3) as $u \to \infty$. Then the random vectors

$$\mathbf{\Delta_u}(x^{(u)}) = \begin{pmatrix} \sqrt{n_1^{(u)}}\pi_1\Big(\tilde\theta_u(x^{(u)}) - \theta\Big) \\ \vdots \\ \sqrt{n_q^{(u)}}\pi_q\Big(\tilde\theta_u(x^{(u)}) - \theta\Big) \end{pmatrix} = \mathbf{D_u}(\tilde\theta_u(x^{(u)}) - \theta)$$

are bounded in probabilities $P = P_\theta^{(u)}$, i.e., $\mathbf{\Delta_u}(x^{(u)}) = \mathcal{O}_p(1)$.

Proof. Choose a number $\delta > 0$ such that $\{\,\theta^* \in R^{mq};\ \|\theta^* - \theta\| < \delta\,\} \subset \Theta$ and put

$$A_u = \{\, x^{(u)};\ L(x^{(u)},\Omega) = L(\,x^{(u)}, \tilde\theta_u(x^{(u)})\,)\,,\ \ \|\tilde\theta_u(x^{(u)}) - \theta\| < \delta\,\}\ .$$

An application of the Taylor theorem yields that for every $x^{(u)} \in A_u$

$$\begin{aligned} &\log L(x^{(u)}, \tilde\theta_u) \\ &= \log L(x^{(u)},\theta) + \frac{\partial \log L(x^{(u)},\theta)}{\partial\theta}^T (\tilde\theta_u - \theta) + \frac{1}{2}(\tilde\theta_u - \theta)^T \mathbf{B}(x^{(u)}, \theta_u^*)(\tilde\theta_u - \theta)\,, \end{aligned} \tag{2.8}$$

where $\|\theta_u^* - \theta\| \le \|\tilde\theta_u - \theta\|$ and (cf. (1.17))

$$\frac{1}{2}(\tilde\theta_u - \theta)^T \mathbf{B}(x^{(u)}, \theta_u^*)(\tilde\theta_u - \theta) \ = \ -\frac{1}{2}\mathbf{\Delta_u}(x^{(u)})^T \mathbf{J}(\theta)\mathbf{\Delta_u}(x^{(u)}) \ + z_u(x^{(u)})\,. \tag{2.9}$$

Making use of the Cauchy-Schwarz inequality and the notation (2.4) we get the inequality

$$\begin{aligned} |z_u(x^{(u)})| \ &\le \ \|\mathbf{\Delta_u}(x^{(u)})\|^2 S_u(x^{(u)})\,, \\ S_u(x^{(u)}) \ &= \ \sum_{j=1}^q m^2 d\Big(x(j, n_j^{(u)})\,,\ \pi_j(\theta)\,,\ \|\tilde\theta_u - \theta\|\Big)\ . \end{aligned} \tag{2.10}$$

However, Lemma 2.2(II) implies that $S_u(x^{(u)}) = o_P(1)$, which together with (2.8) – (2.10) means that on A_u

$$0 \ \le\ \log \frac{L(x^{(u)}, \tilde\theta_u)}{L(x^{(u)},\theta)}$$

$$\le\ \frac{\partial \log L(x^{(u)},\theta)}{\partial\theta}^T (\tilde\theta_u - \theta) - \frac{1}{2}\mathbf{\Delta_u}(x^{(u)})^T \mathbf{J}(\theta)\mathbf{\Delta_u}(x^{(u)}) + \|\mathbf{\Delta_u}(x^{(u)})\|^2 o_p(1)\,. \tag{2.11}$$

Finally, let λ stand for the smallest characteristic root of $\mathbf{J}(\theta)$. Then from (2.11) by means of *(C 4)* and the central limit theorem we obtain that on A_u

$$\|\mathbf{\Delta_u}(x^{(u)})\|^2 \left(\frac{\lambda}{2} - o_p(1)\right) \leq \left\| \mathbf{D_u}^{-1} \frac{\partial \log L(x^{(u)}, \theta)}{\partial \theta} \right\| \|\mathbf{\Delta_u}(x^{(u)})\| =$$

$$= \mathcal{O}_P(1) \|\mathbf{\Delta_u}(x^{(u)})\| ,$$

and since $P_\theta^{(u)}(A_u) \to 1$ as $u \to \infty$, the rest of the proof is obvious. □

The statements (2.12), (2.13) of the next corollary are well-known properties of the maximum likelihood estimators and have been proved under various sets of conditions. The set of conditions *(C 1) – (C 5)* differs in some way from currently used ones, amongst which one can mention the conditions used in [6], [8], in chapter 4 of [12], the conditions in Section 6e.1 in [9] or the ones used on p. 88 in [1]. For the sake of completeness we therefore prefer to include the proof of the following corollary into the text.

Corollary 2.1 *Suppose that the conditions (C 1) - (C 5) are fulfilled. Then for the maximum likelihood estimator $\hat{\gamma}_n$ from (C 5) and for every parameter $\gamma \in \Xi$*

$$\sqrt{n}(\, \hat{\gamma}_n(x_1, \dots, x_n) - \gamma) = \mathbf{J}(\gamma)^{-1} \frac{1}{\sqrt{n}} \frac{\partial \log \, L(x_1, \dots, x_n, \gamma)}{\partial \gamma} + o_P(1) \;, \tag{2.12}$$

where $P = \overline{P}_\gamma^{(n)}$. Hence the weak convergence to the normal distribution

$$\mathcal{L}[\, \sqrt{n}(\hat{\gamma}_n - \gamma) \,|\, \overline{P}_\gamma^{(n)} \,] \;\to\; N(\, \mathbf{0}, \mathbf{J}(\gamma)^{\mathbf{-1}} \,) \tag{2.13}$$

holds as $n \to \infty$.

Proof. Since $\hat{\gamma}_n \to \gamma$ in probability, the Taylor theorem and (1.11) imply that with probability $P = \overline{P}_\gamma^{(n)}$ tending to 1

$$\frac{1}{n} \frac{\partial \log L(x_1, \dots, x_n, \gamma)}{\partial \gamma(i)} = \frac{1}{n} \sum_{j=1}^{m} \frac{\partial^2 \log L(x_1, \dots, x_n, \gamma_i^*)}{\partial \gamma_i^*(j) \partial \gamma_i^*(i)} \left(\gamma(j) - \hat{\gamma}_n(j)\right),$$

where $a(j)$ denotes the jth coordinate of a and $\|\, \gamma_i^* - \gamma \,\| \;\leq\; \|\, \hat{\gamma}_n - \gamma \,\|$. Thus

$$\frac{1}{n} \frac{\partial \log L(x_1, \dots, x_n, \gamma)}{\partial \gamma} = \mathbf{J}(\gamma)(\, \hat{\gamma}_n - \gamma \,) + z_n(x_1, \dots, x_n) \,, \tag{2.14}$$

$$\|z_n(x_1, \dots, x_n)\| \leq S_n(x_1, \dots, x_n) \|\hat{\gamma}_n - \gamma \,\| \,,$$

$$S_n(x_1, \dots, x_n) = m^2 d(x_1, \dots, x_n, \gamma, \|\hat{\gamma}_n - \gamma\|) \,,$$

and d is defined in (2.4). But according to Lemma 2.2 and Lemma 2.3

$$d(x_1,\ldots,x_n,\gamma,\|\hat{\gamma}_n-\gamma\|)=o_P(1)\,,\qquad \sqrt{n}(\,\hat{\gamma}_n(x_1,\ldots,x_n)-\gamma)=\mathcal{O}_P(1)\,,$$

and we see that $z_n(x_1,\ldots,x_n)=o_P(n^{-\frac{1}{2}})$. The rest of the proof follows from (2.14) and the central limit theorem. □

Lemma 2.4 *Suppose that the assumptions of Lemma 2.3 are fulfilled, (C 5) holds, and put*

$$v(z,K)\;=\;\inf\{\,|||z-y|||;\,y\in K\,\}\;,\qquad |||x|||=\sqrt{x\,'\mathbf{J}(\theta)x}\,,\tag{2.15}$$

where $\mathbf{J}(\theta)$ is the matrix (1.17). Let

$$\hat{\theta}_{(u)}=(\hat{\theta}^{\,T}_{n_1^{(u)}},\ldots,\hat{\theta}^{\,T}_{n_q^{(u)}})^T\,,\tag{2.16}$$

where $\hat{\theta}_{n_j^{(u)}}=\hat{\theta}_{n_j^{(u)}}(x(j,n_j^{(u)}))$ is the MLE from (C 5).

(I) Let G_u denote the set of those $x^{(u)}$ for which

$$L(x^{(u)},\Omega)\;=\;L(\,x^{(u)},\tilde{\theta}_u(x^{(u)})\,)\;,\qquad \|\mathbf{D_u}(\,\tilde{\theta}_u(x^{(u)})-\theta\,)\|\;\le\;\tilde{M}\;,\tag{2.17}$$

$$L(x^{(u)},\Theta)\;=\;L(\,x^{(u)},\hat{\theta}_{(u)}(x^{(u)})\,)\;,\qquad \|\mathbf{D_u}(\,\hat{\theta}_{(u)}(x^{(u)})-\theta\,)\|\;\le\;M\;,\tag{2.18}$$

where $\mathbf{D_u}$ is defined in (2.7) and $\tilde{M}$, M are fixed positive constants. Then the relation

$$\left|\log L(x^{(u)},\Omega)-\left[\log L(x^{(u)},\hat{\theta}_{(u)})-\frac{1}{2}\,v^2\Big(\,\mathbf{D_u}\hat{\theta}_{(u)},\mathbf{D_u}\Omega^{(u)}(\tilde{M})\,\Big)\right]\right|I_{G_u}(x^{(u)})=o_P(1)\tag{2.19}$$

holds with $P=P_\theta^{(u)}$, $\Omega^{(u)}(\tilde{M})\;=\;\{\,\theta^*\in\Omega;\,\|\mathbf{D_u}(\theta^*-\theta)\|\le\tilde{M}\,\}$ and I_{G_u} denoting the indicator function of this set.

(II) Suppose further that $\Omega\;=\;\{\,\theta\in\Theta;\,\mathbf{A}\theta=\mathbf{b}\,\}$, $\mathbf{A}$ is a $k\times mq$ matrix of rank k, $\mathbf{b}$ is a vector from R^k and $C=\{\,z\in R^{mq};\,\mathbf{A}z=0\,\}$. Then

$$\log L(x^{(u)},\Omega)-\left[\log L(x^{(u)},\hat{\theta}_{(u)})-\frac{1}{2}\,v^2(\mathbf{D_u}(\hat{\theta}_{(u)}-\theta),\mathbf{D_u}C)\right]=o_P(1)\;.\tag{2.20}$$

Proof.

(I) Since the set $\Theta=\Xi^q$ is open, there is a neighbourhood U of the parameter θ such that $U\subset\Theta$. Obviously, $\Omega^{(u)}(\tilde{M})\subset U$ and $\hat{\theta}_{(u)}(x^{(u)})\in U$ for all $x^{(u)}\in G_u$ and all u sufficiently large. Hence for such an integer u and θ^* belonging to $\Omega^{(u)}(\tilde{M})$ the Taylor theorem yields that on G_u

$$\log L(x^{(u)},\theta^*)=\log L(x^{(u)},\hat{\theta}_{(u)})-\frac{1}{2}\left[\mathbf{D_u}(\theta^*-\hat{\theta}_{(u)})\right]^T\mathbf{J}(\theta)\left[\mathbf{D_u}(\theta^*-\hat{\theta}_{(u)})\right]+z_u\,.\tag{2.21}$$

Here

$$z_u = z_u(x^{(u)}) = \frac{1}{2}\left[\mathbf{D_u}(\theta^* - \hat{\theta}_{(u)})\right]^T \left[\mathbf{D_u^{-1}}\mathbf{B}(x^{(u)}, \theta^{**})\mathbf{D_u^{-1}} + \mathbf{J}(\theta)\right]\left[\mathbf{D_u}(\theta^* - \hat{\theta}_{(u)})\right], \tag{2.22}$$

$\|\theta^{**} - \hat{\theta}_{(u)}\| \le \|\theta^* - \hat{\theta}_{(u)}\|$, $\mathbf{B}$ is the matrix (2.6) and

$$\|\mathbf{D_u}(\theta^* - \hat{\theta}_{(u)})\| \le \|\mathbf{D_u}(\theta^* - \theta)\| + \|\mathbf{D_u}(\theta - \hat{\theta}_{(u)})\| \le \tilde{M} + M\,, \tag{2.23}$$

$$\|\theta^{**} - \theta\| \le \|\theta^* - \hat{\theta}_{(u)}\| + \|\hat{\theta}_{(u)} - \theta\| \le \alpha_u = \|\mathbf{D_u^{-1}}\|(\tilde{M} + 2M)\,. \tag{2.24}$$

From (2.22) – (2.24) and Lemma 2.2 one finds out that

$$|z_u| \le (\tilde{M} + M)^2 \sum_{j=1}^{q} m^2 d(x(j, n_j^{(u)}), \theta_j, \alpha_u) = o_P(1)\,,$$

which together with (2.21) and (2.17) implies (2.19).

(II) Obviously, for all u sufficiently large $\Omega^{(u)}(\tilde{M}) = \theta + \{\, z \in C;\ \|\mathbf{D_u} z\| \le \tilde{M}\,\}$ and

$$v^2\Big(\mathbf{D_u}\hat{\theta}_{(u)}, \mathbf{D_u}\Omega^{(u)}(\tilde{M})\Big) = v^2\Big(\mathbf{D_u}(\hat{\theta}_{(u)} - \theta), K_u\Big)\,, \tag{2.25}$$

where $K_u = \{\, z \in R^{mq};\ \mathbf{A}\mathbf{D_u^{-1}} z = \mathbf{0},\ \|z\| \le \tilde{M}\,\}$. Let $\mathbf{\Pi}_u$ be the matrix of projection on the linear subspace $\mathbf{D_u} C = \{\, z \in R^{mq};\ \mathbf{A}\mathbf{D_u^{-1}} z = \mathbf{0}\,\}$ in the norm $|||z|||$ from (2.15). Then $|||\mathbf{\Pi}_u y||| \le |||y|||$ and with λ_1 denoting the greatest and λ_{mq} the smallest characteristic root of $\mathbf{J}(\theta)$ the inequalities

$$\sqrt{\lambda_{mq}}\,\|\mathbf{\Pi}_u \mathbf{D_u}(\hat{\theta}_{(u)} - \theta)\| \le |||\mathbf{\Pi}_u \mathbf{D_u}(\hat{\theta}_{(u)} - \theta)||| \le \sqrt{\lambda_1}\,\|\mathbf{D_u}(\hat{\theta}_{(u)} - \theta)\|$$

hold. Hence inserting into (2.17) the constant $\tilde{M} = \sqrt{\lambda_1/\lambda_{mq}}\,M$, we see that on G_u

$$v^2(\mathbf{D_u}(\hat{\theta}_{(u)} - \theta), K_u) = v^2(\mathbf{D_u}(\hat{\theta}_{(u)} - \theta), \mathbf{D_u} C)\,. \tag{2.26}$$

With this choice of $\tilde{M}$ validity of (2.19), (2.25) and (2.26) yields the relation

$$g_u(x^{(u)}) I_{G_u}(x^{(u)}) = o_P(1)\,, \tag{2.27}$$

where $g_u(x^{(u)})$ denotes the left-hand side of (2.20).

Finally, let $\varepsilon > 0$ be an arbitrary but fixed real number. If $\delta > 0$, then (2.27), the assumptions on $\tilde{\theta}_u$, Lemma 2.3 and *(C 5)* imply that for $M > 0$ sufficiently large

$$\limsup_{u\to\infty} P_\theta^{(u)}(\,|g_u(x^{(u)})| \ge \varepsilon\,) \le \limsup_{u\to\infty}\Big[1 - P_\theta^{(u)}(\,G_u\,)\Big] < \delta$$

and the relation (2.20) is proved. □

Lemma 2.5 *Let the distributions* $\{\, \mathcal{L}(\xi_u)\,\}_{u=1}^{\infty}$ *of p-dimensional random vectors converge weakly to the normal distribution* $N(\mathbf{0}, \mathbf{I_p})$*, where* $\mathbf{I_p}$ *is the unit matrix. If* $\{\, \mathbf{W_u}\,\}_{u=1}^{\infty}$ *are idempotent symmetric* $p \times p$ *matrices and* $tr(\mathbf{W_u}) = s$ *for all* u*, then*

$$\mathcal{L}(\xi_u^T\, \mathbf{W_u}\, \xi_u) \;\longrightarrow\; \chi_s^2 \tag{2.28}$$

weakly as $u \to \infty$.

Proof. Assume first that for $i, j = 1, \dots, p$

$$\lim_{u\to\infty} \mathbf{W_u}(i,j) = \mathbf{W}(i,j)\,, \tag{2.29}$$

where $\mathbf{W}$ is a real-valued $p \times p$ matrix. Then the functions $h_u(x) = x^T \mathbf{W_u} x$, $h(x) = x^T \mathbf{W} x$ are measurable and since $x_u \to x$ obviously implies that $h_u(x_u) \to h(x)$, according to Theorem 5.5 in [3]

$$\mathcal{L}(\xi_u^T\, \mathbf{W_u}\, \xi_u) = \mathcal{L}(\, h_u(\xi_u)\,) \longrightarrow \mathcal{L}(\, h(x)\,|\, N(\mathbf{0}, \mathbf{I_p})\,)\ .$$

Taking into account (2.29) we see that $\mathrm{tr}(\mathbf{W}) = \lim_{u\to\infty} \mathrm{tr}(\mathbf{W_u}) = s$ and the matrix $\mathbf{W}$ is symmetric and idempotent. This according to Lemma 9.1.2 on p. 169 in [10] means that $\mathcal{L}(\, x^T \mathbf{W} x \,|\, N(\mathbf{0}, \mathbf{I_p})\,) = \chi_s^2$ and (2.28) in this case holds.

Let us drop validity of the assumption (2.29). Since the matrices $\{\, \mathbf{W_u}\,\}_{u=1}^{\infty}$ are symmetric and idempotent, they are positive semidefinite and for all $i, j = 1, \dots, p$

$$0 \le \mathbf{W_u}(i,i) \;\le \mathrm{tr}(\mathbf{W_u}) = s\,, \quad |\mathbf{W_u}(i,j)| \;\le \sqrt{\mathbf{W_u}(i,i)\mathbf{W_u}(j,j)} \le s\ .$$

Hence every increasing sequence $\{u_v\}_{v=1}^{\infty}$ of positive integers contains a subsequence $\{u_{v_t}\}_{t=1}^{\infty}$ such that the matrices $\{\mathbf{W_{u_{v_t}}}\}_{t=1}^{\infty}$ converge to a real valued $p \times p$ matrix $\mathbf{W}$, and according to the previous part of the proof $\mathcal{L}(\xi_{u_{v_t}}^T\, \mathbf{W_{u_{v_t}}}\, \xi_{u_{v_t}}) \to \chi_s^2$ as $t \to \infty$, which proves (2.28). □

P r o o f o f T h e o r e m 1. 1. Making use of *(C 5)* we obtain that (cf. (2.16))

$$\log L(x^{(u)}, \Theta) \;=\; \log L(x^{(u)}, \hat{\theta}_{(u)}) + o_P(1)\,,$$

where $P = P_{\theta}^{(u)}$. This together with (2.20) implies that

$$2\,\log \frac{L(x^{(u)}, \Theta)}{L(x^{(u)}, \Omega_i)} = v^2(\, \mathbf{D_u}(\hat{\theta}_{(u)} - \theta)\,,\, \mathbf{D_u} C_i\,) + o_P(1) = \rho^2(\xi_u, \mathbf{J}(\theta)^{1/2}\mathbf{D_u} C_i) + o_P(1)\,,$$

where $C_i = \{\, z \in R^{mq};\, \mathbf{A_i} z = \mathbf{0}\,\}$, ρ is the distance (1.16) and $\xi_u = \mathbf{J}(\theta)^{\frac{1}{2}} \mathbf{D_u}(\hat{\theta}_{(u)} - \theta)$. Owing to (1.6) and (2.13)

$$\mathcal{L}(\, \xi_u\,) \;\to\; N(\mathbf{0}, \mathbf{I_{mq}}) \tag{2.30}$$

as $u \to \infty$. Since according to the assertion *(i)* on p. 23 of [9] the matrix $\mathbf{\Psi_u^{(i)}}$ of projection on $\mathbf{J}(\theta)^{1/2}\mathbf{D_u}C_i$ is symmetric and idempotent,

$$2 \log \frac{L(x^{(u)}, \Theta)}{L(x^{(u)}, \Omega_i)} = \xi_u^T \, \mathbf{W_u^{(i)}} \xi_u + o_P(1) \,, \tag{2.31}$$

$$\mathbf{W_u^{(i)}} = \mathbf{I_{mq}} - \mathbf{\Psi_u^{(i)}} \,, \quad \mathrm{tr}(\mathbf{\Psi_u^{(i)}}) = \dim(\mathbf{J}(\theta)^{1/2}\mathbf{D_u}C_i) = mq - k_i \,, \tag{2.32}$$

and the matrix $\mathbf{W_u^{(i)}}$ is symmetric and idempotent.

(I) This assertion follows from (2.30) -(2.32) and Lemma 2.5.

(II) By (2.31) and (2.32)

$$2 \log \frac{L(x^{(u)}, \Omega_1)}{L(x^{(u)}, \Omega_0)} = \xi_u{}' \, \mathbf{W_u} \xi_u + o_P(1) \,, \tag{2.33}$$

where $\mathbf{W_u} = \mathbf{\Psi_u^{(1)}} - \mathbf{\Psi_u^{(0)}}$. But $\Omega_0 \subset \Omega_1$ implies that $C_0 \subset C_1$, which together with symmetry of the projection matrices leads to the equalities $\mathbf{\Psi_u^{(1)}} \mathbf{\Psi_u^{(0)}} = \mathbf{\Psi_u^{(0)}} = \mathbf{\Psi_u^{(0)}} \mathbf{\Psi_u^{(1)}}$. Thus the matrix $\mathbf{W_u}$ is symmetric and idempotent, and the rest of the proof follows from (2.30), (2.32), (2.33) and Lemma 2.5. □

In the following text we shall use the concept of contiguity. We recall that a sequence $\{P_u\}_{u=1}^{\infty}$ of probabilities is said to be *contiguous* to propabilities $\{P_u^*\}_{u=1}^{\infty}$, if $\lim_{u\to\infty} P_u(A_u) = 0$ whenever $\lim_{u\to\infty} P_u^*(A_u) = 0$. This is denoted by $\{P_u\} \lhd \{P_u^*\}$ and these sequences of probabilities are said to be contiguous, if both $\{P_u\} \lhd \{P_u^*\}$ and $\{P_u^*\} \lhd \{P_u\}$.

Lemma 2.6 *Suppose that (C 1) - (C 4) and (1.6) hold,* $\theta \in \Theta$, $\lim_{u\to\infty} h_u = h \in R^{mq}$, *and in accordance with (1.21), (2.3) put* $P_u = P_\theta^{(u)}$.

(I) The probabilities $\{P_u\}_{u=1}^{\infty}$, $\{P_u^*\}_{u=1}^{\infty}$ *are contiguous.*

(II) If also (C 5) holds, then for the maximum likelihood estimate (2.16) and the matrices (2.7), (1.17)

$$\mathcal{L}[\, \mathbf{D_u}(\hat{\theta}_{(u)} - \theta) \,|\, P_u^* \,] \;\to\; N(\, h, \mathbf{J}(\theta)^{-\mathbf{1}} \,) \tag{2.34}$$

as $u \to \infty$.

Proof. (I) The proof coincides with its one-sample counterpart used for proving Proposition 3 on p. 17 in [2]. Indeed, let $\theta_{(u)} = \theta + \mathbf{D_u}^{-1} h_u$ and

$$\Lambda_u^* = \log \, \frac{L(x^{(u)}, \theta)}{L(x^{(u)}, \theta_{(u)})} \,. \tag{2.35}$$

Since by the uniform weak convergence of probabilities one understands that integrals of every bounded continuous function converge uniformly, from Proposition 1 on p. 13 in [2] and from (12) – (14) on p. 16 ibidem one easily finds out that

$$\mathcal{L}(\Lambda_u^* \,|\, P_u^* \,) \;\to\; N(-\frac{\sigma^2}{2} \,, \sigma^2 \,) \,, \qquad \sigma^2 = h^T \mathbf{J}(\theta) h \,. \tag{2.36}$$

This together with Le Cam's first lemma (cf. [2] p. 499) means that $\{P_u\} \lhd \{P_u^*\}$, the relation $\{P_u^*\} \lhd \{P_u\}$ can be proved similarly.

(II) Let

$$S_u(\theta) = \left(S_{n_1^{(u)}}(x(1, n_1^{(u)}), \pi_1(\theta))^T, \ldots, S_{n_q^{(u)}}(x(q, n_q^{(u)}), \pi_q(\theta))^T \right)^T,$$

$$S_n(x_1, \ldots, x_n, \gamma) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{\partial \log f(x_i, \gamma)}{\partial \gamma}$$

and $\Lambda_u = -\Lambda_u^*$, where Λ_u^* is defined in (2.35). According to Proposition 2 on p. 16 of [2]

$$\Lambda_u = h_u^T S_u(\theta) - \frac{1}{2} h_u^T \mathbf{J}(\theta) h_u + o_P(1) = h^T S_u(\theta) - \frac{\sigma^2}{2} + o_P(1),$$

where $P = P_u$ and σ^2 is defined in (2.36). This together with (2.12) means that for a fixed vector $g \in R^{mq}$ and $T_u = g^T \mathbf{D_u}(\hat{\theta}_{(u)} - \theta)$

$$\begin{pmatrix} T_u \\ \Lambda_u \end{pmatrix} = \begin{pmatrix} g^T \mathbf{J}(\theta)^{-1} \\ h^T \end{pmatrix} S_u(\theta) - \begin{pmatrix} 0 \\ \sigma^2/2 \end{pmatrix} + o_P(1).$$

Hence $\mathcal{L}(T_u \,|P_u^*) \longrightarrow N(g^T h, g^T \mathbf{J}(\theta)^{-1} g)$ by Le Cam's third lemma (cf. p. 503 in [2] or [7], p. 208), and (2.34) is proved. □

Lemma 2.7 *Suppose that $C \subset R^p$ is a cone, $\lim_{u\to\infty} \mathbf{M_u} = \mathbf{M}$ is a regular $p \times p$ matrix, $p = mq$ and v is the distance (2.15).*

(I) If $\lim_{u\to\infty} y_u = y \in R^p$, then $\lim_{u\to\infty} v(y_u, \mathbf{M_u}C) = v(y, \mathbf{M}C)$.

(II) $\sup\{\,|v(y, \mathbf{M_u}C) - v(y, \mathbf{M}C)|;\ y \in K\,\} \longrightarrow 0$ as $u \to \infty$ provided that the non-empty set $K \subset R^p$ is compact.

Proof. (I) Since $|v(y, \mathbf{M_u}C) - v(y_u, \mathbf{M_u}C)| \leq |||y - y_u|||$, we may assume that $y_u \equiv y$. But if Π, Π_u denotes projection on $\mathbf{M}C$ and $\mathbf{M_u}C$ respectively, then $\Pi(y) = \mathbf{M}z$, $\Pi_u(y) = \mathbf{M_u}z_u$ where z, z_u belong to C. Hence

$$v(y, \mathbf{M_u}C) \leq |||y - \mathbf{M_u}z||| \leq v(y, \mathbf{M}C) + |||\mathbf{M}z - \mathbf{M_u}z|||,$$

and

$$\limsup_{u\to\infty} v(y, \mathbf{M_u}C) \leq v(y, \mathbf{M}C).$$

Similarly, $v(y, \mathbf{M}C) \leq v(y, \mathbf{M_u}C) + |||\mathbf{M_u}z_u - \mathbf{M}z_u|||$. But

$$|||\mathbf{M_u}z_u - \mathbf{M}z_u||| \leq \|\mathbf{J}(\theta)\|^{1/2} ||\mathbf{M_u} - \mathbf{M}||\, \|z_u\|,$$

$$\|z_u\| \leq \|\left(\mathbf{J}(\theta)^{1/2}\mathbf{M_u}\right)^{-1}\|\, |||\mathbf{M_u}z_u||| \leq \|\left(\mathbf{J}(\theta)^{1/2}\mathbf{M_u}\right)^{-1}\|\, |||y|||,$$

where the last inequality holds owing to the inequality $|||\Pi_u(y)||| \leq |||y|||$, following from Theorem 8.2.5 on p. 376 of [13]. Thus also the inequality

$$v(y, \mathbf{M}C) \leq \liminf_{u\to\infty} v(y, \mathbf{M_u}C)$$

holds.

(II) Let $\delta_u = \sup\{\, |v(y, \mathbf{M_u}C) - v(y, \mathbf{M}C)|;\ y \in K\,\}$. Since the function $v(., A)$ is continuous and the set K is compact, there exists a point $y_u \in K$ with the property that $|v(y_u, \mathbf{M_u}C) - v(y_u, \mathbf{M}C)| = \delta_u$. Choose a subsequence $\{u_t\}_{t=1}^{\infty}$ for which

$$\limsup_{u\to\infty} \delta_u = \lim_{t\to\infty} \delta_{u_t}\,.$$

Since the set K is compact, there exists a subsubsequence $\{u_{t_n}\}_{n=1}^{\infty}$ such that

$$\lim_{n\to\infty} y_{u_{t_n}} = y \in K\,,$$

and by (I)

$$\begin{aligned}\limsup_{u\to\infty} \delta_u &= \lim_{n\to\infty} |v(y_{u_{t_n}}, \mathbf{M_{u_{t_n}}}C) - v(y_{u_{t_n}}, \mathbf{M}C)| \\ &= |v(y, \mathbf{M}C) - v(y, \mathbf{M}C)| = 0.\end{aligned}$$

□

Lemma 2.8 *Under validity of the assumptions of Theorem 1.2*

$$\log L(x^{(u)}, \Omega_i) - \left[\log L(x^{(u)}, \hat\theta_{(u)}) - \frac{1}{2}\, v^2\Big(\mathbf{D_u}(\hat\theta_{(u)} - \theta), \mathbf{D}(\mathbf{p})^{1/2} C_i\Big)\right] = o_P(1)\,. \tag{2.37}$$

Here $P = P_\theta^{(u)}$, and v, $\hat\theta_{(u)}$ and the involved matrices are defined in (2.15), (2.16), (2.7) and (1.23), respectively.

Proof. Let G_u be the set described with (2.17) and (2.18), where $\Omega = \Omega_i$ and $\tilde\theta_u = \tilde\theta^{(i)}_{n_1^{(u)},\ldots,n_q^{(u)}}$. Let $\varepsilon > 0$. By means of Lemma 2.3 we easily obtain that for all M, $\tilde M$ sufficiently large the inequality $\limsup_{u\to\infty}\left[1 - P_\theta^{(u)}(G_u)\right] < \varepsilon$ holds. Hence if we show that for the set $\Omega_i^{(u)}(\tilde M) = \{\,\theta^* \in \Omega_i;\ \|\mathbf{D_u}(\theta^* - \theta)\| \leq \tilde M\,\}$ and for M, $\tilde M$ sufficiently large

$$\left|v^2\Big(\mathbf{D_u}\hat\theta_{(u)}, \mathbf{D_u}\Omega_i^{(u)}(\tilde M)\Big) - v^2\Big(\mathbf{D_u}(\hat\theta_{(u)} - \theta), \mathbf{D}(\mathbf{p})^{1/2} C_i\Big)\right| I_{G_u}(x^{(u)}) = o_P(1)\,, \tag{2.38}$$

then with $g_u(x^{(u)})$ standing for the left-hand side of (2.37) we get from (2.19) that for every $\delta > 0$ the inequalities

$$\limsup_{u\to\infty} P_\theta^{(u)}\left[\,|g_u(x^{(u)})| \geq \delta\right] \leq \limsup_{u\to\infty}\left[1 - P_\theta^{(u)}(G_u)\right] < \varepsilon$$

hold, and (2.37) will be proved.

Let

$$\overline{C}_i^{(u)}(M) = \{y \in C_i;\ \|\mathbf{D_u} y\| \le \sqrt{\frac{\lambda_1}{\lambda_{mq}}} M\,\}\,,$$

where λ_1 denotes the largest and λ_{mq} the smallest characteristic root of $\mathbf{J}(\theta)$. According to Theorem 8.2.5 in [13], projection on cone does not enlarge the norm, therefore on the set G_u the equality

$$v^2\Big(\mathbf{D_u}(\hat{\theta}_{(u)} - \theta), \mathbf{D_u} C_i\Big) = v^2\Big(\mathbf{D_u}(\hat{\theta}_{(u)} - \theta), \mathbf{D_u}\overline{C}_i^{(u)}(M)\Big)$$

holds. Since $\Big|\, |||x|||^2 - |||y|||^2 \,\Big| \le \|x+y\|\,\|\mathbf{J}(\theta)\|\,\|x-y\|$, for $x^{(u)} \in G_u$

$$v^2\Big(\mathbf{D_u}\hat{\theta}_{(u)}, \mathbf{D_u}\Omega_i^{(u)}(\tilde{M})\Big) \;-\; v^2\Big(\mathbf{D_u}(\hat{\theta}_{(u)} - \theta), \mathbf{D_u} C_i\Big)$$

$$\le \sup_{y \in \overline{C}_i^{(u)}(M)} \;\inf_{\theta^* \in \Omega_i^{(u)}(\tilde{M})} \|\mathbf{D_u}(\hat{\theta}_{(u)} - \theta^*) + \mathbf{D_u}(\hat{\theta}_{(u)} - \theta - y)\|\,\|\mathbf{J}(\theta)\|\,\|\mathbf{D_u}(\theta + y - \theta^*)\|$$

$$\le \sup_{y \in \overline{C}_i^{(u)}(M)} (2M + \tilde{M} + \sqrt{\frac{\lambda_1}{\lambda_{mq}}} M)\,\|\mathbf{J}(\theta)\|\,\|\mathbf{D_u}\|\,\rho(\,\theta + y, \Omega_i^{(u)}(\tilde{M})\,)\,. \tag{2.39}$$

Put

$$n^{(u)} = n_1^{(u)} + \ldots + n_q^{(u)}.$$

From (1.19) we obtain that for all u sufficiently large

$$\sup\{\,\|y\|;\ y \in \overline{C}_i^{(u)}(M)\,\} \;\le\; \|\mathbf{D_u^{-1}}\| \sqrt{\frac{\lambda_1}{\lambda_{mq}}} M \;\le\; \frac{QM}{\sqrt{n^{(u)}}}\,, \qquad Q = \left[\frac{\lambda_1}{\lambda_{mq}} \sum_{j=1}^{q} \frac{2m}{p_j}\right]^{1/2}\,.$$

Hence employing (1.15) we see that for all u sufficiently large the relations

$$y \in \overline{C}_i^{(u)}(M)\,, \quad \theta^* \in \Omega_i\,, \quad \rho(\theta + y, \theta^*) < \rho(\theta + y, \Omega_i) + 1/n^{(u)}$$

imply that

$$\|\theta^* - \theta\| \;\le\; \|\theta^* - (\theta + y)\| \;+\; \|y\| \;<\; \frac{3QM}{\sqrt{n^{(u)}}}\,,$$

$$\|\mathbf{D_u}(\theta^* - \theta)\| \;\le\; \|\mathbf{D_u}\|\,\|\theta^* - \theta\| \;<\; 3QM\sqrt{m}\,,$$

and if $\tilde{M} > 3QM\sqrt{m}$, then for $y \in \overline{C}_i^{(u)}(M)$ the equality $\rho(y + \theta, \Omega_i) = \rho(\,y + \theta, \Omega_i^{(u)}(\tilde{M})\,)$ holds. This together with (2.39) and the definition of approximability means that given $M > 0$ there exists $\tilde{M} > 0$ such that for all u sufficiently large on G_u

$$v^2\Big(\,\mathbf{D_u}\hat{\theta}_{(u)}, \mathbf{D_u}\Omega_i^{(u)}(\tilde{M})\,\Big) \;-\; v^2\Big(\,\mathbf{D_u}(\hat{\theta}_{(u)} - \theta), \mathbf{D_u}C_i\,\Big) \;\leq\; o(1)\,. \tag{2.40}$$

If u is sufficiently large, then for every $\theta^* \in \Omega_i^{(u)}(\tilde{M})$

$$\|\theta^* - \theta\| \;\leq\; \|\mathbf{D_u^{-1}}\|\,\|\mathbf{D_u}(\theta^* - \theta)\| \;\leq\; \frac{Q\tilde{M}}{\sqrt{n^{(u)}}}$$

and the distance $\rho(\theta^* - \theta, C_i)$ is attained at a vector $y \in C_i$, for which

$$\|\mathbf{D_u}y\| \;\leq\; \|\mathbf{D_u}\|\,\|y\| \;\leq\; \|\mathbf{D_u}\|\,\|\theta^* - \theta\| \;\leq\; Q\tilde{M}\sqrt{m}\,.$$

Thus similarly as in (2.39) for all u sufficiently large on G_u

$$\begin{aligned} &v^2\Big(\,\mathbf{D_u}(\hat{\theta}_{(u)} - \theta), \mathbf{D_u}C_i\,\Big) \;-\; v^2\Big(\,\mathbf{D_u}\hat{\theta}_{(u)}, \mathbf{D_u}\Omega_i^{(u)}(\tilde{M})\,\Big) \\ &\leq v^2\Big(\,\mathbf{D_u}(\hat{\theta}_{(u)} - \theta), \mathbf{D_u}C_i^{(u)}(Q\tilde{M}\sqrt{m})\,\Big) \;-\; v^2\Big(\,\mathbf{D_u}\hat{\theta}_{(u)}, \mathbf{D_u}\Omega_i^{(u)}(\tilde{M})\,\Big) \\ &\leq\; \mathcal{O}(1)\,\|\mathbf{D_u}\| \sup_{\theta^* \in \Omega_i^{(u)}(\tilde{M})} \rho\Big(\,\theta^* - \theta\,,\, C_i^{(u)}(Q\tilde{M}\sqrt{m})\,\Big) = o(1)\,, \end{aligned}$$

where the last equality follows from definition of approximability. Hence (2.40) remains true also when the left-hand side is taken with the absolute value. Finally, let $\hat{p}_j = n_j^{(u)}/n^{(u)}$, $j = 1, \ldots, q$ denote relative sample sizes from particular populations. Then $\mathbf{D_u}C_i = \mathbf{D}(\hat{\mathbf{p}})^{1/2}C_i$, according to Lemma 2.7

$$\Big|\,v^2\Big(\,\mathbf{D_u}(\,\hat{\theta}_{(u)} - \theta\,), \mathbf{D_u}C_i\,\Big) \;-\; v^2\Big(\,\mathbf{D_u}(\,\hat{\theta}_{(u)} - \theta\,), \mathbf{D}(\mathbf{p})^{1/2}C_i\,\Big)\,\Big|\,I_{G_u}(x^{(u)}) = o_P(1)$$

and validity of (2.38) is proved. □

P r o o f o f T h e o r e m 1. 2. By Lemma 2.8,

$$\begin{aligned} &2\log\frac{L(x^{(u)}, \Omega_1)}{L(x^{(u)}, \Omega_0)} \\ =\;\; & v^2\Big(\mathbf{D_u}(\,\hat{\theta}_{(u)} - \theta\,), \mathbf{D}(\mathbf{p})^{1/2}C_0\Big) \;-\; v^2\Big(\mathbf{D_u}(\,\hat{\theta}_{(u)} - \theta\,), \mathbf{D}(\mathbf{p})^{1/2}C_1\Big) \;+\; o_P(1) \\ =\;\; & \rho^2(\xi_u, G_0) - \rho^2(\xi_u, G_1) \;+\; o_P(1)\,, \end{aligned}$$

where $P = P_\theta^{(u)}$ and $\xi_u = \mathbf{J}(\theta)^{1/2}\mathbf{D_u}(\,\hat{\theta}_{(u)} - \theta\,)$. Since the functions $\rho^2(.,G_i)$ are continuous, (1.22) follows from Lemma 2.6.

Further, let $C_0 \subset C_1$ be linear subspaces of R^{mq}. Then also $G_0 \subset G_1$ are linear subspaces and similarly as in the proof of Theorem 1.1(II) one easily finds out that

$$\rho^2(x, G_0) - \rho^2(x, G_1) = x^T\mathbf{A}x\,,$$

where the matrix $\mathbf{A} = \boldsymbol{\Psi}_\mathbf{1} - \boldsymbol{\Psi}_\mathbf{0}$ is symmetric, idempotent and $\boldsymbol{\Psi}_\mathbf{i}$ denotes the matrix of projection on G_i. Hence the assumptions of Theorem 9.2.1 in [10] are in this case fulfilled with $\boldsymbol{\Sigma} = \mathbf{I_{mq}}$, $\mu = \mathbf{J}(\theta)^{1/2}h$, which implies that $\mathcal{L}(x^T\mathbf{A}x \,|\, N(\mu, \boldsymbol{\Sigma})) = \chi^2_s(\lambda)$, where the degrees of freedom $s = \mathrm{tr}(\mathbf{A}\boldsymbol{\Sigma}) = \mathrm{rank}(\boldsymbol{\Psi}_\mathbf{1}) - \mathrm{rank}(\boldsymbol{\Psi}_\mathbf{0}) = k_1 - k_0$ and the non-centrality parameter $\lambda = \mu^T\mathbf{A}\mu = \rho^2(\mu, G_0) - \rho^2(\mu, G_1)$. □

Proof of the Corollary 1.2 will be based on the following lemma, which probably does not contain new results, because the involved cones are termed in the literature as tangent cones. Since the property (1.15) of the sequential approximability was not previously mentioned in the available literature, we prefer to include the assertion into the text.

Lemma 2.9 *Let $\Theta \subset R^{mq}$ be an open set.*

(I) Let $\theta \in \Omega \subset \Theta$ and

$$\Omega \cap W = \left\{ \begin{pmatrix} x \\ \eta(x) \end{pmatrix} ; x \in V \right\},$$

where $W \subset R^{mq}$ is an open set containing θ, $V \subset R^s$ is an open set, $s < mq$ and $\eta : V \to R^{mq-s}$ belongs to $\mathcal{C}_1$. Then Ω is at θ sequentially approximable by the cone

$$C = \left\{ z \in R^{mq}; \begin{pmatrix} z_{s+1} \\ \vdots \\ z_{mq} \end{pmatrix} = \mathbf{d} \begin{pmatrix} z_1 \\ \vdots \\ z_s \end{pmatrix} \right\}, \tag{2.41}$$

where

$$\mathbf{d} = \mathbf{d}\,[\,\eta\,]\,(\vartheta) = \begin{pmatrix} \frac{\partial \eta_1(\vartheta)}{\partial \vartheta_1}, & \dots & , \frac{\partial \eta_1(\vartheta)}{\partial \vartheta_s} \\ \vdots & & \vdots \\ \frac{\partial \eta_{mq-s}(\vartheta)}{\partial \vartheta_1}, & \dots & , \frac{\partial \eta_{mq-s}(\vartheta)}{\partial \vartheta_s} \end{pmatrix} \tag{2.42}$$

and $\vartheta = (\theta_1, \dots, \theta_s)^T$ consists of the first s coordinates of θ.

(II) If the matrix (1.27) is of rank k_i and its elements are functions continuous on Θ, then the set (1.26) is at θ sequentially approximable by the cone

$$C_i = \{\, y \in R^{mq};\, \boldsymbol{\partial_i}(\theta) y = \mathbf{0} \,\}. \tag{2.43}$$

Proof. (I) The proof can be easily carried out by means of the definition of differentiable real-valued function.

(II) Since it is only a matter of notation, we may assume that the last k_i columns of (1.27) are linearly independent. Since $g = (g_1, \dots, g_{k_i})^T$ belongs to $\mathcal{C}_1$ and $g(\theta) = \mathbf{0}$, from theorem on implicit functions one obtains that there exist a neighbourhood

$U \subset R^{mq-k_i}$ of $(\theta_1, \ldots, \theta_{mq-k_i})^T$, a neighbourhood $V \subset R^{k_i}$ of $(\theta_{mq-k_i+1}, \ldots, \theta_{mq})^T$ and a mapping $\eta : U \to V$ belonging to $\mathcal{C}_1$ such that

$$W = \{\, (x^T\,,\, y^T\,)^T;\ x \in U\,,\, y \in V\,\}$$

is a subset of Θ, for every $x \in U$ the only point $y \in V$ satisfying $g(\,(x^T, y^T)^T\,) = \mathbf{0}$ is $y = \eta(x)$ and the matrix (2.42) has for every $\vartheta \in U$ the form

$$\mathbf{d}\,[\,\eta\,]\,(\vartheta) = -\left[\mathbf{D}\begin{pmatrix}\vartheta\\ \eta(\vartheta)\end{pmatrix}\right]^{-1}\mathbf{H}\begin{pmatrix}\vartheta\\ \eta(\vartheta)\end{pmatrix}. \tag{2.44}$$

Here $s = mq - k_i$ and $\boldsymbol{\partial_i}(\theta) = (\,\mathbf{H}(\theta)\,\mathbf{D}(\theta)\,)$ is the partition of the matrix (1.27) into the blocks determined by the last k_i columns. Thus

$$\Omega_i \cap W = \left\{\begin{pmatrix}x\\ \eta(x)\end{pmatrix};\ x \in U\right\}$$

and (2.44) means that the cone (2.43) equals (2.41). □

P r o o f o f C o r o l l a r y 1. 2. The approximating cone C_0 in this case equals (2.43) with $i = 0$, and putting $\Omega_1 = \Theta$, $C_1 = R^{mq}$ we see that (1.24) holds with $s = mq - (mq - k_0)$,

$$\lambda \;=\; \rho^2(\mathbf{J}(\theta)^{1/2}h, G_0)\,, \quad G_0 = \{\, z \in R^{mq};\ \mathbf{A}z = \mathbf{0}\,\}\,, \quad \mathbf{A} = \mathbf{F_0}\mathbf{J}(\theta)^{-1/2}\,.$$

Since $\rho^2(x, G_0) \;=\; x^T\mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}\mathbf{A}x$, validity of (I) is proved. The assertion (II) can be proved similarly and validity of (III) is obvious. □

References

[1] O. E. Barndorf-Nielsen and D. R. Cox: *Inference and Asymptotics.* Chapmann & Hall, London, 1994.

[2] P. J. Bickel, C. A. J. Klaassen, Y. Ritov and J. A. Wellner: *Efficient and Adaptive Estimation for Semiparametric Models.* John Hopkins University Press, Baltimore, 1993.

[3] P. Billingsley: *Convergence of Probability Measures.* Wiley & Sons, New York, 1968.

[4] H. Chernoff: On the distribution of the likelihood ratio. *Ann. Math. Statist.*, 25(1954), 573–578.

[5] R. R. Davidson and W. L. Lever: The limiting distribution of the likelihood ratio statistic under a class of local alternatives. *Sankhya, Ser. A*, 32(1970), 209–224.

[6] P. I. Feder: On the distribution of the log likelihood ratio test statistic when the true parameter is "near" the boundaries of the hypothesis regions. *Ann. Math. Statist.*, 39(1968), 2044–2055.

[7] J. Hájek and Z. Šidák: *Theory of Rank Tests*. Academia, Prague, 1967.

[8] N. Inagaki: Asymptotic relations between the likelihood estimating function and the maximum likelihood estimate. *Ann. Inst. Stat. Math.*, 25(1973), 1–26.

[9] C. R. Rao: *Linear Statistical Inference and Its Applications*. Wiley & Sons, New York, 1965.

[10] C. R. Rao and S. K. Mitra: *Generalized Inverse of Matrices and its Applications*. Wiley & Sons, New York, 1971.

[11] P. K. Sen and M. Singer: *Large Sample Methods in Statistics*. Chapmann & Hall, London, 1993.

[12] R. J. Serfling: *Approximation Theorems of Mathematical Statistics*. Wiley & Sons, New York, 1980.

[13] F. T. Wright T. Robertson and R. L. Dykstra: *Order Restricted Statistical Inference*. Wiley & Sons, New York, 1988.

Institute of Measurement of SAS
Dúbravská cesta 9
84104 Bratislava
Slovakia