Tatra Mt. Math. Publ. **22** (2001), 115–141

TATRA MOUNTAINS
Mathematical Publications

ON CONSISTENCY OF THE MAXIMUM LIKELIHOOD PRINCIPLE BASED TESTS AND ESTIMATES IN THE CASE OF EXPONENTIAL-LIKE DISTRIBUTIONS

František Rublík

Dedicated to Lubomír Kubáček on the occasion of his 70th birthday

ABSTRACT. Consistency of the likelihood ratio tests is established under the assumption that the infimum of the Kullback-Leibler information number on the hypothesis of interest is a real number and certain regularity conditions are fulfilled. Under these conditions also consistency of the MLE is proved in the case that the parameter minimizing the Kullback-Leibler information number is unique. The results are proved in the multisample setting and applied to exponential distributions with unknown lower bound.

1. Introduction

One of the most frequently used estimators is the maximum likelihood estimator. One of the reasons is that it has good properties not only as an estimator, but also as a basis for construction of the likelihood ratio tests as well. Asymptotic optimality of the likelihood ratio tests (described for example in [2], [11], [12], [13] and [15]), contributes to the fact that these tests belong to the most popular tools of hypotheses testing and supports the use of the maximum likelihood estimate in a wide range of cases.

One of the basic properties of any estimator is its consistency. While in the unrestricted case the consistency of the maximum likelihood estimator (MLE) was one of the classical results of the estimation theory, in the case of estimation under the restriction this consistency was in an abstract setting treated only later

2000 Mathematics Subject Classification: 62F12, 62F05.
Keywords: existence and consistency of the maximum likelihood estimator, consistency of the likelihood ratio tests, exponential distribution with unknown lower bounds, multisample setting.

Supported by Grant VEGA 1/7295/20.

in the paper [17], where the hypothesis under the consideration is assumed to be a compact set. Even though many situations do not meet this requirement, the ideas presented in [17] turned out to be very useful and later were successfully used in many statistical papers. Here one should emphasize that the study of the existence and consistency has not only an esthetic value, but is also of practical importance, especially in the situations when an explicit formula for the MLE is unknown and its computation is carried out by some numerical method.

The existence and the uniqueness of the MLE is in an abstract setting (including differentiability of the likelihood function) studied in [8], results on this topic in the case of exponential densities can be also found in [10]. The existence and algorithms for computation of the MLE in the exponential families of distributions are studied in [9]. Classical results on consistency of the unrestricted MLE in the differentiable case can be found in Section 4.2 of [6], sufficient conditions for the consistency of the unrestricted MLE can also be found in [5]. Consistency of the MLE under restrictions is for the exponential families of densities in the one-sample case proved in [3]. In [4] the consistency of the MLE is proved in the one-sample case in an abstract setting, allowing the set of the restrictions to be any set closed in the Euclidean space, and not requiring continuity of the underlying densities, but no assertions on existence of the concerned estimates are included there. Consistency of the MLE in the setting including exponential family of densities is in the multisample case proved in [14], where in the case of the exponential families also the existence of the MLE is proved.

In this paper the consistency of the MLE is established in the multisample case in the setting, requiring special type of densities but allowing them to have discontinuity points. The presented results include also consistency of the likelihood ratio tests.

The regularity conditions and general results on consistency are presented in Section 2 and their application to the exponential distributions with unknown lower bounds (defined by means of the densities (3.2)) is presented in the Section 3. The proofs of these results are in the Section 4, where also a generalization of the Le Cam assertion on existence of MLE in the compact case can be found.

2. Regularity conditions and the main assertions

Throughout the paper we shall deal with inference on $q \geq 1$ statistical populations. Distribution of each population is supposed to belong to the class of probabilities $\{\overline{P}_\gamma \colon \gamma \in \Xi\}$, which are defined on $(X, \mathcal{F})$ by means of the densities

$$f(x,\gamma) = \frac{d\overline{P}_\gamma}{d\nu}(x)\,. \tag{2.1}$$

Thus the parameter space is the q-fold Cartesian product

$$\Theta = \Xi^q \,, \tag{2.2}$$

i.e., $\theta = (\theta_1, \ldots, \theta_q)$ is the overall parameter and its component $\theta_j \in \Xi$ describes the state of the jth population.

Let $x(j, n_j) = \big(x_1^{(j)}, \ldots, x_{n_j}^{(j)}\big)$ denote the outcome of the sampling from the jth population. Then

$$x_{(n_1,\ldots,n_q)} = \big(x(1, n_1), \ldots, x(q, n_q)\big) \tag{2.3}$$

is the pooled sample and its distribution is the product measure

$$P_\theta = \overline{P}_{\theta_1}^{(n_1)} \times \cdots \times \overline{P}_{\theta_q}^{(n_q)} \,. \tag{2.4}$$

In this notation $\overline{P}_{\theta_1}^{(n_1)}$ denotes the Cartesian product of n_1 copies of this probability, for the sake of brevity the numbers $n_1, \ldots, n_q$ are not included into the left-hand side of (2.4). Another meaning of this symbol is defined in (2.27), but these two meanings do not contradict each other because (2.27) is an extension of (2.4), and it is always clear from the context, which of the meanings is involved in the concerned text.

If $(x_1, \ldots, x_n)$ denotes random sample from one population, then for $\gamma \in \Xi$ the likelihood function is defined by the formula

$$L(x_1, \ldots, x_n, \gamma) = \prod_{i=1}^{n} f(x_i, \gamma) \,.$$

For $\Omega \subset \Theta$ its multisample analogue

$$L(x_{(n_1,\ldots,n_q)}, \Omega) = \sup \left\{ \prod_{j=1}^{q} \prod_{i=1}^{n_j} f(x_i^{(j)}, \theta_j) \colon (\theta_1, \ldots, \theta_q) \in \Omega \right\} , \tag{2.5}$$

i.e., (2.5) denotes the likelihood function of the pooled sample (2.3).

In the following considerations an important role will belong to the Kullback-Leibler information quantity, which is defined by the formula

$$K(\gamma, \gamma^*) = \int \left(\log \frac{f(x, \gamma)}{f(x, \gamma^*)} \right) f(x, \gamma)\, \nu(x) \,, \tag{2.6}$$

where log denotes the logarithm to the base e and

$$\log \frac{a}{b} = \begin{cases} \log \frac{a}{b} \,, & a > 0,\ b > 0 \,, \\ +\infty \,, & a > 0,\ b = 0 \,, \\ -\infty \,, & a = 0,\ b > 0 \,, \\ 0 \,, & a = 0,\ b = 0 \,. \end{cases} \tag{2.7}$$

It is well-known that the integral (2.6) exists, is nonnegative with $+\infty$ as a possible value and equals zero if and only if $\overline{P}_{\gamma} = \overline{P}_{\gamma^*}$ (e.g., [1] Theorem 4.1, p. 9).

The densities (2.1) will be subjected to regularity assumptions.

(AI) Ξ is an open subset of $\mathbb{R}^m$, the probabilities $\{\overline{P}_{\gamma}\colon \gamma \in \Xi\}$ are mutually different and the function (2.5) is measurable whenever $\Omega = \Theta \cap C$, where $C \subset \mathbb{R}^{mq}$ is either open or closed.

(AII) Let γ, $\gamma^* \in \Xi$. There exist a neighbourhood

$$V = V(\gamma^*, \delta^*) = \big\{\tilde{\gamma}\colon \|\gamma^* - \tilde{\gamma}\| < \delta^*\big\}$$

and a nonnegative measurable function $G \in L_1(\overline{P}_{\gamma})$ such that

$$\sup\big\{\big|\log f(x,\tilde{\gamma})\big|\colon \tilde{\gamma} \in V,\ K(\gamma,\tilde{\gamma}) < +\infty\big\} \le G(x)\,,$$
$$\log L(x,V) \le G(x)\,,$$

where $\sup\emptyset = 0$ and $L(x,V) = \sup\big\{f(x,\tilde{\gamma})\colon \tilde{\gamma} \in V\big\}$. Moreover, if $f(x,\gamma_n) > 0$ and $\gamma_n \to \gamma$, then $f(x,\gamma_n) \to f(x,\gamma)$.

(AIII) Let $\gamma_n \to \gamma$, $\gamma_n^* \to \gamma^*$ belong to Ξ. If $K(\gamma_n,\gamma_n^*) < +\infty$ for every n, then $K(\gamma_n,\gamma_n^*) \longrightarrow K(\gamma,\gamma^*)$.

(AIV) If Γ is a non-empty compact subset of Ξ and c is a positive real number, then

$$\big\{\gamma^* \in \Xi\colon \text{there exists } \gamma \in \Gamma \text{ such that } K(\gamma,\gamma^*) \le c\big\}$$

is a compact subset of $\mathbb{R}^m$.

(AV) There exist measurable sets $A_n \subset X^n$, a positive integer N and measurable mappings

$$\hat{\gamma}_n\colon A_n \to \Xi\,, \qquad g_n\colon X^n \to \mathbb{R}^1$$

such that the following assertions (a) and (b) hold.

(a) If $n \ge N$, then

$$\big\{(x_1,\dots,x_n,x_{n+1})\colon (x_1,\dots,x_n) \in A_n\big\} \subset A_{n+1} \tag{2.8}$$

and on A_n the equality (cf. (2.6))

$$\log L(x_1,\dots,x_n,\gamma) = g_n(x_1,\dots,x_n) - nK\big(\hat{\gamma}_n(x_1,\dots,x_n),\gamma\big) \tag{2.9}$$

holds for each $\gamma \in \Xi$.

(b) The postulated sets and mappings fulfill for each $\gamma \in \Xi$ the equalities

$$\lim_{n\to\infty} \overline{P}_{\gamma}^{(n)}(A_n) = 1\,, \qquad \lim_{n\to\infty} \hat{\gamma}_n(x_1,\dots,x_n) = \gamma \ \text{ a.e. } \overline{P}_{\gamma}^{\infty}. \tag{2.10}$$

If $K(\hat{\gamma}_n(x_1,\dots,x_n),\gamma)=+\infty$, then (2.9) means that $L(x_1,\dots,x_n,\gamma)=0$. Here it should be also noted that the notation $V(\gamma^*,\delta^*)$ in (AII) is not intended to express that the postulated neighbourhood does not depend on γ (typically it does!); the aim of this notation is to introduce a symbol for the neighbourhood, which will be used also in the further text.

The following Theorems 2.1 and 2.2 are an extension of results of [14] into the setting postulated in this paper.

Theorem 2.1. *Suppose that the assumptions* (AI)–(AV) *hold and* (*cf.* (2.2))

$$\Omega=\Theta\cap C\,,\qquad C\subset\mathbb{R}^{mq}\ \text{is a closed set.}\tag{2.11}$$

Let

$$D_{n_1,\dots,n_q}=A_{n_1}\times\cdots\times A_{n_q}\,,\tag{2.12}$$

where A_n *is the set from* (AV). *Then there exist measurable mappings*

$$\tilde{\theta}_{n_1,\dots,n_q}\colon D_{n_1,\dots,n_q}\to\Omega\tag{2.13}$$

such that the equality (*cf.* (2.5))

$$L\Big(x_{(n_1,\dots,n_q)}\,,\,\tilde{\theta}_{n_1,\dots,n_q}\big(x_{(n_1,\dots,n_q)}\big)\Big)=L\Big(x_{(n_1,\dots,n_q)},\Omega\Big)\tag{2.14}$$

holds for every point $x_{(n_1,\dots,n_q)}\in D_{n_1,\dots,n_q}$.

Since the Kullback-Leibler information quantity is nonnegative, the mapping $\hat{\gamma}_n$ postulated in (AV) is the maximum likelihood estimate of the unknown parameter. Thus an interpretation of the previous theorem is that the MLE under the (relatively closed) restriction exists in the points when the unrestricted MLE exists.

Before proceeding further, let us introduce some more notations related to the Kullback-Leibler information number. Let E^q denote the set of q-dimensional row vectors and

$$\mathcal{P}=\Big\{p=\big(p(1),\dots,p(q)\big)\in E^q\colon\ \sum_{j=1}^{q}p(j)=1\,,\ \min_j p(j)>0\Big\}\,.\tag{2.15}$$

For $\theta=(\theta_1,\dots,\theta_q)$, $\theta^*=(\theta_1^*,\dots,\theta_q^*)$ belonging to Θ and for a vector $p=(p_1,\dots,p_q)$ from (2.15) put (cf. (2.6))

$$K(\theta,\theta^*,p)=\sum_{j=1}^{q}p_jK(\theta_j,\theta_j^*)\,,\tag{2.16}$$

and for $\Omega\subset\Theta$ let

$$K(\theta,\Omega,p)=\inf\big\{K(\theta,\theta^*,p)\colon\ \theta^*\in\Omega\big\}\,.\tag{2.17}$$

Suppose now that a sequence of experiments is indexed by the integer $u = 1, 2, \dots$. In accordance with the previous notation assume that $n_u^{(j)}$ denotes the size of the sample from the jth population and

$$x(j, n_u^{(j)}) = \left(x_1^{(j)}, \dots, x_{n_u^{(j)}}^{(j)}\right)$$

is the sample from the jth population, everything in the uth experiment. In this multisample setting

$$x^{(u)} = \left(x(1, n_u^{(1)}), \dots, x(q, n_u^{(q)})\right) \tag{2.18}$$

denotes the pooled sample in the uth experiment and in accordance with (2.5)

$$L(x^{(u)}, \Omega) = L\left(x_{(n_u^{(1)}, \dots, n_u^{(q)})}, \Omega\right) \tag{2.19}$$

stands for the likelihood function of the pooled sample.

The total sample size in the uth experiment

$$n_u = \sum_{j=1}^{q} n_u^{(j)}, \tag{2.20}$$

and the relative sample size from the jth population

$$p_u^{(j)} = n_u^{(j)} / n_u \,. \tag{2.21}$$

It will be assumed throughout the rest of the section that

$$\lim_{u\to\infty} n_u = +\infty\,, \qquad \lim_{u\to\infty} p_u^{(j)} = p_j \in (0, 1\rangle\,, \qquad j = 1, \dots, q\,. \tag{2.22}$$

THEOREM 2.2. *Suppose that the assumptions* (AI)–(AV) *hold and* (*cf.* (2.2))

$$\Omega = \Theta \cap C\,, \qquad C \subset \mathbb{R}^{mq} \text{ is a closed set}\,.$$

Let

$$D_u = A_{n_u^{(1)}} \times \cdots \times A_{n_u^{(q)}}\,, \tag{2.23}$$

where A_n *is a set from* (AV). *Further, let the mappings* $\tilde\theta_u = \tilde\theta_{n_u^{(1)}, \dots, n_u^{(q)}} :$ $D_u \to \Omega$ *be such that the equality*

$$L\left(x^{(u)}, \tilde\theta_u(x^{(u)})\right) = L\left(x^{(u)}, \Omega\right) \tag{2.24}$$

is fulfilled for all $x^{(u)} \in D_u$. *Finally, let* (2.22) *hold.*

(I) *If* $\theta \in \Theta$, $K(\theta, \Omega, p) < +\infty$ *and there exists a unique parameter* $\eta = \eta(\theta, p) \in \Omega$ *such that*

$$K(\theta, \Omega, p) = K\big(\theta, \eta(\theta, p), p\big)\,, \tag{2.25}$$

then

$$\tilde{\theta}_u \longrightarrow \eta(\theta, p) \tag{2.26}$$

almost everywhere

$$P_\theta = \overline{P}_{\theta_1}^{\infty} \times \cdots \times \overline{P}_{\theta_q}^{\infty}. \tag{2.27}$$

(II) *If* $\theta \in \Omega$, *then* (*cf.* (2.27))

$$\tilde{\theta}_u \longrightarrow \theta \tag{2.28}$$

almost everywhere P_θ.

The following theorem deals with the consistency of the likelihood ratio tests based on the statistic

$$T_u = 2 \log \frac{L\big(x^{(u)}, \Omega_1\big)}{L\big(x^{(u)}, \Omega_0\big)}, \tag{2.29}$$

where the likelihood functions are defined by means of (2.19).

Theorem 2.3. *Suppose that the assumptions* (AI)–(AV) *hold and* (*cf.* (2.2))

$$\emptyset \neq \Omega_0 \subset \Omega_1 \subset \Theta, \quad \Omega_0 = \Theta \cap C, \qquad C \subset \mathbb{R}^{mq} \text{ is a closed set}. \tag{2.30}$$

Let for each $\gamma \in \Xi$ *there exist a probability* $\mathcal{L}_\gamma^*$ *on* $(\mathbb{R}^1, \mathcal{B}^1)$ *such that the weak convergence of distributions*

$$\mathcal{L}\left[\log \frac{L(x_1, \ldots, x_n, \Xi)}{L(x_1, \ldots, x_n, \gamma)} \,\Big|\, \overline{P}_\gamma\right] \longrightarrow \mathcal{L}_\gamma^* \tag{2.31}$$

holds as $n \to \infty$, *where* $L(x_1, \ldots, x_n, \Xi) = \sup\{L(x_1, \ldots, x_n, \gamma) \colon \gamma \subset \Xi\}$. *Assume that* (2.22) *is fulfilled and tests* $\varphi_u = \varphi_u\big(x^{(u)}\big)$ *of the hypothesis* Ω_0 *against the alternative* $\Omega_1 - \Omega_0$ *and the constants* c_u *are such that* (*cf.* (2.29))

$$\varphi_u = \begin{cases} 1, & T_u > c_u, \\ 0, & T_u \leq c_u, \end{cases} \tag{2.32}$$

where 1 *denotes the rejection of* Ω_0 *in favor of* $\Omega_1 - \Omega_0$ *and* 0 *denotes the acceptance of* Ω_0. *If* (*cf.* (2.4))

$$\sup_{\theta \in \Omega_0} \lim_{u \to \infty} P_\theta(\varphi_u = 1) > 0, \tag{2.33}$$

then

$$\lim_{u \to \infty} P_\theta(\varphi_u = 1) = 1 \tag{2.34}$$

for every parameter $\theta \in \Omega_1 - \Omega_0$ *such that* $K(\theta, \Omega_0, p) < +\infty$, *where* $p = (p_1, \ldots, p_q)$ *is the limiting vector from* (2.22).

3. Applications to exponential distribution

Suppose throughout this section that

$$\Xi = \left\{ \begin{pmatrix} \mu \\ \sigma \end{pmatrix} \in \mathbb{R}^2 \colon \sigma > 0 \right\}, \tag{3.1}$$

and for $\gamma = (\mu, \sigma)' \in \Xi$ now the functions (2.1)

$$f(x, \gamma) = \frac{d\overline{P}_\gamma}{d\nu}(x) = \frac{1}{\sigma} \exp\Big(-\frac{x-\mu}{\sigma}\Big) \chi_{[\mu, +\infty)}(x) \tag{3.2}$$

are densities with respect to the Lebesgue measure on the real line; in (3.2) the symbol $\chi_{[\mu,+\infty)}$ denotes the indicator function of the interval $[\mu, +\infty)$. Thus the probabilities $\{\overline{P}_\gamma \colon \gamma \in \Xi\}$ defined on the Borel subsets of the real line by means of these densities are the exponential distributions with unknown lower bound.

The existence of the MLE will be formulated by means of the set

$$A_n = \big\{(x_1, \dots, x_n) \in \mathbb{R}^n \colon \min_i x_i < \max_i x_i \big\}, \tag{3.3}$$

whose complement has Lebesgue measure 0 for every integer $n > 1$.

For these exponential probabilities and the multisample setting described by (2.2)–(2.5) the following assertions hold.

THEOREM 3.1. *Let* (3.1)–(3.3) *hold and* (*cf.* (2.2))

$$\Omega = \Theta \cap C\,, \qquad C \subset \mathbb{R}^{2q} \text{ is a closed set}\,.$$

Let

$$D_{n_1,\dots,n_q} = A_{n_1} \times \dots \times A_{n_q}\,, \tag{3.4}$$

where A_n is the set (3.3). *Then there exist measurable mappings*

$$\tilde{\theta}_{n_1,\dots,n_q} \colon D_{n_1,\dots,n_q} \to \Omega \tag{3.5}$$

such that the equality (*cf.* (2.5))

$$L\Big(x_{(n_1,\dots,n_q)}, \tilde{\theta}_{n_1,\dots,n_q}\big(x_{(n_1,\dots,n_q)}\big)\Big) = L\big(x_{(n_1,\dots,n_q)}, \Omega\big) \tag{3.6}$$

holds for every point $x_{(n_1,\dots,n_q)} \in D_{n_1,\dots,n_q}$.

Thus the previous theorem says that if the sample sizes from each of the q exponential distributions with unknown lower bound are greater than 1, then for every set of restrictions Ω closed in the parameter space $\Theta = \Xi^q$ the MLE exists with probability 1.

Theorem 3.2. *Let (3.1)–(3.3) hold and (cf. (2.2))*

$$\Omega = \Theta \cap C\,, \qquad C \subset \mathbb{R}^{2q} \text{ is a closed set}\,.$$

Let for $u = 1, 2, \dots$

$$D_u = A_{n_u^{(1)}} \times \cdots \times A_{n_u^{(q)}}\,, \tag{3.7}$$

where A_n is the set (3.3). Further, let the mappings $\tilde\theta_u = \tilde\theta_{n_u^{(1)},\dots,n_u^{(q)}} : D_u \to \Omega$ be such that the equality

$$L\Big(x^{(u)}, \tilde\theta_u(x^{(u)})\Big) = L\Big(x^{(u)}, \Omega_0\Big) \tag{3.8}$$

is fulfilled for all $x^{(u)} \in D_u$. Finally, let (2.22) hold.

(I) *Suppose that $\theta = \big((\mu_1, \sigma_1)', \dots, (\mu_q, \sigma_q)'\big) \in \Theta$ and there exists a parameter $\theta^* = \big((\mu_1^*, \sigma_1^*)', \dots, (\mu_q^*, \sigma_q^*)'\big) \in \Omega$ such that*

$$\mu_j^* \le \mu_j \qquad \text{for all } \ j\,. \tag{3.9}$$

If there exists a unique parameter $\eta = \eta(\theta, p) \in \Omega$ such that (cf. (2.17))

$$K(\theta, \Omega, p) = K\Big(\theta, \eta\,(\theta, p)\,,\ p\Big)\,,$$

then, as $u \to \infty$,

$$\tilde\theta_u \longrightarrow \eta(\theta, p) \tag{3.10}$$

almost everywhere P_θ (here P_θ is defined by means of (2.27), where the involved probabilities are determined with (3.2)).

(II) *If $\theta \in \Omega$, then, as $u \to \infty$,*

$$\tilde\theta_u \longrightarrow \theta \tag{3.11}$$

almost everywhere P_θ, where P_θ is described in (I).

Theorem 3.3. *Suppose that (3.1)–(3.3) hold and (cf. (2.2))*

$$\emptyset \ne \Omega_0 \subset \Omega_1 \subset \Theta\,, \quad \Omega_0 = \Theta \cap C\,, \qquad C \subset \mathbb{R}^{2q} \text{ is a closed set}\,.$$

Assume that (2.22) is fulfilled and tests $\varphi_u = \varphi_u\big(x^{(u)}\big)$ of the hypothesis Ω_0 against the alternative $\Omega_1 - \Omega_0$ and the constants c_u are such that (cf. (2.29))

$$\varphi_u = \begin{cases} 1\,, & T_u > c_u\,, \\ 0\,, & T_u \le c_u\,, \end{cases} \tag{3.12}$$

where 1 denotes the rejection of Ω_0 in favor of $\Omega_1 - \Omega_0$ and 0 denotes the acceptance of Ω_0. Let

$$\sup_{\theta \in \Omega_0} \lim_{u \to \infty} P_\theta(\varphi_u = 1) > 0\,. \tag{3.13}$$

If $\theta = \big((\mu_1,\sigma_1)',\dots,(\mu_q,\sigma_q)'\big)$ *is a parameter from* $\Omega_1 - \Omega_0$ *and there exists a parameter* $\theta^* = \big((\mu_1^*,\sigma_1^*)',\dots,(\mu_q^*,\sigma_q^*)'\big) \in \Omega_0$ *such that* (3.9) *holds, then*

$$\lim_{u\to\infty} P_\theta(\varphi_u = 1) = 1\,. \tag{3.14}$$

The previous theorem says that (under the stated assumptions on the true parameter) the likelihood ratio tests are consistent in the multiparameter setting (the relation (3.14)). We remark that the exponential distributions do not fulfill the regularity conditions of Theorem 2 in [16], under which the consistence of the likelihood ratio tests in the one sample setting was proved ibidem.

4. Proofs

LEMMA 4.1. *Suppose that* Ξ *is an open subset of* $\mathbb{R}^m$ *and* (AIV) *holds.*

(I) *Assume that for* $r = 1,2,\dots$ *the parameters* θ_r, θ_r^* *belong to the set* (2.2), *the vectors* p_r *belong to* (2.15) *and there exists a compact set* $C \subset \Theta$ *such that* $\theta_r \in C$ *for all* r. *If* (*cf.* (2.16)

$$\limsup_{r\to\infty} K(\theta_r,\theta_r^*,p_r) < +\infty\,, \tag{4.1}$$

then there exists a compact set $\Gamma \subset \Theta$ *such that*

$$\theta_r^* \in \Gamma\,, \qquad r = 1,2,\dots\,. \tag{4.2}$$

(II) *Assume that also* (AIII) *holds. If* $\{\gamma_r\}_{r=1}^\infty$ *are parameters from* Ξ *and either*

$$\lim_{r\to\infty} \|\gamma_r\| = +\infty \tag{4.3}$$

or

$$\lim_{r\to\infty} \rho\big(\gamma_r,\mathcal{H}(\Xi)\big) = 0\,, \tag{4.4}$$

where $\mathcal{H}(\Xi)$ *denotes the boundary of the set* Ξ *in the usual Euclidean topology and* ρ *is the usual Euclidean distance from the set, then for every parameter* $\gamma \in \Xi$ (*cf.* (2.6))

$$\lim_{r\to\infty} K(\gamma,\gamma_r) = +\infty\,. \tag{4.5}$$

P r o o f . (I) For $\theta = (\theta_1,\dots,\theta_q) \in \Theta$ let

$$\Pi_j(\theta) = \theta_j \tag{4.6}$$

denote the projection on the jth subparametric set. The set

$$C_j = \big\{\Pi_j(\theta)\colon \theta \in C\big\} = \Pi_j(C)$$

is compact, because the projection on coordinate set is a continuous mapping. Since the Kullback-Leibler information number is nonnegative,

$$K\big(\Pi_j(\theta_r)\,,\,\Pi_j(\theta_r^*)\big) \le \frac{1}{p_r(j)} K(\theta_r, \theta_r^*, p_r)\,,$$

where $p_r(j)$ denotes the jth coordinate of p_r. Hence if $\tilde{c}$ denotes the left-hand side of (4.1), then for $j = 1, \dots, q$

$$\limsup_{r\to\infty} K\big(\Pi_j(\theta_r)\,,\,\Pi_j(\theta_r^*)\big) \le c\,, \quad c = \Big(\max_{j=1,\dots,q} \frac{1}{p(j)}\Big)\,\tilde{c}\,. \tag{4.7}$$

According to (AIV) the set

$$\tilde{\Gamma}_j = \big\{\gamma^* \in \Xi\colon \text{there exists } \tilde{\gamma} \in C_j \text{ such that } K(\tilde{\gamma}, \gamma^*) \le c + 1\big\}$$

is compact, hence also

$$\tilde{\Gamma} = \tilde{\Gamma}_1 \times \cdots \times \tilde{\Gamma}_q$$

is compact. Since (4.7) means that $\theta_r \in \tilde{\Gamma}$ for all r sufficiently large and by adding finitely many points to the compact set one obtains again a compact set, the assertion (I) is proved.

(II) Suppose that $\gamma \in \Xi$ is fixed and M is a positive real number. According to (AIII) and (AIV) the set

$$\Gamma = \big\{\tilde{\gamma} \in \Xi\colon\; K(\gamma, \tilde{\gamma}) \le M\big\}$$

is compact.

Let (4.3) hold. Since compact set is also bounded, $\gamma_r \notin \Gamma$ for all r sufficiently large. Hence for these indices r

$$K(\gamma, \gamma_r) > M \tag{4.8}$$

and (4.5) is proved.

Now let (4.4) hold. Since $\rho\big(\tilde{\gamma}, \mathcal{H}(\Xi)\big)$ is a continuous function of $\tilde{\gamma}$ and is positive on Γ because each point of this compact is in interior of Ξ, there exists a point $\gamma^* \in \Gamma$ such that for

$$\delta = \inf\big\{\rho\big(\tilde{\gamma},\, \mathcal{H}(\Xi)\big) : \tilde{\gamma} \in \Gamma\big\}$$

the relations

$$\rho\big(\gamma^*, \mathcal{H}(\Xi)\big) = \delta > 0$$

hold. But (4.4) implies that

$$\rho\big(\gamma_r, \mathcal{H}(\Xi)\big) < \delta$$

for all r sufficiently large. Hence for these indices r the inequality (4.8) holds. □

The following lemma is a generalization of the Lemma 3 from p. 298 of [7] on existence of the measurable MLE in the case of the compact parametric set.

Lemma 4.2. *Suppose that $\mathcal{S}$ is a σ-algebra of subsets of S, Γ is a compact metric space, $\mathcal{B}$ are Borel subsets of Γ and*

$$h\colon S \times \Gamma \to [0, +\infty]$$

is a function having the following properties (I) *and* (II).

(I) *If $s \in S$, $\gamma_n \to \gamma$ and $h(s, \gamma_n) < +\infty$ for all n, then $h(s, \gamma_n) \longrightarrow h(s, \gamma)$.*

(II) *If A is a set open in Γ, then*

$$h(s, A) = \inf\{h(s, \gamma)\colon \gamma \in A\} \tag{4.9}$$

is a measurable function of the argument s.

Then there exists a measurable mapping $T\colon S \to \Gamma$ such that

$$h\big(s, T(s)\big) = h(s, \Gamma) \tag{4.10}$$

for each $s \in S$.

P r o o f . Suppose first that

$$h(s, \Gamma) < +\infty \tag{4.11}$$

for all $s \in S$.

We shall proceed similarly as in the proof of Lemma 2.1 in [14] and we shall use the mappings defined as follows. If

$$S = \bigcup_{i=1}^{m} V_i$$

is a measurable partition of the set S, then

$$Z(s) = \sum_{i=1}^{m} \gamma_i \chi_{V_i}(s)$$

will denote the mapping whose value equals γ_i whenever $s \in V_i$.

Since Γ is a compact metric space, there exist $\{\gamma_j\}_{j=1}^{\infty}$ from Γ and a non-decreasing sequence $\{m_n\}_{n=1}^{\infty}$ of positive integers such that for each n

$$S_n = \{\gamma_1, \dots, \gamma_{m_n}\}$$

is a 2^{-n} net in Γ, i.e.,

$$\Gamma = \bigcup_{j=1}^{m_n} U(\gamma_j, 2^{-n}),$$

where $U(\gamma_j, \delta) = \big\{\gamma^* \in \Gamma\colon \rho(\gamma_j, \gamma^*) < \delta\big\}$. Since the functions $h\big(s, U(\gamma_i, \frac{1}{2})\big)$ are measurable, the sets

$$W_i^{(1)} = \left\{ s \in S\colon h(s, \Gamma) = h\Big(s, U\Big(\gamma_i, \frac{1}{2}\Big)\Big)\right\}$$

are measurable, and

$$V_i^{(1)} = W_i^{(1)} - \bigcup_{j=1}^{i-1} W_j^{(1)}, \qquad i = 1, \dots, m_1,$$

form a measurable partition of S. Thus the mapping

$$T_1(s) = \sum_{i=1}^{m_1} \gamma_i \chi_{V_i^{(1)}}(s)$$

is a measurable mapping from S into Γ.

Let us assume that for $n = 1, \dots, k$ we have already constructed measurable mappings $T_n \colon S \to \Gamma$ such that for all n

$$T_n(s) = \sum_{i=1}^{m_n} \gamma_i \chi_{V_i^{(n)}}(s), \qquad \bigcup_{i=1}^{m_n} V_i^{(n)} = S, \tag{4.12}$$

for all $s \in V_i^{(n)}$

$$h(s, \Gamma) = h\left(s, U\left(\gamma_i, \frac{1}{2^n}\right)\right) \tag{4.13}$$

and for $n = 1, \dots, k-1$

$$\rho\big(T_n(s), T_{n+1}(s)\big) < \frac{1}{2^{n-1}}. \tag{4.14}$$

Denoting

$$W_j^{(k+1,i)} = \left\{ s \in V_i^{(k)} \colon h(s, \Gamma) = h\left(s, U\left(\gamma_i, \frac{1}{2^k}\right) \cap U\left(\gamma_j, \frac{1}{2^{k+1}}\right)\right)\right\},$$

$$V_j^{(k+1,i)} = W_j^{(k+1,i)} - \bigcup_{r=1}^{j-1} W_r^{(k+1,i)}, \qquad j = 1, \dots, m_{k+1},$$

taking into account the fact that $\{V_j^{(k+1,i)} \colon i = 1, \dots, m_k,\ j = 1, \dots m_{k+1}\}$ form a measurable partition of S, and putting

$$T_{k+1}(s) = \sum_{j=1}^{m_{k+1}} \left[\sum_{i=1}^{m_k} \gamma_j \chi_{V_j^{(k+1,i)}}(s) \right] = \sum_{j=1}^{m_{k+1}} \gamma_j \chi_{V_j^{(k+1)}}(s),$$

where

$$V_j^{(k+1)} = \bigcup_{i=1}^{m_k} V_j^{(k+1,i)},$$

we see that for $s \in V_j^{(k+1)} \cap V_i^{(k)}$

$$\rho\big(T_k(s), T_{k+1}(s)\big) \le \frac{1}{2^k} + \frac{1}{2^{k+1}} < \frac{1}{2^{k-1}}\,,$$

and for $s \in V_j^{(k+1)}$

$$h(s,\Gamma) = h\left(s,\, U\Big(\gamma_j, \frac{1}{2^{k+1}}\Big)\right).$$

Hence the existence of measurable mappings $\{T_n\}_{n=1}^{\infty}$ satisfying (4.12)–(4.14) is proved. It is obvious from (4.14) that the limit

$$T(s) = \lim_{n\to\infty} T_n(s) \tag{4.15}$$

exists; this mapping is obviously measurable. Now if $s \in S$ and an integer $n \ge 1$ are fixed, then according to (4.11)–(4.13) there exists an index $i \in \{1,\dots,m_n\}$ such that $s \in V_i^{(n)}$ and

$$+\infty > h(s,\Gamma) = h\left(s,\, U\Big(\gamma_i, \frac{1}{2^n}\Big)\right).$$

Hence for some $\tilde\gamma_i^{(n)} \in U\big(\gamma_i, \frac{1}{2^n}\big)$

$$h(s,\Gamma) \le h\Big(s, \tilde\gamma_i^{(n)}\Big) < h(s,\Gamma) + \frac{1}{n} < +\infty \tag{4.16}$$

and

$$\rho\Big(\tilde\gamma_i^{(n)},\, T_n(s)\Big) = \rho\Big(\tilde\gamma_i^{(n)}, \gamma_i\Big) < \frac{1}{2^n} \longrightarrow 0$$

as $n \to \infty$. This together with (4.15) means that $\tilde\gamma_i^{(n)} \to T(s)$. Therefore by means of (I) and (4.16)

$$h(s,\Gamma) \le h\big(s, T(s)\big) = \lim_{n\to\infty} h\Big(s, \tilde\gamma_i^{(n)}\Big) \le h(s,\Gamma)\,,$$

and (4.10) is proved.

Now do not assume that (4.11) holds. Since the set

$$D = \big\{s \in S \colon h(s,\Gamma) = +\infty\big\}$$

is measurable, also the set $S_0 = S - D$ is measurable and according to the previous part of the proof there exists a measurable mapping $T_0 \colon S_0 \to \Gamma$ such that

$$h\big(s, T_0(s)\big) = h(s,\Gamma)$$

for all $s \in S_0$. Now it is sufficient to fix any $\gamma \in \Gamma$ and put

$$T(s) = \begin{cases} T_0(s)\,, & s \in S_0\,,\\ \gamma\,, & s \in D\,. \end{cases}$$

□

Proof of Theorem 2.1. According to the assumption (AI) the function $L(x_{(n_1,\dots,n_q)},\Omega)$ is measurable and therefore the sets

$$B_1 = \{x_{(n_1,\dots,n_q)} \in D_{n_1,\dots,n_q} : L(x_{(n_1,\dots,n_q)},\Omega) = 0\}, \tag{4.17}$$

$$B_2 = \{x_{(n_1,\dots,n_q)} \in D_{n_1,\dots,n_q} : L(x_{(n_1,\dots,n_q)},\Omega) > 0\} \tag{4.18}$$

are measurable. As before, let $\mathcal{H}$ denote the boundary of the set. The sets

$$\Gamma_k = \Big\{\theta \in \Theta \colon \rho(\theta,\mathcal{H}(\Theta)) \ge \frac{1}{k}\,,\ \|\theta\| \le k\Big\}$$

are closed, bounded and therefore compact (if $\mathcal{H}(\Theta) = \emptyset$, then only the second inequality is considered here). Obviously

$$\Gamma_k \subset \Gamma_{k+1}\,, \qquad \bigcup_{k=1}^{\infty} \Gamma_k = \Theta\,.$$

This means that

$$\Omega = \Theta \cap C = \bigcup_{k=1}^{\infty} (\Gamma_k \cap C)\,, \qquad (\Gamma_k \cap C) \subset (\Gamma_{k+1} \cap C)\,, \tag{4.19}$$

where the sets $\Gamma_k \cap C$ are compact. Let $x_{(n_1,\dots,n_q)} \in B_2$. To obtain a contradiction, suppose that for all $n = 1, 2, \dots$

$$L(x_{(n_1,\dots,n_q)},\Omega) > L(x_{(n_1,\dots,n_q)},\Gamma_k \cap C)\,. \tag{4.20}$$

Then for all k there exist $\theta_k \in \Omega - (\Gamma_k \cap C) = \Omega - \Gamma_k$ such that

$$L(x_{(n_1,\dots,n_q)},\theta_k) > L(x_{(n_1,\dots,n_q)},\Gamma_k \cap C)\,. \tag{4.21}$$

However, if $\tilde\theta \in \Theta$, then according to (AV)

$$L(x_{(n_1,\dots,n_q)},\tilde\theta) = G_{n_1,\dots,n_q}(x_{(n_1,\dots,n_q)}) - nK(\hat\theta,\tilde\theta,\hat p)\,, \tag{4.22}$$

where (cf. (4.6))

$$n = n_1 + \dots + n_q\,, \quad K(\hat\theta,\tilde\theta,\hat p) = \sum_{j=1}^{q} \hat p_j K(\hat\theta_{n_j},\Pi_j(\tilde\theta))\,, \quad \hat p_j = n_j/n\,, \tag{4.23}$$

and $\hat\theta_{n_j} : A_{n_j} \to \Xi$ is the mapping, postulated in (AV) and possessing the properties (2.9)–(2.10). But (4.22) together with Lemma 4.1 (II) imply that

$$\log L(x_{(n_1,\dots,n_q)},\theta_k) \longrightarrow -\infty$$

as $k \to \infty$. Hence the right-hand side of (4.21) tends to 0 as $k \to \infty$ and (4.19) implies that $L(x_{(n_1,\dots,n_q)}, \Omega) = 0$. Thus $x_{(n_1,\dots,n_q)} \notin B_2$, which is a contradiction and (4.20) holds with the equality sign for all k sufficiently large. Therefore

$$B_2 = \bigcup_{k=1}^{\infty} H_k\,, \qquad H_k = \{x_{(n_1,\dots,n_q)} \in D_{n_1,\dots,n_q} : L(x_{(n_1,\dots,n_q)}, \Omega) \\ = L(x_{(n_1,\dots,n_q)}, \Gamma_k \cap C)\}\,,$$

where owing to (AI) the sets H_k are measurable. Since $H_k \subset H_{k+1}$,

$$B_2 = \bigcup_{k=1}^{\infty} B_{2,k}\,, \qquad B_{2,k} = H_k - H_{k-1}\,, \quad H_0 = \emptyset\,. \tag{4.24}$$

Let k be an arbitrary but fixed positive integer. The validity of (AI), (AIII) and the fact that a set open in the Euclidean space is a countable union of closed sets mean that the assumptions of the Lemma 4.2 are fulfilled with

$$S = B_{2,k}\,, \quad \Gamma = \Gamma_k \cap C\,, \quad h(x_{(n_1,\dots,n_q)}, \tilde\theta) = nK(\hat\theta, \tilde\theta, \hat p)\,.$$

Hence according to the mentioned lemma there exist measurable mappings

$$\tilde\theta^{(k)}_{n_1,\dots,n_q} : B_{2,k} \to \Gamma_k \cap C$$

such that on $B_{2,k}$ (cf. (4.23), (2.17))

$$nK\Big(\hat\theta(x_{(n_1,\dots,n_q)})\,,\ \tilde\theta^{(k)}_{n_1,\dots,n_q}(x_{(n_1,\dots,n_q)})\,,\ \hat p\Big) = nK\Big(\hat\theta(x_{(n_1,\dots,n_q)})\,,\ \Gamma_k \cap C,\ \hat p\Big)\,, \tag{4.25}$$

where $\hat\theta(x_{(n_1,\dots,n_q)}) = \big(\hat\theta_{n_1}[x(1,n_1)], \dots, \hat\theta_{n_q}[x(q,n_q)]\big)$. Substituting (4.25) into the equality (4.22) yields that on $B_{2,k}$

$$L(x_{(n_1,\dots,n_q)}, \Omega) = L(x_{(n_1,\dots,n_q)}, \Gamma_k \cap C) = L\Big(x_{(n_1,\dots,n_q)},\ \tilde\theta^{(k)}_{n_1,\dots,n_q}(x_{(n_1,\dots,n_q)})\Big)\,.$$

Hence the validity of (4.24), (4.17) and (4.18) means that it is sufficient to put

$$\tilde\theta_{n_1,\dots,n_q}(x_{(n_1,\dots,n_q)}) = \begin{cases} \tilde\theta^{(k)}_{n_1,\dots,n_q}(x_{(n_1,\dots,n_q)})\,, & x_{(n_1,\dots,n_q)} \in B_{2,k}\,, \\ \theta_0\,, & x_{(n_1,\dots,n_q)} \in B_1\,, \end{cases} \tag{4.26}$$

where θ_0 is any fixed point from Ω. □

The considerations in rest of the paper will be carried out in the multisample setting and based on the notation (2.18) and (2.19).

Lemma 4.3. *Suppose that both the assumptions* (AI)–(AV) *and the equalities* (2.22) *hold. Given* $\theta \in \Theta$ *and a positive real number* c *there exists a compact set* $\Gamma_c \subset \Theta$ *such that*

$$\limsup_{u\to\infty} \frac{1}{n_u} \log L\big(x^{(u)}, \Theta - \Gamma_c\big) \leq -c \tag{4.27}$$

a.e., P_θ.

P r o o f. (a) Let $q = 1$. Since $\theta \in \Xi$, there exists a positive number δ such that

$$\Gamma = \big\{\gamma^* \in \mathbb{R}^m \colon \gamma^* - \theta\| \leq \delta\big\} \subset \Xi\,.$$

This according to (AIV) means that for any fixed positive real number c_1 the set

$$\Gamma_0 = \{\tilde\gamma \in \Xi\colon \text{there exists } \gamma^* \in \Gamma \text{ such that } K(\gamma^*, \tilde\gamma) \leq c_1\big\}$$

is compact. Put

$$\tilde A_n = \big\{(x_1, \dots, x_n)\colon L(x_1, \dots, x_n, \theta) > 0\big\}\,,$$

where A_n is the set described in (AV). Since (2.8) and (2.10) imply that $\chi_{A_n} \to 1$ a.e., $\overline{P}_\theta^{\infty}$, also $\chi_{\tilde A_n} \to 1$ a.e., $\overline{P}_\theta^{\infty}$. But if $(x_1, \dots, x_n) \in \tilde A_n$, then according to (2.9)

$$\log \frac{L(x_1, \dots, x_n, \Xi - \Gamma_0)}{L(x_1, \dots, x_n, \theta)} = n\Big[K\big(\hat\gamma_n(x_1, \dots, x_n), \theta\big) - K\big(\hat\gamma_n(x_1, \dots, x_n)\,, \Xi - \Gamma_0\big)\Big]$$

and if $\hat\gamma_n \in \Gamma$, then

$$\frac{1}{n} \log \frac{L(x_1, \dots, x_n, \Xi - \Gamma_0)}{L(x_1, \dots, x_n, \theta)} \leq K(\hat\gamma_n, \theta) - c_1\,,$$

where for the sake of brevity the notation $\hat\gamma_n = \hat\gamma(x_1, \dots, x_n)$ is used. But (AII) implies that the integral

$$c_0 = \int \log f(x, \theta)\, \overline{P}_\theta(x)$$

is a real number, which together with the law of large numbers means that

$$\frac{1}{n} \log L(x_1, \dots, x_n, \theta) \longrightarrow c_0$$

a.e., $\overline{P}_\theta^{\infty}$. By (2.9) therefore $K(\hat\gamma_n, \theta) < +\infty$ a.e., for all n sufficiently large, and (2.10) and (AIII) yield that

$$K(\hat\gamma_n, \theta) \longrightarrow K(\theta, \theta) = 0$$

a.e., $\overline{P}_\theta^\infty$. Hence a.e., $\overline{P}_\theta^\infty$

$$\limsup_{n\to\infty} \frac{1}{n} \log \frac{L(x_1,\dots,x_n,\Xi-\Gamma_0)}{L(x_1,\dots,x_n,\theta)} \le -c_1$$

and putting $c_1 = c + c_0$, $\Gamma_c = \Gamma_0$ one obtains that the Lemma holds for $q = 1$.

(b) Before proceeding to the final part of the proof, it is necessary to show that given $\gamma \in \Xi$ there exists a real number d, such that

$$\limsup_{n\to\infty} \frac{1}{n} \log L(x_1,\dots,x_n,\Xi) < d \tag{4.28}$$

a.e., $\overline{P}_\gamma^\infty$.

According to the previous part of the proof there exists a compact set $\tilde{\Gamma} \subset \Xi$ such that (4.27) holds a.e., $\overline{P}_\gamma^\infty$ with $c = 1$. If $\gamma^* \in \tilde{\Gamma}$, then from (AII) one obtains the existence of the neighbourhood $V(\gamma^*,\delta^*)$ such that the function $\log L\big(x, V(\gamma^*,\delta^*)\big)$ is $\overline{P}_\gamma$ integrable (with $-\infty$ as a possible value of the integral). But the set $\tilde{\Gamma}$ is compact and therefore there exist finitely many sets $V(\gamma_k,\delta_k)$, $k = 1,\dots,r$, such that

$$\tilde{\Gamma} \subset \bigcup_{k=1}^{r} V(\gamma_k,\delta_k), \quad d_k = \int \log L\big(x, V(\gamma_k,\delta_k)\big)\, \overline{P}_\gamma(x) < +\infty .$$

Thus almost surely

$$\begin{aligned}\limsup_{n\to\infty} \frac{1}{n} \log L(x_1,\dots,x_n,\Xi) \le \limsup_{n\to\infty} \max\Bigg\{ &\frac{1}{n} \log L(x_1,\dots,x_n,\Xi-\tilde{\Gamma}), \\ &\frac{1}{n} \log\Bigg(\prod_{i=1}^{n} L\big(x_i, V(\gamma_k,\Delta_k)\big)\Bigg),\ k = 1,\dots,r \Bigg\} \\ \le \max\{-1, d_1,\dots,d_k\} < \max\{-1, d_1,\dots,d_k\} + 1 = d .\end{aligned}$$

(c) Suppose finally that $q > 1$ and $\theta = (\theta_1,\dots,\theta_q)$. According to (b) of this proof there exists a number d such that the inequality (4.28) holds a.e., $\overline{P}_{\theta_j}^\infty$ for $j = 1,\dots,q$. Now if $c_1,\dots,c_q$ are positive real numbers, then according to (a) of this proof there exist compact sets $\Gamma_j \subset \Xi$ such that for $j = 1,\dots,q$

$$\limsup_{n\to\infty} \frac{1}{n} \log L(x_1,\dots,x_n,\Xi-\Gamma_j) < -c_j \tag{4.29}$$

$\overline{P}_{\theta_j}^\infty$ a.e. Obviously, $\Gamma = \Gamma_1 \times \dots \times \Gamma_q$ is a compact subset of Θ and

$$\frac{1}{n_u} \log L\big(x^{(u)}, \Theta - \Gamma\big) \le \max_{j=1,\dots,q} \frac{1}{n_u} \log L\big(x^{(u)}, D_j\big) , \tag{4.30}$$

where $D_j = \Xi \times \cdots \times \Xi \times (\Xi - \Gamma_j) \times \Xi \times \cdots \times \Xi$. Hence if $\sum^*$ denotes the sum over the indices $1 \le i \le q$, $i \ne j$, then in accordance with (4.28) and (2.19)

$$\begin{aligned}
&\limsup_{u\to\infty} \frac{1}{n_u} \log L\big(x^{(u)}, D_j\big) \\
&\le \sum\nolimits^* \limsup_{u\to\infty} \frac{1}{n_u} \log L\big(x\big(i, n_u^{(i)}\big), \Xi\big) + \limsup_{u\to\infty} \frac{1}{n_u} \log L\big(x\big(j, n_u^{(j)}\big), \Xi - \Gamma_j\big) \\
&\overset{(4.29)}{\le} \sum\nolimits^* p_i d - p_j c_j\,.
\end{aligned}$$

Thus choosing the c_i's sufficiently large one obtains by (4.30) that (4.27) holds.
□

P r o o f o f T h e o r e m 2.2. (I) Let $\gamma \in \Xi$. As it has already been remarked, the integral (cf. (4.6))

$$K\big(\Pi_j(\theta), \gamma\big) = \int \log \frac{f\big(x, \Pi_j(\theta)\big)}{f(x,\gamma)}\, f\big(x, \Pi_j(\theta)\big)\, d\nu(x)$$

exists with $+\infty$ as a possible value. According to (A2) the function $\log f\big(x, \Pi_j(\theta)\big)$ is $\overline{P}_{\Pi_j(\theta)}$ integrable, and therefore the integral

$$I_j(\gamma) = \int \big(\log f(x,\gamma)\big)\, f\big(x, \Pi_j(\theta)\big)\, d\nu(x) \tag{4.31}$$

exists with $-\infty$ as a possible value and

$$K\big(\Pi_j(\theta), \gamma\big) = I_j\big(\Pi_j(\theta)\big) - I_j(\gamma)\,.$$

Now let

$$I(\theta^*) = \sum_{j=1}^{q} p_j I_j\big(\Pi_j(\theta^*)\big)\,, \tag{4.32}$$

where p is the limiting vector from (2.22). Then (cf. (2.16))

$$K(\theta, \theta^*, p) = \sum_{j=1}^{q} p_j K(\theta_j, \theta_j^*) = I(\theta) - I(\theta^*) \tag{4.33}$$

and by means of the notation

$$I(\Omega) = \sup\big\{I(\theta^*) : \theta^* \in \Omega\big\} \tag{4.34}$$

the quantity (2.17) can be written as

$$K(\theta, \Omega, p) = I(\theta) - I(\Omega)\,. \tag{4.35}$$

Since $K(\theta,\eta,p) = K(\theta,\Omega,p)$ is a real number and $I(\theta)$ is a real number, the relations (4.33) and (4.35) imply that η is the unique parameter from Ω fulfilling the equality

$$I(\eta) = I(\Omega)\,. \tag{4.36}$$

Since $I(\eta)$ is a real number, there is a positive real number c such that $-c < I(\eta)$. Further, according to Lemma 4.3 there exists a compact set $\Gamma = \Gamma_c \subset \Theta$ such that (4.27) holds a.e. P_θ. Hence a.e.

$$\begin{aligned}\limsup_{u\to\infty} \frac{1}{n_u} \log \frac{L(x^{(u)}, \Omega - \Gamma)}{L(x^{(u)}, \Omega)} &\le \limsup_{u\to\infty} \frac{1}{n_u} \log \frac{L(x^{(u)}, \Theta - \Gamma)}{L(x^{(u)}, \eta)} \\ &\le -c - I(\eta) < 0\,,\end{aligned}$$

which means that with probability P_θ equals 1

$$L\big(x^{(u)}, \Omega\big) = L\big(x^{(u)}, \Omega \cap \Gamma\big)\,, \qquad \tilde\theta_u\big(x^{(u)}\big) \in \Omega \cap \Gamma \tag{4.37}$$

for all u sufficiently large. Let $\varepsilon > 0$ be a fixed number such that the set

$$U_\varepsilon = \big\{\theta^* \in \Omega \cap \Gamma \colon \rho(\theta^*, \eta) \ge \varepsilon\big\} \tag{4.38}$$

is non-empty. For $M < I(\Omega)$ put

$$\Omega(M) = \big\{\theta^* \in \Omega \cap \Gamma \colon I(\theta^*) \le M\big\}\,. \tag{4.39}$$

Let real numbers $M_k \nearrow I(\Omega)$. If

$$U_\varepsilon - \Omega(M_k) \neq \emptyset \tag{4.40}$$

for every k, then one can choose $\theta_k^* \in \Theta$ such that

$$M_k < I(\theta_k^*) \le I(\Omega)\,, \qquad \theta_k^* \in U_\varepsilon\,. \tag{4.41}$$

But the set $\Omega \cap \Gamma = C \cap \Gamma$ is compact and without the loss of generality we may assume that

$$\theta_k^* \longrightarrow \theta^* \in \Omega \cap \Gamma\,. \tag{4.42}$$

Since (4.41) holds, $I(\theta_k^*) \to I(\Omega)$ and the validity of (4.33), (4.35) implies that

$$K(\theta, \theta_k^*, p) \longrightarrow I(\theta) - I(\Omega) = K(\theta, \Omega, p)\,.$$

Further, the finiteness of $K(\theta, \theta_k^*, p)$ and (AIII) yield that

$$K(\theta, \theta_k^*, p) \longrightarrow K(\theta, \theta^*, p)\,.$$

Thus $K(\theta, \theta^*, p) = K(\theta, \Omega, p)$ and by the second relation in (4.41) also $\theta^* \neq \eta$, which contradicts the uniqueness of η. Since the contradiction has been deduced from (4.40),

$$U_\varepsilon \subset \Omega(M) \tag{4.43}$$

if $M < I(\Omega)$ is sufficiently large.

Let $\gamma^* \in \Xi$. Since (AII) holds, the integral

$$I_j(\gamma^*, \delta^*) = \int \log L\big(x, V(\gamma^*, \delta^*)\big)\, d\overline{P}_{\Pi_j(\theta)}(x)$$

exists with $-\infty$ as a possible value if $\delta^* > 0$ is sufficiently small, and by means of the monotone convergence theorem

$$\lim_{\delta \to 0+} I_j(\gamma^*, \delta) = I_j(\gamma^*)\,,$$

where I_j is defined in (4.31). Hence if $\theta^* \in U_\varepsilon$, then in accordance with (4.32)

$$\lim_{\delta \to 0+} \sum_{j=1}^{q} p_j\, I_j\big(\Pi_j(\theta^*), \delta\big) = \sum_{j=1}^{q} p_j\, I_j\big(\Pi_j(\theta^*)\big) = I(\theta^*)$$

and (4.43) together with the inequality $M < I(\Omega)$ yield the existence of a positive real number $\delta^* = \delta^*(\theta^*)$ such that

$$\sum_{j=1}^{q} p_j\, I_j\big(\Pi_j(\theta^*), \delta^*\big) < I(\Omega)\,. \tag{4.44}$$

Since the set $\Omega \cap \Gamma = C \cap \Gamma$ is compact, also U_ε is compact and the validity of (4.44), (4.36) implies the existence of finitely many open sets

$$V_i = V\big(\Pi_1(\theta_i^*), \delta_i^*\big) \times \cdots \times V\big(\Pi_q(\theta_i^*), \delta_i^*\big)\,, \qquad i = 1, \ldots, r$$

such that

$$\max\left\{ \sum_{j=1}^{q} p_j\, I_j\big(\Pi_j(\theta_i^*), \delta_i^*\big) :\ i = 1, \ldots, r \right\} < I(\Omega)\,,$$

$$U_\varepsilon \subset \bigcup_{i=1}^{r} V_i \subset \Theta - \eta\,.$$

Hence P_θ a.e.,

$$\begin{aligned}\limsup_{u \to \infty} \frac{1}{n_u} \log \frac{L(x^{(u)}, U_\varepsilon)}{L(x^{(u)}, \Omega)} &\le \max_{i=1,\ldots,r}\ \limsup_{u \to \infty} \frac{1}{n_u} \log \frac{L(x^{(u)}, V_i)}{L(x^{(u)}, \eta)} \\ &= \max_{i=1,\ldots,r} \left[\sum_{j=1}^{q} p_j\, I_j\big(\Pi_j(\theta_i^*), \delta_i^*\big) - I(\eta) \right] < 0\,,\end{aligned} \tag{4.45}$$

because (4.36) and the law of large numbers hold. Validity of (4.45) obviously implies that a.e., P_θ for all u sufficiently large

$$L\big(x^{(u)}, U_\varepsilon\big) < L\big(x^{(u)}, \Omega\big)\,.$$

But this, (4.37) and (4.38) yield that with probability P_θ equal 1 the inequality

$$\|\tilde{\theta} - \theta\| < \varepsilon$$

holds for all u sufficiently large.

(II) Since $\theta \in \Omega$, in (2.25)

$$\eta(\theta, p) = \theta$$

and (II) follows from (I). □

LEMMA 4.4. *Suppose that* (2.30), (2.22) *and the assumptions* (AI)–(AV) *hold. If* $\theta \in \Omega_1 - \Omega_0$, $K(\theta, \Omega_0, p) < +\infty$ *and* T_u *denote the statistics* (2.29), *then*

$$\liminf_{u\to\infty} \frac{T_u}{n_u} > 0 \tag{4.46}$$

a.e., P_θ.

P r o o f . The proof will be based on the relations (4.31)–(4.35) from the previous proof, where Ω is replaced with Ω_0.

To obtain a contradiction, suppose that

$$K(\theta, \Omega_0, p) = 0\,. \tag{4.47}$$

This implies that there exist parameters $\{\theta_k^*\}_{k=1}^{\infty}$ from Ω_0 such that

$$K(\theta, \theta_k^*, p) \longrightarrow K(\theta, \Omega_0, p) = 0\,, \tag{4.48}$$

and by Lemma 4.1 there exists a compact set $\tilde{\Gamma} \subset \Theta$ such that $\theta_k^* \in \tilde{\Gamma}$ for all k. Thus for some subsequence $\theta_{k_r}^* \to \theta^* \in \Omega_0$ which together with (4.48) and (AIII) means that $K(\theta, \theta^*, p) = 0$. Hence $\overline{P}_{\Pi_j(\theta)} = \overline{P}_{\Pi_j(\theta^*)}$ for all j and (AI) yields that $\theta = \theta^* \in \Omega_0$. This is a contradiction and therefore

$$K(\theta, \Omega_0, p) > 0\,. \tag{4.49}$$

Further, (4.31), (4.32) and the law of large numbers imply that

$$\frac{1}{n_u} \log L\big(x^{(u)}, \theta\big) \longrightarrow I(\theta)$$

a.e., P_θ as $u \to \infty$. Since $I(\theta)$ is a real number, putting (cf. (2.23))

$$\tilde{D}_u = \big\{x^{(u)} \in D_u : L(x^{(u)}, \theta) > 0\big\} \tag{4.50}$$

and utilizing (AV) one easily obtains that a.e., P_θ

$$\lim_{u\to\infty} \chi_{\tilde{D}_u}\big(x^{(u)}\big) = 1\,, \tag{4.51}$$

where χ denotes the indicator function of the concerned set.

Now we are going to use the notation (cf. (2.21))

$$p_u = \left(p_u^{(1)}, \ldots, p_u^{(q)}\right)$$

and

$$\hat{\theta}_{(u)} = \hat{\theta}_{(u)}(x^{(u)}) = \left(\hat{\theta}_{n_u^{(1)}}\left(x(1, n_u^{(1)})\right), \ldots, \hat{\theta}_{n_u^{(q)}}\left(x(q, n_u^{(q)})\right)\right),$$

where the mapping $\hat{\theta}_n \colon A_n \to \Xi$ is described in (AV) (it is denoted there by the symbol $\hat{\gamma}_n$). Since $\theta \in \Omega_1$ and (AV) is fulfilled, on the set $\tilde{D}_u$ the relations

$$\frac{T_u}{n_u} \geq \frac{2}{n_u} \log \frac{L(x^{(u)}, \theta)}{L(x^{(u)}, \Omega_0)} = 2\left[K(\hat{\theta}_{(u)}, \Omega_0, p_u) - K(\hat{\theta}_{(u)}, \theta, p_u)\right] \tag{4.52}$$

hold with $K(\hat{\theta}_{(u)}, \theta, p_u) < +\infty$. Since $\hat{\theta}_{(u)} \to \theta$ a.e., P_θ, from (AIII) and (4.51)

$$K\left(\hat{\theta}_{(u)}, \theta, p_u\right) \longrightarrow K(\theta, \theta, p) = 0$$

a.e., P_θ, which together with (4.51) and (4.52) leads to

$$\liminf_{u\to\infty} \frac{T_u}{n_u} \geq \liminf_{u\to\infty}\ 2K\left(\hat{\theta}_{(u)}, \Omega_0, p_u\right) \tag{4.53}$$

a.e., P_θ. To prove (4.46) it is sufficient to consider the case when

$$+\infty > \liminf_{u\to\infty} K\left(\hat{\theta}_{(u)}, \Omega_0, p_u\right) = \lim_{v\to\infty} K\left(\hat{\theta}_{(u_v)}, \Omega_0, p_{u_v}\right) \tag{4.54}$$

and $\hat{\theta}_{(u_v)} \longrightarrow \theta$.

Since the quantities behind the limit symbol on the right-hand side of (4.54) are real numbers, one may choose $\eta_v \in \Omega_0$ such that

$$\left|K\left(\hat{\theta}_{(u_v)}, \Omega_0, p_{u_v}\right) - K\left(\hat{\theta}_{(u_v)}, \eta_{u_v}, p_{u_v}\right)\right| < \frac{1}{v}, \qquad v = 1, 2, \ldots$$

According to Lemma 4.1 (I) for some subsequence $\eta_{u_{v_r}}$ the convergence

$$\eta_{u_{v_r}} \longrightarrow \eta^* \in \Theta$$

holds. Hence $\eta^* \in \Omega_0$ and from (AIII)

$$\begin{aligned} K(\theta, \Omega_0, p) \leq K(\theta, \eta^*, p) &= \lim_{r\to\infty} K\left(\hat{\theta}_{(u_{v_r})}, \eta_{u_{v_r}}, p_{u_{v_r}}\right) \\ &= \liminf_{u\to\infty}\ 2K\left(\hat{\theta}_{(u)}, \Omega_0, p_u\right) \end{aligned}$$

and (4.49), (4.53) imply (4.46). □

P r o o f o f T h e o r e m 2.3. Let

$$B_u = \left\{x^{(u)} \colon \frac{T_u(x^{(u)})}{n_u} > \frac{c_u}{n_u}\right\}.$$

Obviously

$$P_\theta\Big(\varphi_u(x^{(u)}) = 1\Big) = P_\theta\Big(T_u(x^{(u)}) > c_u\Big) = P_\theta(B_u) \ge P_\theta\Big(\bigcap_{r=u}^{\infty} B_r\Big) . \qquad (4.55)$$

Since for $\theta^* \in \Omega_0$

$$T_u \le 2\log\frac{L(x^{(u)},\Theta)}{L(x^{(u)},\theta^*)} ,$$

by (2.31)–(2.33)

$$\limsup_{u\to+\infty} c_u < +\infty ,$$

and by means of (4.55)

$$\begin{aligned}\liminf_{u\to\infty} P_\theta(\varphi_u = 1) &\ge \lim_{u\to\infty} P_\theta\left(\bigcap_{r=u}^{\infty} B_r\right) = P_\theta\left(\bigcup_{u=1}^{\infty}\bigcap_{r=u}^{\infty} B_r\right)\\ &\ge P_\theta\left(\liminf_{u\to\infty}\frac{T_u}{n_u} > 0\right) = 1 ,\end{aligned}$$

where the last equality follows from Lemma 4.4. □

P r o o f o f T h e o r e m s 3.1–3.3. These theorems will easily follow from the Theorems 2.1–2.3, provided that one knows that the imposed assumptions hold also in the case of the exponential distributions with unknown lower bounds. This validity is the topic of the following lemma.

LEMMA 4.5. *Suppose that the class of probabilities* $\{\overline{P}_\gamma : \gamma \in \Xi\}$ *is defined by means of the parameter set* (3.1) *and the densities* (3.2).

(I) *Let* $\gamma = (\mu,\sigma)'$, $\gamma^* = (\mu^*,\sigma^*)'$ *belong to* (3.1). *Then* $K(\gamma,\gamma^*) < +\infty$ *if and only if* $\mu \ge \mu^*$.

(II) *The assumptions* (AI)–(AV) *hold. In* (AV) *in this case* A_n *are the sets* (3.3).

(III) *If* $\gamma \in \Xi$, *then* (2.31) *holds.*

P r o o f . (I) This assertion follows from the fact that (2.6) is now determined by the formula

$$K(\gamma,\gamma^*) = \begin{cases} \log\frac{\sigma^*}{\sigma} + \frac{\mu+\sigma-(\mu^*+\sigma^*)}{\sigma^*} , & \mu \ge \mu^* ,\\ +\infty , & \mu < \mu^* . \end{cases} \qquad (4.56)$$

(II) As it has already been established in the proof of Theorem 4.1 of [12], the measurability postulated in (AI) follows from Lemma 2.4 of [11].

From (4.56) one can easily prove the validity of (AII)–(AIV). Further, for $(x_1,\ldots,x_n)\in\mathbb{R}^n$ put

$$\hat\mu=\min\{x_1,\ldots,x_n\}\,,\quad \hat\sigma=\overline{x}-\hat\mu\,,\quad \hat\gamma(x_1,\ldots,x_n)=(\hat\mu,\hat\sigma)'\,, \tag{4.57}$$

where $\overline{x}=(x_1+\ldots x_n)/n$. Let A_n be the set (3.3). Making use of (4.56) and (3.2) it is easy to see that on A_n

$$\log L(x_1,\ldots,x_n,\gamma)=-n\log(\overline{x}-\hat\mu)-n-nK\big(\hat\gamma(x_1,\ldots,x_n),\gamma\big)\,.$$

If the sampling is made from the exponential $\overline{P}_\gamma$ distribution with $\gamma=(\mu,\sigma)'$, then obviously

$$\hat\mu=\mu+o_P\big(n^{-1/2}\big) \tag{4.58}$$

and $\hat\mu\to\mu$ almost surely, because $\hat\mu$ depends on n in a non-increasing way. Since the mean and the variance of the exponential $\overline{P}_\gamma$ distribution are

$$E_\gamma(x)=\mu+\sigma\,,\quad \mathrm{Var}_\gamma(x)=\sigma^2\,, \tag{4.59}$$

by the law of large numbers a.s. $\overline{x}\to\mu+\sigma$. Thus it is easy to see that also the relations (2.10) hold.

(III) On A_n the equality

$$\log\frac{L(x_1,\ldots,x_n,\Xi)}{L(x_1,\ldots,x_n,\gamma)}=n\left[\frac{\overline{x}-(\mu+\sigma)}{\sigma}-\log\frac{\overline{x}-\hat\mu}{\sigma}\right] \tag{4.60}$$

holds. Further, by the central limit theorem and (4.59)

$$\sqrt{n}\big[\overline{x}-(\mu+\sigma)\big]\longrightarrow N_1(0,\sigma^2) \tag{4.61}$$

in distribution. This together with (4.58) and the Taylor expansion of (4.60) for $\overline{x}$ about $\mu+\sigma$ and $\hat\mu$ about μ leads to

$$\begin{aligned} &n\left[\frac{\overline{x}-(\mu+\sigma)}{\sigma}-\log\frac{\overline{x}-\hat\mu}{\sigma}\right]\\ &=n\,\frac{\hat\mu-\mu}{\sigma}+\frac{n}{2}\left[\frac{\big(\overline{x}-(\mu+\sigma)\big)^2}{\sigma^2}-\frac{2}{\sigma^2}\big(\overline{x}-(\mu+\sigma)\big)(\hat\mu-\mu)+\frac{(\hat\mu-\mu)^2}{\sigma^2}\right]+o_P(1)\\ &=n\,\frac{\hat\mu-\mu}{\sigma}+\frac{n}{2}\frac{\big(\overline{x}-(\mu+\sigma)\big)^2}{\sigma^2}+o_P(1)\,. \end{aligned} \tag{4.62}$$

Since

$$\big(\overline{x}-(\mu+\sigma)\big)^2=(\overline{x}-\hat\mu-\sigma)^2+2(\overline{x}-\hat\mu-\sigma)(\hat\mu-\mu)+(\hat\mu-\mu)^2\,,$$

by (4.58) and (4.61)

$$\big(\overline{x}-(\mu+\sigma)\big)^2=(\hat\sigma-\sigma)^2+o_P(n^{-1})\,.$$

Hence by (4.60)–(4.62)

$$n\,\frac{(\hat{\sigma}-\sigma)^2}{\sigma^2} \longrightarrow \chi_1^2 \tag{4.63}$$

and

$$\log\frac{L(x_1,\dots,x_n,\Xi)}{L(x_1,\dots,x_n,\gamma)} = n\,\frac{\hat{\mu}-\mu}{\sigma} + \frac{n}{2}\,\frac{(\hat{\sigma}-\sigma)^2}{\sigma^2} + o_P(1)\,. \tag{4.64}$$

But for every $\varepsilon > 0$

$$P\Big[n\,\frac{\hat{\mu}-\mu}{\sigma} > \varepsilon\Big] = \exp(-\varepsilon) \tag{4.65}$$

and since $\hat{\mu}$, $\hat{\sigma}$ are independent, from (4.63)–(4.65) we obtain the validity of (2.31). □

REFERENCES

[1] BAHADUR, R. R.: *Some Limit Theorems in Statistics*, SIAM, Philadelphia, 1971.

[2] BAHADUR, R. R.: *An Optimal Property of the Likelihood Ratio Statistic*, in: Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability (L. Le Cam and J. Neyman, eds.), 1965, pp. 13–26.

[3] BERK, R. H.: *Consistency and asymptotic normality of* MLE*'s for exponential models*, Ann. Math. Statist. **43** (1972), 193–204.

[4] HUBER, P. J.: *The Behaviour of Maximum Likelihood Estimates Under Nonstandard Conditions*, in: Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability (L. Le Cam and J. Neyman, eds.), 1965, pp. 221–233.

[5] IBRAGIMOV, I. A.—CHASMINSKIJ, R. Z.: *Asymptotic Theory of Estimation*, Nauka, Moscow, 1979. (In Russian)

[6] KUBÁČEK, L.: *Foundations of Estimation Theory*, Elsevier, Amsterdam, 1988.

[7] Le CAM, L.: *On some asymptotic properties of maximum likelihood estimates and related Bayes' estimates*, Univ. California Publ. Statist. **1** (1953), 277–329.

[8] MÄKELÄINEN, T.—SCHMIDT, K.—STYAN, G. P. H.: *On the existence and uniqueness of the maximum likelihood estimate of a vector-valued parameter in fixed-size samples*, Ann. Statist. **9** (1981), 758–767.

[9] MIAO, W.—HAHN, M.: *Existence of maximum likelihood estimates for multidimensional exponential families*, Scand. J. Statist. **24** (1997), 371–386.

[10] PÁZMAN, A.: *On the uniqueness of the ML estimates in curved exponential families*, Kybernetika **22** (1986), 124–132.

[11] RUBLÍK, F.: *On optimality of the LR tests in the sense of exact slopes. Part I. General Case*, Kybernetika **25** (1989), 13–25.

[12] RUBLÍK, F.: *On optimality of the LR tests in the sense of exact slopes. Part II. Application to individual distributions*, Kybernetika **25** (1989), 117–135.

[13] RUBLÍK, F.: *On Hodges-Lehman optimality of LR tests*, Kybernetika **30** (1994), 199–210.

[14] RUBLÍK, F.: *On consistency of the* MLE, Kybernetika **31** (1995), 45–64.

[15] RUBLÍK, F.: *On d-optimality of the LR tests*, Kybernetika **31** (1995), 189–206.

[16] VAJDA, I.: *Global information in statistical experiments and consistency of likelihood based estimates and tests*, Kybernetika **34** (1998), 245–263.

[17] WALD, A.: *Note on the consistency of the maximum likelihood estimate*, Ann. Math. Statist. **20** (1949), 595–601.

Received January 4, 2001

Institute of Measurement Science
Slovak Academy of Sciences
Dúbravská cesta 9
SK–84219 Bratislava
SLOVAKIA

E-mail: umerrubl@savba.sk