Siberian Mathematical Journal, Vol.42, No. 3, 2001.
Translated from:
Sibirskiĭ Matematicheskiĭ Zhurnal, Vol. 42, No. 3, pp. 610–633.

ASYMPTOTICALLY NORMAL EXPLICIT ESTIMATION OF PARAMETERS IN THE MICHAELIS–MENTEN EQUATION

YU. YU. LINKE AND A. I. SAKHANENKO

1. Introduction

1.1. In many biochemical and physical experiments, the data of an experiment are assumed to satisfy the following two-parameter equation:

$$v(s) = \frac{Vs}{K+s}, \quad K, V, s > 0, \tag{1.1}$$

where K and V are some unknowns. For example, in study of chemical reactions, equation (1.1) is known as the Michaelis–Menten equation which describes the theoretical interrelation between the speed of reaction $v(s)$ and the amount of reagent s. In this case the parameter V has the meaning of a maximally possible speed of reaction and the parameter K is the value of s for which the speed of reaction is at twice as less than the maximal. In 1913 Michaelis and Menten published their research [1] on this equation, and now they are considered as founders of modern enzymology. The interest in study of this seemingly particular equation is aroused, first of all, by its various applications in practice of natural sciences. By way of example, we mention the article [2] in which the systems that can be described by equation (1.1) are considered in more detail.

In this article we deal with the problem of estimation of the unknown parameters $K > 0$ and $V > 0$ for the regressive Michaelis–Menten model

$$v_i = \frac{Vs_i}{K+s_i} + \sigma_i\xi_i, \quad i = 1, \dots, N, \tag{1.2}$$

when in result of a series of N trials, $N \to \infty$, we observe a sequence of independent random variables v_i, $i = 1, \dots, N$. The numbers $s_i > 0$ in (1.2) are assumed to be known. Concerning the measurement errors $\sigma_i\xi_i$, we assume that ξ_i are independent identically distributed random variables satisfying the following natural assumption:

$$\mathbf{E}\xi_i = 0, \quad \mathbf{D}\xi_i = 1, \quad i = 1, \dots, N.$$

The values of the random variables ξ_i and, in general, the behavior of the variances $\sigma_i^2 = \mathbf{D}X_i > 0$ are assumed unknown.

This problem is a particular instance of a nonlinear regression problem. However, solution of nonlinear regression problems faces a row of essential difficulties.

Date: July 24, 2000. Revision submitted November 30, 2000.

For example, finding estimators by the least squares method involves various iteration procedures, which raises a number of problems related to the choice of an initial approximation and the properties of the so-constructed estimators.

In this article, which is an immediate continuation of the preceding two articles [3, 4] by the authors, the problem of estimation of the parameters K and V in the Michaelis–Menten equation is solved by some new method that enables us to obtain asymptotically normal explicit estimators of the unknown parameters in a rather simple way for the model (1.2) as well as for more general linear-fractional regression problems; moreover, the estimators often have a minimal asymptotic covariance matrix.

It turns out that under very general assumptions we can choose numbers c_{Ki} and c_{Vi} so that the (rather simple) statistics

$$K^* = -\sum_{i=1}^N c_{Ki} v_i \left(\sum_{i=1}^N c_{Ki} v_i / s_i\right)^{-1}, \quad V^* = \sum_{i=1}^N c_{Vi}(1 + K^*/s_i) v_i \tag{1.3}$$

are asymptotically normal estimators of the parameters K and V. In the case when we have some information on the behavior of the variances $\{\sigma_i^2\}$ we can choose functions $\{\gamma_{Ki}(K, V)\}$ and $\{\gamma_{Vi}(K, V)\}$ so that for

$$\gamma^*_{Ki} = \gamma_{Ki}(K^*, V^*), \quad \gamma^*_{Vi} = \gamma_{Vi}(K^*, V^*)$$

the "improved" estimators

$$K^{**} = -\sum_{i=1}^N \gamma^*_{Ki} v_i \left(\sum_{i=1}^N \gamma^*_{Ki} v_i / s_i\right)^{-1}, \quad V^{**} = \sum_{i=1}^N \gamma^*_{Vi} \Big(1 + K^{**}/s_i\Big) v_i \tag{1.4}$$

are in some sense asymptotically effective.

The main goal of this article is to study the asymptotic properties of the estimators (1.3) and (1.4).

All ideas of the estimation method under consideration are exposed in detail in [4] for a multidimensional linear-fractional regression problem. The most important properties of these estimators are described in the same article under certain additional assumptions. We note that [3] contains a complete mathematically rigorous study and justification of this method for a simplest one-dimensional linear-fractional regression problem. In the present article, majority of the conditions in assertions for a two-parameter model (1.2) are obtained in a more lucid form, which was frequently impossible for a general multidimensional problem.

In particular, in the construction of the estimators K^* and V^* we recommend choosing the constants of (1.3) in the form

$$c_{Vi} = c_i \Big/ \sum_{i=1}^N c_i, \quad c_{Ki} = c_i/s_i - c_{Vi} \sum_{i=1}^N c_i/s_i, \tag{1.5}$$

where $c_i > 0$ are some constants. Similarly, we recommend restricting in (1.4) to the functions

$$\gamma_{Vi}(K, V) = \gamma_i(K, V) \Big/ \sum_{i=1}^N \gamma_i(K, V),$$

$$\gamma_{Ki}(K, V) = \gamma_i(K, V)/s_i - \gamma_{Vi}(K, V) \sum_{i=1}^N \gamma_i(K, V)/s_i, \tag{1.6}$$

where $\gamma_i(K,V) > 0$ are continuously differentiable functions. As we show in § 3, to increase the accuracy of the so-defined estimators we should choose the numbers c_i and functions $\gamma_i(K,V)$ in (1.5) and (1.6) so that the quantities $c_i(1+K/s_i)^3\sigma_i^2$ and $\gamma_i(K,V)(1+K/s_i)^3\sigma_i^2$ depend on i as less as possible.

Organization of the article. In § 2 we describe a method for constructing two classes of the mentioned estimators. In § 3 we obtain necessary conditions for their optimality and give some practical recommendations. In § 4 we formulate conditions for consistency and asymptotic normality of estimators whose derivation is given in § 5 and § 6. Concluding remarks reside in § 7.

Notation. Summation over i is always carried out from 1 to N. All limits are taken as $N \to \infty$, unless otherwise stated. The symbol $\sum$ without indices is used only for $\sum_{i=1}^N$, i.e., only if the summation is carried out over i in the indicated range. The other notations appear in due course.

2. Construction of Estimators

2.1. First of all, we note that (1.2) can be rewritten as

$$v_i = \frac{V}{1+K/s_i} + \sigma_i\xi_i. \tag{2.1}$$

It is easy to see that (2.1) is a particular instance of the equation

$$v_i = \frac{a_{0i} + a_{1i}\theta_1 + a_{2i}\theta_2}{1 + b_{1i}\theta_1 + b_{2i\theta_2}} + \sigma_i\xi_i \tag{2.2}$$

for

$$m = 2,\ \theta_1 = K,\ \theta_2 = V,\ a_{2i} = 1,\ b_{1i} = 1/s_i,\ a_{0i} = a_{1i} = b_{2i} = 0. \tag{2.3}$$

Thus, the equation (2.1) under consideration is a particular case of (2.2). Some method for estimation of the parameters in the latter equation was proposed in [4].

Following the recommendations of [4], we should choose some numbers $c_{11},\dots,c_{1N}$ and $c_{21},\dots,c_{2N}$ and seek the estimators K^* and V^* as solutions of the equations

$$K^*\sum c_{1i}v_i/s_i + \sum c_{1i}v_i = V^*\sum c_{1i}, \quad K^*\sum c_{2i}v_i/s_i + \sum c_{2i}v_i = V^*\sum c_{2i}. \tag{2.4}$$

It is easy to see that solutions to (2.4) can be represented in the form (1.3) for

$$c_{Vi} = c_{2i}\Big/\sum c_{2i}, \quad c_{Ki} = c_{1i} - c_{Vi}\sum c_{1i}. \tag{2.5}$$

In particular, in this case the estimators K^* and V^* are well defined whenever

$$\sum c_{2i} \neq 0, \quad \sum c_{1i}^2 > 0.$$

2.2. If the estimators K^* and V^* have been already constructed then, following the recommendations of [4], we can enlarge the class of estimators. To this end, we should choose functions $\gamma_{1i}(K,V)$ and $\gamma_{2i}(K,V)$ and take the estimators K^{**} and V^{**} to be solutions of the equations

$$K^{**}\sum \gamma_{1i}^* v_i/s_i + \sum \gamma_{1i}^* v_i = V^{**}\sum \gamma_{1i}^*, \quad K^{**}\sum \gamma_{2i}^* v_i/s_i + \sum \gamma_{2i}^* v_i = V^{**}\sum \gamma_{2i}^* \tag{2.6}$$

for $\gamma_{1i}^* = \gamma_{1i}(K^*, V^*)$ and $\gamma_{2i}^* = \gamma_{2i}(K^*, V^*)$. It is easy to see that solutions to (2.6) can be represented in the form (1.4) for

$$\gamma_{Vi}(K,V) = \gamma_{2i}(K,V) \Big/ \sum \gamma_{2i}(K,V),$$

$$\gamma_{Ki}(K,V) = \gamma_{1i}(K,V) - \gamma_{Vi}(K,V) \sum \gamma_{1i}(K,V). \tag{2.7}$$

2.3. We now give simple arguments that lead readily to estimators of the form (1.3). We first rewrite (2.1) as

$$(1 + K/s_i)v_i = V + \eta_i \quad \text{for } \eta_i = (1 + K/s_i)\sigma_i \xi_i. \tag{2.8}$$

It follows from (2.8) that for arbitrary $c_{K1}, \ldots, c_{KN}$, $c_{V1}, \ldots, c_{VN}$

$$K \sum c_{Ki} v_i/s_i + \sum c_{Ki} v_i = V \sum c_{Ki} + \sum c_{Ki} \eta_i,$$

$$\sum c_{Vi} \Big(1 + K/s_i\Big) v_i = V \sum c_{Vi} + \sum c_{Vi} \eta_i. \tag{2.9}$$

Now, let c_{Ki} and c_{Vi} be arbitrary constants satisfying the conditions

$$\sum c_{Ki} = 0, \quad \sum c_{Vi} = 1. \tag{2.10}$$

Inserting (2.10) in (2.9), we obtain the representations

$$K = -\frac{\sum c_{Ki} v_i}{\sum c_{Ki} v_i/s_i} + \frac{\sum c_{Ki} \eta_i}{\sum c_{Ki} v_i/s_i},$$

$$V = \sum c_{Vi} \Big(1 + K/s_i\Big) v_i - \sum c_{Vi} \eta_i. \tag{2.11}$$

It is natural to suppose that the weighted sums of the measurement errors in (2.11) are small as compared to the other summands in these equations. Therefore, if in (2.11) we discard the weighted sums of the measurement errors and replace the parameters K and V by their estimators then we arrive at the representation (1.3) for the statistics K^* and V^*.

2.4. Subtracting (2.11) from (1.3), we come to the following useful identities:

$$K^* - K = -\sum c_{Ki} \eta_i \Big/ \sum c_{Ki} v_i/s_i,$$

$$V^* - V = \sum c_{Vi} \eta_i + (K^* - K) \sum c_{Vi} v_i/s_i, \tag{2.12}$$

using which we can guess the asymptotic behavior of the random variables $K^* - K$ and $V^* - V$. However, such an investigation leads to cumbersome calculations. Therefore, proceeding along the lines of [4], it is sometimes more convenient to pass to matrix notation.

3. Optimization of Estimators

3.1. As in [4], the notation $\mathbf{A} = A_{m\times n}$ means that $\mathbf{A}$ is a matrix with m rows and n columns. The entry at the intersection of the pth row and qth column of this matrix is denoted by $(\mathbf{A})_{pq}$. Henceforth the symbol $^\top$ stands for transposition of a vector or matrix. If t is a vector with coordinates $t_1,\dots,t_N$ then t is treated a column vector, whereas $t^\top = (t_1,\dots,t_N)$ is treated a row vector. The symbols $\mathbf{I} = I_{n\times n}$ and $\mathbf{0} = 0_{n\times n}$ stand for the identity and zero matrices of corresponding dimensions. We denote by $\mathbf{T}_a$ the diagonal $(N\times N)$-matrix with diagonal entries $a_1,\dots,a_N$. In particular, $\mathbf{T}_{\mathrm{D}\eta}$ is the covariance matrix of a random vector $\eta = (\eta_1,\dots,\eta_N)^\top$ with independent coordinates.

If $\mathbf{A}$ is a random matrix or vector (whose entries may depend on N) then by the convergence $\mathbf{A} \xrightarrow{p} \mathbf{A}_0$ we mean the coordinatewise convergence in probability for the entries of this matrix or vector (as $N\to\infty$). The convergence $\mathbf{A} \Longrightarrow \Phi_2(0,\mathbf{B})$ means that the distribution of the vector $\mathbf{A}$ may depend on N and it converges weakly (as $N\to\infty$) to the 2-dimensional normal distribution with mean zero and covariance matrix $\mathbf{B}$.

If for some statistics ($\widetilde{K}^*$ and $\widetilde{V}^*$) and a nondegenerate matrix $\mathbf{A} = \mathbf{A}_{2\times 2}$ we have

$$\mathbf{A}^{-1}(\widetilde{K}^* - K, \widetilde{V}^* - V)^\top \Longrightarrow \Phi_2(0,\mathbf{I}),$$

then we say that the two-dimensional statistic $(\widetilde{K}^*, \widetilde{V}^*)$ is an asymptotically normal estimator of the parameter (K,V) with the asymptotical covariance matrix $\mathbf{A}\mathbf{A}^\top$.

If $\mathbf{B}$ is a nonnegative definite matrix then we denote by $\mathbf{B}^{1/2}$ the only nonnegative definite symmetric matrix that satisfies the equality

$$(\mathbf{B}^{1/2})^2 = \mathbf{B} = \mathbf{B}^{1/2}\mathbf{B}^{1/2\top}.$$

We agree that the inequality $\mathbf{B}_1 \ge \mathbf{B}_2$ between nonnegative definite matrices $\mathbf{B}_1$ and $\mathbf{B}_2$ means that the difference $\mathbf{B}_1 - \mathbf{B}_2$ is a nonnegative definite matrix too.

3.2. As mentioned (see (2.2) and (2.3)), the problem under consideration in this article is a particular instance of a more general problem of [4]. Therefore, we follow [4] and introduce the matrices

$$\mathbf{C} = \mathbf{C}_{2\times N} \quad \text{and} \quad \mathbf{\Gamma} = \mathbf{\Gamma}_{2\times N} \quad \text{for } (\mathbf{C})_{ki} = c_{ki} \quad \text{and } (\mathbf{\Gamma})_{ki} = \gamma_{ki}(K,V).$$

Taking (2.3) into account, we define the matrix

$$\mathbf{\Lambda} = \mathbf{\Lambda}(K,V)_{2\times N} \quad \text{for } (\mathbf{\Lambda})_{1i} = -V\big/(K+s_i) \quad \text{and } (\mathbf{\Lambda})_{2i} = 1, \tag{3.1}$$

which will play an important role in the sequel. We note that in our case

$$\mathbf{\Lambda} = \mathbf{Q}_o\mathbf{\Lambda}_o\mathbf{T}^{-1}_{1+K/s} \tag{3.2}$$

for

$$\mathbf{Q}_o = \begin{pmatrix} -V & 0 \\ K & 1 \end{pmatrix}, \quad \mathbf{\Lambda}_o = \begin{pmatrix} 1/s_1, & 1/s_2, & \dots, & 1/s_N \\ 1, & 1, & \dots, & 1 \end{pmatrix}.$$

We introduce the following set of matrices: $\mathcal{M} = \big\{\mathbf{F} = F_{2\times N} : \exists\big(\mathbf{F}\Lambda_o^\top\big)^{-1}\big\}$. Henceforth we assume that $\mathcal{M}$ is nonempty. This means that the matrix $\mathbf{\Lambda}_o$ has rank 2 and that all matrices in $\mathcal{M}$ have rank 2 as well. Therefore, for $\mathbf{F}\in\mathcal{M}$ we can introduce the notation

$$\mathbf{B}(\mathbf{F}) \equiv (\mathbf{F}\mathbf{\Lambda}^\top)^{-1}\mathbf{F}\mathbf{T}_{\mathrm{D}\eta}\mathbf{F}^\top(\mathbf{F}\mathbf{\Lambda}^\top)^{-1\top} \quad \text{for } \mathbf{F}\in\mathcal{M}.$$

Throughout the sequel we assume that $\mathbf{C}, \boldsymbol{\Gamma} \in \mathcal{M}$. It was shown in [4] that under some additional assumptions we have

$$(\mathbf{B}(\mathbf{C}))^{-1/2}(K^* - K, V^* - V) \Longrightarrow \Phi_2(0, \mathbf{I}), \tag{3.3}$$

$$(\mathbf{B}(\boldsymbol{\Gamma}))^{-1/2}(K^{**} - K, V^{**} - V) \Longrightarrow \Phi_2(0, \mathbf{I}). \tag{3.4}$$

As mentioned, (3.3) and (3.4) mean that the statistic $(K^*,\, V^*)$ is an asymptotically normal estimator of the two-dimensional parameter (K, V) with the asymptotic covariance matrix $\mathbf{B}(\mathbf{C})$, whereas $(K^{**},\, V^{**})$ is an asymptotically normal estimator of this parameter with the asymptotic covariance matrix $\mathbf{B}(\boldsymbol{\Gamma})$.

3.3. It is clear that, among the constructed classes of statistics $(K^*,\, V^*)$ and $(K^{**},\, V^{**})$, those estimators are preferable whose asymptotic covariance matrix is less. Therefore, the question arises naturally of optimizing the matrices $\mathbf{B}(\mathbf{C})$ and $\mathbf{B}(\boldsymbol{\Gamma})$. We introduce the matrices $\boldsymbol{\Gamma}^o$ and $\mathbf{B}^{opt}$ by setting

$$\gamma_i^o = (1 + K/s_i)^{-3}\sigma_i^2, \quad \boldsymbol{\Gamma}^o = \boldsymbol{\Lambda}_o \mathbf{T}_{\gamma^o}, \quad \mathbf{B}^{opt} \equiv \mathbf{B}(\boldsymbol{\Gamma}^o).$$

Theorem 1. *Let $\mathbf{H} = \mathbf{H}_{2\times 2}$ be an arbitrary nondegenerate matrix. Then*

$$\forall \boldsymbol{\Gamma} \in \mathcal{M} \quad \mathbf{B}(\boldsymbol{\Gamma}) \geq \mathbf{B}^{opt} = \mathbf{B}(\mathbf{H}\boldsymbol{\Gamma}^o).$$

Corollary 1. *For any c put*

$$\gamma_i^{opt} = c\gamma_i^o = \frac{c\sigma_i^2}{(1 + K/s_i)^3}, \quad c \neq 0.$$

Then the matrix

$$\boldsymbol{\Gamma}^{opt} \equiv \boldsymbol{\Lambda}_o \mathbf{T}_{\gamma^{opt}} \equiv c\boldsymbol{\Gamma}^o \equiv \begin{pmatrix} \gamma_1^{opt}/s_1, & \gamma_2^{opt}/s_2, & \ldots, & \gamma_N^{opt}/s_N \\ \gamma_1^{opt}, & \gamma_2^{opt}, & \ldots, & \gamma_N^{opt} \end{pmatrix} \tag{3.5}$$

is one of the optimal matrices for which the equality in Theorem 1 is achieved.

Remark 1. We emphasize that, for determining the optimal entries $(\boldsymbol{\Gamma}^{opt})_{ki}$, $k = 1, 2$, in the Michaelis–Menten equation, it suffices to construct only one sequence $\{\gamma_i^{opt}\}$ and, constructing optimal matrices, to restrict by (3.5) only to the matrices representable in the form

$$\boldsymbol{\Gamma}_\gamma \equiv \boldsymbol{\Lambda}_o \mathbf{T}_\gamma \equiv \begin{pmatrix} \gamma_1/s_1, & \gamma_2/s_2, & \ldots, & \gamma_N/s_N \\ \gamma_1, & \gamma_2, & \ldots, & \gamma_N \end{pmatrix}. \tag{3.6}$$

Such matrices play an important role in the second half of the article.

The members of the sequence $\{\gamma_i^{opt}\}$ may depend in general on the unknown parameters K and V as well as on the variances $\{\sigma_i^2\}$ (see an exception below in Example 2). This circumstance is one of the causes for which it can be reasonable to use the matrices $\boldsymbol{\Gamma}_\gamma$ with specially chosen functions $\{\gamma_i\}$ in place of the optimal matrices $\boldsymbol{\Gamma}^{opt}$ determined by the unknown functions $\{\gamma_i^{opt}\}$.

3.4. Now, we discuss the question of the extent to which the covariance matrix $\mathbf{B}(\boldsymbol{\Gamma})$ differs from the optimal covariance matrix $\mathbf{B}^{opt}$ in the case of using a nonoptimal matrix $\boldsymbol{\Gamma}$. The following theorem gives one of the possible answers to this question.

Theorem 2. *There exists a nonnegative definite symmetric matrix $\mathbf{B}^o$ such that the relation*

$$\mathbf{B}(\boldsymbol{\Gamma}) - \mathbf{B}^{opt} = (\boldsymbol{\Gamma}\boldsymbol{\Lambda}^\top)^{-1}((\boldsymbol{\Gamma} - \mathbf{H}\boldsymbol{\Gamma}^o)\mathbf{B}^o(\boldsymbol{\Gamma} - \mathbf{H}\boldsymbol{\Gamma}^o)^\top)(\boldsymbol{\Gamma}\boldsymbol{\Lambda}^\top)^{-1\top} \tag{3.7}$$

holds for all matrices $\boldsymbol{\Gamma} \in \mathcal{M}$ and all nondegenerate matrices $\mathbf{H} = \mathbf{H}_{2\times 2}$.

Now, we consider in more detail the important particular case in which the matrix $\mathbf{\Gamma} = \mathbf{\Gamma}_\gamma$ is defined by (3.6). Note that in this case

$$\mathbf{\Gamma}_\gamma - \mathbf{\Gamma}^{opt} = \mathbf{\Lambda}_o(\mathbf{T}_\gamma - c\mathbf{T}_{\gamma^o}) = \mathbf{\Gamma}^o \mathbf{T}_{\gamma/\gamma^o - c}.$$

Corollary 2. *The relation*

$$\mathbf{B}(\mathbf{\Gamma}_\gamma) - \mathbf{B}^{opt} = (\mathbf{\Gamma}_\gamma \mathbf{\Lambda}^\top)^{-1} \mathbf{\Gamma}^o (\mathbf{T}_{\gamma/\gamma^o - c} \mathbf{B}^o \mathbf{T}_{\gamma/\gamma^o - c}) \mathbf{\Gamma}^{o\top} (\mathbf{\Gamma}_\gamma \mathbf{\Lambda}^\top)^{-1\top} \tag{3.8}$$

holds for every collection of functions $\{\gamma_i\}$ *and every constant* c.

Remark 2. Identity (3.7) can be interpreted as follows: The more precisely the matrix $\mathbf{\Gamma}$ is approximated by some optimal matrix $\mathbf{\Gamma}^{opt} = \mathbf{H}\mathbf{\Gamma}^o$, the closer the asymptotic covariance matrix $\mathbf{B}(\mathbf{\Gamma})$ is to the corresponding estimator of the optimal matrix $\mathbf{B}^{opt}$. Moreover, the matrix $(\mathbf{\Gamma} - \mathbf{H}\mathbf{\Gamma}^o)\mathbf{B}^o(\mathbf{\Gamma} - \mathbf{H}\mathbf{\Gamma}^o)^\top$, arising in (3.7), characterizes the degree of proximity.

Similarly, the matrix $\mathbf{T}_{\gamma/\gamma^o - c}\mathbf{B}^o\mathbf{T}_{\gamma/\gamma^o - c}$ with the entries

$$\{(\gamma_i/\gamma_i^o - c)(\mathbf{B}^o)_{ij}(\gamma_j/\gamma_j^o - c)\}$$

in (3.8) characterizes the degree of this proximity in terms of the relative errors $\{\gamma_i/\gamma_i^o\}$ of the approximation of the chosen quantities $\{\gamma_i\}$ by the optimal $\{\gamma_i^o\}$. From this fact we can conclude that the estimators K^* and V^* constructed in (1.3) and (1.5) are as more precise as less the quantities $c_i(1 + K/s_i)^3\sigma_i^2$ depend on i. Similarly, to increase the accuracy of the estimators K^{**} and V^{**} defined by (1.4) and (1.6) we recommend choosing the functions $\gamma_i(K, V)$ so that the quantities $\gamma_i(K, V)(1 + K/s_i)^3\sigma_i^2$ be closer to some constant c as much as possible.

3.5. We note that all assertions of this section are of a "conditional" character; i.e., they are valid provided that (3.3) and (3.4) hold. The question of the restrictions under which (3.3) and (3.4) hold will be discussed in the next section, but some practical recommendations for the choice of the entries of the matrices $\mathbf{C}$ and $\mathbf{\Gamma}$ are given in the current subsection.

Example 1 Let

$$\sigma_i^2 \equiv \sigma^2/W_i(K, V),$$

where $W_i(K, V)$ are known functions, while the parameter σ^2 may be unknown. Then we can put $c = 1/\sigma^2$ in Corollary 1 and take

$$\gamma_i^{opt} \equiv W_i(K, V)/(1 + K/s_i)^3 \quad \text{and} \quad (\mathbf{\Gamma}^{opt})_{ki} = \begin{cases} \gamma_i^{opt}/s_i, & k = 1, \\ \gamma_i^{opt}, & k = 2. \end{cases}$$

Observe that this example includes two cases that, according to [5], are most important in practice: one in which the variances $\sigma_i^2 = \sigma^2$ are constant and the other in which $\sigma_i^2 = \sigma^2/(1 + K/s_i)^2$, i.e., the variances σ_i^2 are proportional to the mean of the random variables v_i.

Example 2 Let

$$\sigma_i^2 = \sigma^2/w_{0i}(1 + K/s_i)^3,$$

where w_{0i} are known numbers, while the parameter σ^2 may be unknown. Then the quantities

$$\gamma_i^{opt} \equiv w_{0i}$$

are independent of the unknown parameters K, V, and σ^2; i.e., in this case we can determine the optimal constants that minimize the covariance matrix even at the first step by setting $\gamma_i^{opt} \equiv w_{0i}$ in Corollary 1.

3.6. Thus, in Examples 1 and 2 (of course, assuming validity of (3.3) and (3.4)) in quite general cases we constructed asymptotically normal estimators with the optimal asymptotic covariance matrix $\mathbf{B}^{opt}$.

Remark 3. Constructing the estimators K^* and V^* in Example 1, we can put, for instance, $c_i = \gamma_i^{opt}(K_o, V_o)$ with some K_o and V_o chosen in advance.

Remark 4. If an exact form of the variance σ_i^2 is unknown then we cannot find the sequence $\{\gamma_i^o\}$ and construct the estimators K^{**} and V^{**} for $\mathbf{\Gamma} = c\mathbf{\Gamma}^o$. In this case we can recommend taking the entries of $\mathbf{\Gamma}$ to be functions of which we may assume that they "differ slightly" from the unknown entries of the matrix $c\mathbf{\Gamma}^o$. Then, by Remark 2, the "better" the chosen entries of the matrix $\mathbf{\Gamma}$ approximate the entries of the optimal matrix $c\mathbf{\Gamma}^o$, the less the asymptotic variance of the resultant estimators differs from $\mathbf{B}^{opt}$.

Remark 5. If there is no information about the behavior of the variance σ_i^2 then we can recommend to take at the first step $c_{1i} = s_i^2/(K_o + s_i)^3$ and $c_{2i} = s_i^3/(K_o + s_i)^3$ for some K_o. In this case the estimators K^* and V^* are as more precise as closer the chosen value K_o is to the unknown true value of the parameter K and as closer the unknown values $(1 + K/s_i)^3\sigma_i^2$ are to some σ^2.

3.7. It only remains to prove the above-stated assertions. First of all, we note that Theorem 1 and Corollaries 1 and 2 are immediate from Theorem 2. To prove Theorem 2, we use the results of [4]. Observe that Theorem 4 of [4] differs from our Theorem 2 only by the presence of the matrix $\mathbf{R}\mathbf{\Lambda}(\boldsymbol{\theta})\mathbf{V}^{-1}$ instead of $\mathbf{H}\mathbf{\Gamma}^o$. However, in our case, by (2.8) and (3.2)

$$\mathbf{V} = \mathbf{T}_{D\eta} \equiv \mathbf{T}_{(1+K/s)^2\sigma^2} \quad \text{and} \quad \mathbf{\Lambda} = \mathbf{Q}_o\mathbf{\Lambda}_o\mathbf{T}^{-1}_{1+K/s}.$$

Therefore,

$$\mathbf{R}\mathbf{\Lambda}(\boldsymbol{\theta})\mathbf{V}^{-1} = \mathbf{R}\mathbf{Q}_o\mathbf{\Lambda}_o\mathbf{T}^{-1}_{1+K/s}\mathbf{T}^{-1}_{(1+K/s)^2\sigma^2} \equiv \mathbf{H}\mathbf{\Lambda}_o\mathbf{T}^{-1}_{(1+K/s)^3\sigma^2} = \mathbf{H}\mathbf{\Gamma}^o$$

for $\mathbf{H} = \mathbf{R}\mathbf{Q}_o$.

3.8. Remark 6. Suppose that the errors ξ_i are independent and have standard normal distribution while the variances σ_i^2 are independent of the parameters K and V. Then

$$\mathbf{B}^{opt} = \mathbf{I}_N^{-1},$$

where $\mathbf{I}_N$ is the Fisher information for the sample $v_1, \dots, v_N$ (see [4, Remark 3]). By analogy to the Rao–Cramér inequality, we should thus expect a certain unimprovability of the estimators K^* and V^* in case we choose $\mathbf{C} = \mathbf{\Gamma}^{opt}$, and the estimators K^{**} and V^{**} in case we choose $\mathbf{\Gamma} = \mathbf{\Gamma}^{opt}$.

4. Conditions for Consistency and Asymptotic Normality

4.1. It is obvious that the question of consistency and asymptotic normality of the constructed estimators depends essentially on the choice of the constants $\{c_{ki}\}$, functions $\{\gamma_{ki}\}$, and behavior of the unknown variances $\{\sigma_i\}$. However, consistent and asymptotically normal estimators for K and V may fail to exist without additional assumptions on the behavior of the numbers s_i.

For example, if $s_i \equiv s$ then we can estimate only $\frac{V}{1+K/s}$, but cannot estimate K or V separately.

Of the two main assumptions to be imposed on the behavior of s_i below, the first is as follows: there exists a distribution function $F(x)$ such that

$$\frac{\sharp\{i : s_i < x\}}{N} \to F(x), \tag{4.1}$$

where the symbol $\sharp A$ stands for the number of elements in a set A.

Remark 7. If s_i are chosen at random in a real experiment then we should take as $F(x)$ the distribution function of the random variable s_i.

Remark 8. If we "plan" the choice of s_i (for example, choose half of s_i near K and the other half "as far as possible" (see [6]) then the distribution function $F(x)$ describes the asymptotic behavior of such a "plan" (see, for instance, [7]).

From assumption (4.5) formulated below, it follows in particular that the distribution function $F(x)$ is not concentrated at a single point.

4.2. Suppose that a random variable $\nu(N)$ takes the values $1, \dots, N$ with equal probabilities, i.e.,

$$\mathbf{P}(\nu(N) = i) = 1/N \quad \text{for } i = 1, \dots, N.$$

Denote a random variable with the distribution function $F(x)$ by s. In this case, assumption (4.1) can be rewritten as

$$s_{\nu(N)} \Longrightarrow s, \quad \mathbf{P}(s < x) = F(x) \text{ for every } x. \tag{4.2}$$

We say that *condition* (A) *holds for* $A_0^2(\cdot)$ if (4.2) is satisfied and moreover

$$\frac{1}{N}\sum_{i=1}^{N} A_0^2(s_i) \equiv \mathbf{E}A_0^2(s_{\nu(N)}) \to \mathbf{E}A_0^2(s) < \infty. \tag{4.3}$$

We emphasize that in each of the assertions of this section we use its own specially chosen functions $A_0^2(\cdot) = A_0^2(\cdot, K, V) \ge 0$ which each time are continuous in their first argument. We note that by Helly's theorem (see [8]), convergence (4.3) follows from (4.2) whenever the function $A_0^2(\cdot, K, V)$, continuous in its first argument, is bounded. If the function $A_0^2(\cdot, K, V)$ is unbounded then (4.3) is supplementary to (4.2).

4.3. We now turn to studying the estimators K^* and V^* defined by (1.3). Throughout the sequel, we assume that the numbers c_{Vi} and c_{Ki} in (1.5) are representable in the form

$$c_{Vi} \equiv c_2(s_i)\Big/\sum c_2(s_i), \quad c_{Ki} \equiv c_1(s_i) - c_{Vi}\sum c_1(s_i), \tag{4.4}$$

where $c_1(\cdot)$ and $c_2(\cdot)$ are specially chosen continuous functions. Below we frequently use the following constraints on these functions:

$$\mathbf{E}c_1(s)\mathbf{E}\frac{c_2(s)}{K+s} \ne \mathbf{E}c_2(s)\mathbf{E}\frac{c_1(s)}{K+s}, \quad \mathbf{E}c_2(s) \ne 0. \tag{4.5}$$

In what follows we also assume that the variances of the random variables $\{\eta_i = (1 + K/s_i)\xi_i\}$ satisfy the extra conditions

$$\forall i \quad (1 + K/s_i)^2\sigma_i^2 \equiv \mathbf{D}\eta_i = w_i^2\sigma^2 > 0 \quad \text{for } w_i = w(s_i, K, V),$$

where $w(\cdot, \cdot, \cdot)$ is some positive function continuous in its first argument.

Theorem 3. *Suppose that conditions* (4.5) *and* (A) *hold for*

$$A_0^2(\cdot) = (|c_1(\cdot)| + |c_2(\cdot)|)(1 + w(\cdot, K, V)).$$

Then K^* and V^* are consistent estimators of the parameters K and V.

4.4. We turn to studying the conditions of asymptotic normality of the estimators K^* and V^*. We need the following notation:

$$a_K = \sum \frac{c_{Ki}}{K+s_i}, \quad a_V = \sum \frac{c_{Vi}}{K+s_i}, \quad b_K^2 = \sum c_{Ki}^2 w_i^2, \quad b_V^2 = \sum c_{Vi}^2 w_i^2,$$
$$b_{KV} = \sum c_{Ki} c_{Vi} w_i^2, \quad \rho_{KV} = b_{KV}/b_K^2, \quad b_{VV}^2 = b_V^2 - b_{KV}^2/b_K^2.$$

Theorem 4. *Suppose that conditions* (4.5) *and* (A) *hold for*

$$A_0^2(\cdot) = (|c_1(\cdot)| + |c_2(\cdot)|) + \big(c_1^2(\cdot) + c_2^2(\cdot)\big) w^2(\cdot, K, V).$$

Then

$$\left(\frac{a_K V(K^* - K)}{\sigma b_K}, \frac{(V^* - V) + (a_V/a_K + a_K \rho_{KV} V)(K^* - K)}{\sigma b_{VV}}\right)^\top \Longrightarrow \Phi_2(0, \mathbf{I}).$$

Corollary 3. *Under the assumptions of Theorem* 4*, the estimators K^* and V^* are consistent and* (3.3) *holds.*

The asymptotic covariance matrices in Theorem 4 and Corollary 2 depend essentially on the unknown parameters K and V and perhaps on the unknown variances σ_i^2. Therefore, while constructing confidence intervals and testing hypotheses, it should be good to have an assertion in which at least one of these matrices is replaced by a known matrix. We give some solution of this problem for the estimators of the first step. We put

$$\eta_i^* = v_i + K^* v_i/s_i - V^*, \tag{4.6}$$

$$a_K^* = \sum c_{Ki} v_i/s_i, \quad a_V^* = \sum c_{Vi} v_i/s_i, \quad b_K^{*2} = \sum c_{Ki}^2 \eta_i^{*2}, \quad b_V^{*2} = \sum c_{Vi}^2 \eta_i^{*2},$$
$$b_{KV}^* = \sum c_{Ki} c_{Vi} \eta_i^{*2}, \quad \rho_{KV}^* = b_{KV}^*/b_K^{*2}, \quad b_{VV}^{*2} = b_V^{*2} - b_{KV}^{*2}/b_K^{*2}.$$

Theorem 5. *Suppose that conditions* (4.5) *and* (A) *hold for*

$$A_0^2(\cdot) = \big(c_1^2(\cdot) + c_2^2(\cdot)\big)(1 + w^2(\cdot, K, V)).$$

Then

$$\left(\frac{a_K^*(K^* - K)}{b_K^*}, \frac{(V^* - V) + (a_V^*/a_K^* + a_K^* \rho_{KV}^*)(K^* - K)}{b_{VV}^*}\right)^\top \Longrightarrow \Phi_2(0, \mathbf{I}).$$

Corollary 4. *Suppose that the functions $c_k(x)$ in* (4.4) *are representable in the form*

$$c_1(x) = c(x)/x \quad \text{and} \quad c_2(x) = c(x) \quad \text{for } x > 0; \tag{4.7}$$

moreover,

$$\mathbf{P}(s = t) < 1 \quad \text{for every } t \text{ and } \quad \mathbf{P}(c(s) > 0) = 1, \tag{4.8}$$

where $c(\cdot)$ is some continuous function. Suppose that condition (A) holds for

$$A_0^2(x) = c^2(x)(1 + 1/x^2)(1 + w^2(x, K, V)).$$

Then all assertions of Corollary 3 *and Theorems* 3–5 *are valid.*

4.5. We turn to studying the estimators K^{**} and V^{**} defined by (1.4). Throughout the sequel, we assume that the statistics γ_{Vi}^* and γ_{Ki}^* in (1.4) are representable in the form

$$\gamma_{Vi}^* \equiv \gamma_2(s_i, K^*) \Big/ \sum_{i=1}^N \gamma_2(s_i, K^*), \quad \gamma_{Ki}^* \equiv \gamma_1(s_i, K^*) - \gamma_{Vi}^* \sum_{i=1}^N \gamma_1(s_i, K^*), \tag{4.9}$$

where $\gamma_j(s, K) > 0$, $j = 1, 2$, are specially chosen positive functions.

We say that *condition* (D) *holds* if the functions $\{\gamma_j(s_i,K)\}$ are continuous in the first argument, differentiable in the second, their partial derivatives with respect to the second argument satisfy the condition

$$\sup_{K/2\le t\le 2K}\left|\frac{\partial\gamma_j(s_i,t)}{\partial t}\right|\le M(s_i,K)<\infty,\quad j=1,2,$$

and the function $M(\cdot,\cdot)$ is continuous in its first argument.

We need the following notations:

$$\tilde a_K=\sum\frac{\gamma_{Ki}}{K+s_i},\quad \tilde a_V=\sum\frac{\gamma_{Vi}}{K+s_i},\quad \tilde b_K^2=\sum\gamma_{Ki}^2w_i^2,\quad \tilde b_V^2=\sum\gamma_{Vi}^2w_i^2,$$

$$\tilde b_{KV}=\sum\gamma_{Ki}\gamma_{Vi}w_i^2,\quad \tilde\rho_{KV}=\tilde b_{KV}/\tilde b_K^2,\quad \tilde b_{VV}^2=\tilde b_V^2-\tilde b_{KV}^2/\tilde b_K^2.$$

Also, we will use the condition

$$\mathbf{E}\gamma_1(s)\mathbf{E}\frac{\gamma_2(s)}{K+s}\ne\mathbf{E}\gamma_2(s)\mathbf{E}\frac{\gamma_1(s)}{K+s},\quad \mathbf{E}\gamma_2(s)\ne 0. \tag{4.10}$$

Theorem 6. *Suppose that conditions* (4.5)*,* (4.10)*,* (D)*, and* (A) *hold for*

$$A_0^2(\cdot)=(|c_1(\cdot)|+|c_2(\cdot)|+|\gamma_1(\cdot,K)|+|\gamma_2(\cdot,K)|+M(\cdot,K))$$
$$+\big(c_1^2(\cdot)+c_2^2(\cdot)+\gamma_1^2(\cdot,K)+\gamma_2^2(\cdot,K)+M(\cdot,K)^2\big)w^2(\cdot,K,V).$$

Then

$$\left(\frac{\tilde a_K V(K^{**}-K)}{\sigma\tilde b_K},\frac{(V^{**}-V)+(\tilde a_V/\tilde a_K+\tilde a_K\tilde\rho_{KV}V)(K^{**}-K)}{\sigma\tilde b_{VV}}\right)^\top\Longrightarrow\Phi_2(0,\mathbf{I}).$$

Corollary 5. *If all assumptions of Theorem* 6 *are satisfied then* (3.4) *holds.*

Remark 9. Suppose that the functions $\gamma_j(\cdot,\cdot)$ in (4.9) are representable in the form

$$\gamma_1(x,K)=\gamma(x,K)/x\quad\text{and } \gamma_2(x,K)=\gamma(x,K)\quad\text{for } x>0,$$

where $\gamma(\cdot,K)\ge 0$ is some function continuous in its first argument. As follows from the forthcoming Lemma 21, in this case the conditions

$$\mathbf{P}(s=t)<1\quad\forall t\quad\text{and}\quad 0<\mathbf{E}(1+1/s)\gamma(s,K)<\infty$$

guarantee validity of (4.10).

4.6. We now describe the idea behind the proofs of Theorems 3–6 and their corollaries. The proofs themselves are exposed in § 5 and § 6. As we have repeated constantly, the problem under consideration is a special particular case of a general problem of [4] under the assumptions (2.3). And the proofs below are as a rule based on verification of the conditions of more general assertions of [4].

By analogy to [4], we introduce the following notations:

$$\eta=(\eta_1,\dots,\eta_N)^\top,\quad v=(v_1,\dots,v_N)^\top,$$

$$\mathbf{X}=\mathbf{X}_{2\times N},\quad (\mathbf{X})_{1i}=-v_i/s_i,\quad (\mathbf{X})_{2i}=1,\quad \mathbf{\Psi}=-(\mathbf{X}-\mathbf{\Lambda}),$$

where $\mathbf{\Lambda}=\mathbf{E}\mathbf{X}$ is defined by (3.1). Moreover, we need the matrices

$$\mathbf{U}=\begin{pmatrix}N^{1/2}/b_K & 0\\ -N^{1/2}\rho_{KV}/b_{VV} & N^{-1/2}/b_{VV}\end{pmatrix},\quad \mathbf{R}=\begin{pmatrix}1 & -\sum c_1(s_i)/\sum c_2(s_i)\\ 0 & N/\sum c_2(s_i)\end{pmatrix}. \tag{4.11}$$

Finally, we denote by $\mathbf{\Gamma}(K)$, $\mathbf{R}(K)$, and $\widetilde{\mathbf{U}}$ the matrices that result from substituting $\gamma_1(s,K)$ for $c_1(s)$ and $\gamma_2(s,K)$ for $c_2(s)$ in the definitions of the respective matrices $\mathbf{C}$, $\mathbf{R}$, and $\mathbf{U}$.

Remark 10. Note that the matrices $\mathbf{R}$ and $\mathbf{R}(K)$ are chosen so that

$$\mathbf{RC} = \begin{pmatrix} c_{K1}, \dots, c_{KN} \\ Nc_{V1}, \dots, Nc_{VN} \end{pmatrix}, \quad \mathbf{R}(K)\mathbf{\Gamma}(K) = \begin{pmatrix} \gamma_{K1}, \dots, \gamma_{KN} \\ N\gamma_{V1}, \dots, N\gamma_{VN} \end{pmatrix}, \tag{4.12}$$

and the matrices $\mathbf{U}$ and $\widetilde{\mathbf{U}}$ are chosen so that the assertions of Theorems 4 and 6 can be rewritten in the form

$$\sigma^{-1}N^{-1/2}\mathbf{URC\Lambda}^\top(K^*-K, V^*-V) \Longrightarrow \Phi_2(0,\mathbf{I}), \tag{4.13}$$

$$\sigma^{-1}N^{-1/2}\widetilde{\mathbf{U}}\mathbf{R}(K)\mathbf{\Gamma}(K)\mathbf{\Lambda}^\top(K^{**}-K, V^{**}-V) \Longrightarrow \Phi_2(0,\mathbf{I}). \tag{4.14}$$

Moreover, the random variables $\sigma^{-1}N^{-1/2}\mathbf{URC}\eta$ and $\sigma^{-1}N^{-1/2}\widetilde{\mathbf{U}}\mathbf{R}(K)\mathbf{\Gamma}(K)\eta$ have the identity covariance matrices; i.e.,

$$\mathbf{UR}(\mathbf{CT}_w^2\mathbf{C}^\top)\mathbf{R}^\top\mathbf{U}^\top/N = \mathbf{I}, \quad \widetilde{\mathbf{U}}\mathbf{R}(K)(\mathbf{\Gamma}(K)\mathbf{T}_w^2\mathbf{\Gamma}^\top(K))\mathbf{R}^\top(K)\widetilde{\mathbf{U}}^\top/N = \mathbf{I}. \tag{4.15}$$

Thus, with the above notations, equations (2.4) and (2.6) defining the estimators θ^* and θ^{**} can be rewritten in the following equivalent matrix form:

$$\mathbf{CX}^\top(K^*,V^*)^\top = \mathbf{C}v, \quad \mathbf{\Gamma}(K^*)\mathbf{X}^\top(K^{**},V^{**})^\top = \mathbf{\Gamma}(K^*)v. \tag{4.16}$$

As noted in Remark 5 of [4], these equations do not change if we multiply them from the left by arbitrary nondegenerate matrices:

$$\mathbf{RCX}^\top(K^*,V^*)^\top = \mathbf{RC}v, \quad \mathbf{R}(K^*)\mathbf{\Gamma}(K^*)\mathbf{X}^\top(K^{**},V^{**})^\top = \mathbf{R}(K^*)\mathbf{\Gamma}(K^*)v. \tag{4.17}$$

On using (4.12), it is easy to check that the first equation in (4.17) coincides with (2.9) for c_{Ki} and c_{Vi} defined by (2.5); i.e., it coincides with a system of equations whose coefficients satisfy (2.10) and immediately yield the estimators K^* and V^* of the first step. The second equation in (4.17) leads immediately to the improved estimators K^{**} and V^{**}.

Therefore, in this article, while studying estimators we will ground on equations (4.17) rather than on less convenient equations (4.16). Using the general results of [4], we will replace in them the matrices $\mathbf{C}$ by $\mathbf{RC}$, $\mathbf{\Gamma}(\theta^*)$ by $\mathbf{R}(K^*)\mathbf{\Gamma}(K^*)$, and the matrices $\mathbf{U}$ and $\mathbf{U}_\Gamma$ by $\sigma^{-1}N^{-1/2}\mathbf{U}$ and $\sigma^{-1}N^{-1/2}\widetilde{\mathbf{U}}$ respectively.

Remark 11. Observe that, in view of relations (4.13)–(4.15), Corollaries 3 and 5 are immediate from the corresponding Corollaries 5 and 6 of [4].

4.7. To simplify proofs, henceforth we do not indicate dependence of some quantities on the true values of the parameters K and V which are assumed to be fixed. In particular, we will use the notations

$$w(s) = w(s,K,V),\ w_i = w(s_i,K,V),\ k(s) \equiv (K+s)^{-1},\ k_i = k(s_i) = (K+s_i)^{-1},$$

$$M(s) = M(s,K,V),\ \gamma_j(s) = \gamma_j(s,K),\ \gamma_{Ki} = \gamma_{Ki}(s_i,K),\ \gamma_{Vi} = \gamma_{Vi}(s_i,K).$$

Furthermore, we assume that

$$\Delta = \mathbf{\Gamma}(K^*) - \mathbf{\Gamma}(K), \quad \delta_{ji} \equiv (\Delta)_{ji} = \mathbf{\Gamma}_j(s_i,K^*) - \mathbf{\Gamma}_j(s_i,K), \quad j=1,2.$$

5. Proofs of Theorems 3–5

5.1. As a preliminary, we prove some auxiliary assertions.

Lemma 1. *If condition* (A) *holds then*

$$\frac{1}{N}\sum_{i=1}^{N} A(s_i) \to \mathbf{E}A(s) \tag{5.1}$$

for every function $A(s)$ continuous in the first argument and satisfying the condition

$$\forall x \quad |A(x)| \le C_0\big(1 + A_0^2(x)\big) \tag{5.2}$$

with some constant C_0.

See a proof in, for instance, [9, Proposition 5].

Remark 12 Thus, if some continuous function $A(\cdot)$ satisfies conditions (5.2) and (A) then, in addition to (5.1), we also have the convergences

$$\frac{1}{N}\sum A(s_i)k_i \to \mathbf{E}A(s)k(s), \quad \frac{1}{N}\sum A(s_i)k_i^2 \to \mathbf{E}A(s)k^2(s),$$

since $k(s_i) \le 1/K$. If for all x

$$A_1^2(x) + A_2^2(x) \le C_0(1 + A_0^2(x))$$

for some continuous functions $A_1(\cdot)$ and $A_2(\cdot)$ then Lemma 1 trivially implies the convergences

$$\frac{1}{N}\sum A_j(s_i) \to \mathbf{E}A_j(s), \quad \frac{1}{N}\sum A_j^2(s_i) \to \mathbf{E}A_j^2(s) \quad \text{for } j = 1, 2,$$

$$\frac{1}{N}\sum A_1(s_i)A_2(s_i) \to \mathbf{E}A_1(s)A_2(s).$$

Lemma 2. *Suppose that $\mathbf{E}|\zeta| < \infty$ and that condition* (A) *holds for $A_0^2(s) = |a(s)|$, where $a(s)$ is some continuous function. Then for every $\alpha > 0$*

$$D_{0,N} \equiv \sum \mathbf{E}\min\left\{\left|\frac{a(s_i)\zeta}{N}\right|^{1+\alpha}, \left|\frac{a(s_i)\zeta}{N}\right|\right\} \to 0. \tag{5.3}$$

Proof. Rewrite (5.3) in the following equivalent form:

$$D_{0,N} = \frac{1}{N}\sum \mathbf{E}\left(|a(s_i)||\zeta|\min\left\{\left(\frac{|a(s_i)||\zeta|}{N}\right)^{\alpha}, 1\right\}\right), \tag{5.4}$$

and use the inequality

$$\min\left\{\frac{(|a(s_i)|\,|\zeta|)^{\alpha}}{N^{\alpha}}, 1\right\} \le \min\left\{\frac{|a(s_i)|^{\alpha}}{N^{\alpha/2}}, 1\right\} + \min\left\{\frac{|\zeta|^{\alpha}}{N^{\alpha/2}}, 1\right\}. \tag{5.5}$$

Then, from (5.4) and (5.5) obtain

$$D_{0,N} \le \frac{1}{N}\sum f_N(s_i)\mathbf{E}|\zeta| + \frac{1}{N}\sum |a(s_i)|\mathbf{E}\min\left\{\frac{|\zeta|^{1+\alpha}}{N^{\alpha/2}}, |\zeta|\right\}, \tag{5.6}$$

where

$$f_M(s) = \min\left\{\frac{|a(s)|^{1+\alpha}}{M^{\alpha/2}}, |a(s)|\right\}. \tag{5.7}$$

Observe that, for each fixed $M > 0$, the function $f_M(s)$ is continuous in s and that $|f_M(s)| \le |a(s)| \equiv A_0^2(s)$. Hence, by Lemma 1

$$\frac{1}{N}\sum |a(s_i)| \to \mathbf{E}|a(s)| \quad \text{and} \quad \frac{1}{N}\sum f_M(s_i) \to \mathbf{E} f_M(s).$$

Thus, from the last two relations we find that

$$\frac{1}{N}\sum f_N(s_i) \le \frac{1}{N}\sum f_M(s_i) \to \mathbf{E} f_M(s) \quad \forall M > 0.$$

However, by Lebesgue's theorem and in view of definition (5.7)

$$\mathbf{E} f_M(s) = \mathbf{E}\min\Big\{|a(s)|^{1+\alpha}/M^{\alpha/2}, |a(s)|\Big\} \to \mathbf{E}\min\{0, |a(s)|\} = 0$$

as $M \to \infty$. Hence,

$$\lim \frac{1}{N}\sum f_N(s_i) = 0. \tag{5.8}$$

From Lebesgue's theorem we also immediately obtain

$$\mathbf{E}\min\Big\{|\zeta|^{1+\alpha}/N^{\alpha/2}, |\zeta|\Big\} \to 0. \tag{5.9}$$

Inserting (5.8) and (5.9) in (5.6), we now arrive at (5.4).

Lemma 3. *Suppose that condition* (A) *holds for* $A_0^2(s) = |a(s)|$*, where* $a(s)$ *is some continuous function. Then*

$$\frac{1}{N}\sum a(s_i)\xi_i \xrightarrow{p} 0.$$

Proof. This statement is a particular instance of the Law of Large Numbers for a scheme of series (see [8, Chapter 8, Theorem 3]). Therefore, it suffices to demonstrate that

$$D_{1,N} \equiv \sum \mathbf{E}\min\left\{\left(\frac{a(s_i)\xi_i}{N}\right)^2, \left|\frac{a(s_i)\xi_i}{N}\right|\right\} \to 0. \tag{5.10}$$

The quantities ξ_i in this expression can be replaced with ξ_1, as their distributions coincide. Now, the needed convergence in (5.10) follows from Lemma 2 with $\zeta = \xi_1$ and $\alpha = 1$. □

Lemma 4. *Suppose that condition* (A) *holds for* $A_0^2 = b^2(s)$*, where* $b(s)$ *is some continuous function. Then*

$$\frac{1}{\sqrt{N}}\sum b(s_i)\xi_i \Longrightarrow \mathbf{\Phi}_1(0, \mathbf{E}b^2(s)).$$

Proof. Since $\mathbf{E}b(s_i)\xi_i = 0$ for all i and

$$\mathbf{D}\left(\frac{1}{\sqrt{N}}\sum b(s_i)\xi_i\right) = \frac{1}{N}\sum b^2(s_i) \to \mathbf{E}b^2(s) < \infty$$

by (4.3), this statement is a particular instance of the Central Limit Theorem for a scheme of series. Therefore, it suffices to check (see [8, Chapter 8, Theorem 5]) fulfillment of the conditions

$$D_{2,N} \equiv \sum \mathbf{E}\min\left\{\left(\frac{b(s_i)\xi_i}{\sqrt{N}}\right)^2, \left(\frac{b(s_i)\xi_i}{\sqrt{N}}\right)^3\right\} \to 0.$$

However, since the sequence $\{\xi_i\}$ is identically distributed, the representation for $D_{2,N}$ can be rewritten in the following equivalent form:

$$D_{2,N} = \sum \mathbf{E}\min\left\{ \left(\frac{b^2(s_i)\xi_1^2}{N} \right)^{1+\frac{1}{2}}, \frac{b^2(s_i)\xi_1^2}{N} \right\} \to 0. \tag{5.11}$$

Convergence (5.11) is now immediate from Lemma 2 with $a(s) = b^2(s)$, $\zeta = \xi_1^2$, and $\alpha = 1/2$. □

Lemma 5. *Suppose that the conditions of Lemma* 3 *are satisfied. Then*

$$\frac{1}{N} \sum a(s_i)(\xi_i^2 - 1) \xrightarrow{p} 0.$$

Proof. The claim of the lemma is a particular instance of the Law of Large Numbers for a scheme of series. Therefore, by analogy to the derivation of Lemma 3, it suffices to check the convergence

$$D_{1,N} \equiv \sum \mathbf{E}\min\left\{ \left(\frac{a(s_i)(\xi_i^2 - 1)}{N} \right)^2, \left| \frac{a(s_i)(\xi_i^2 - 1)}{N} \right| \right\} \to 0.$$

Since the sequence $\{\xi_i\}$ is identically distributed, the latter convergence is immediate from Lemma 2 with $\zeta = \xi_1^2 - 1$ and $\alpha = 1$. □

5.2. We now state some corollaries to the proven assertions.

Lemma 6. *Suppose that condition* (A) *holds for* $A_0^2(s) = |c_1(s)| + |c_2(s)|$. *Then*

$$\frac{1}{N} \sum c_j(s_i) \to \mathbf{E}c_1(s), \quad \frac{1}{N} \sum c_j(s_i)k_i \to \mathbf{E}c_j(s)k(s), \quad j = 1, 2.$$

Lemma 7. *If condition* (A) *holds for* $A_0^2(s) = (|c_1(s)| + |c_2(s)|)w(s)$ *then*

$$\frac{1}{N} \sum c_j(s_i)w(s_i)\xi_i \xrightarrow{p} 0, \quad \frac{1}{N} \sum c_j(s_i)w(s_i)k_i\xi_i \xrightarrow{p} 0, \quad j = 1, 2.$$

Lemma 6 is immediate from Remark 12 and Lemma 7, from Lemma 3 and the same Remark 12.

Lemma 8. *Suppose that conditions* (4.5) *and* (A) *hold for* $A_0^2(s) = (c_1^2(s) + c_2^2(s))w^2(s)$. *Then*

$$\mathbf{C}T_{D\eta}\mathbf{C}^\top/N \to \mathbf{V}_0,$$

where $\mathbf{V}_0$ *is some nondegenerate function.*

Proof. The convergence

$$\mathbf{C}T_{D\eta}\mathbf{C}^\top/N \to \sigma^2 \begin{pmatrix} \mathbf{E}c_1^2(s)w^2(s) & \mathbf{E}c_1(s)c_2(s)w^2(s) \\ \mathbf{E}c_1(s)c_2(s)w^2(s) & \mathbf{E}c_2^2(s)w^2(s) \end{pmatrix} \equiv \mathbf{V}_0$$

is immediate from Remark 12. Were the function $\mathbf{V}_0$ degenerate, we would have the equality

$$\mathbf{E}c_1^2(s)w^2(s)\mathbf{E}c_2^2(s)w^2(s) = (\mathbf{E}c_1(s)c_2(s)w^2(s))^2. \tag{5.12}$$

However, as indicated in the well-known derivation of Schwartz's inequality, equality (5.12) is possible only in the case of

$$\mathbf{P}((c_2(s) - tc_1(s))^2 w^2(s) = 0) = 1$$

for some nonrandom t. The last could hold only if $\mathbf{P}(c_2(s) = tc_1(s)) = 1$, which contradicts (4.5). □

Lemma 9. *If the conditions of Lemma* 8 *are satisfied then*

$$\mathbf{C}\eta/\sqrt{N} \Longrightarrow \Phi_2(0, \mathbf{V}_0). \tag{5.13}$$

Proof. We use the Cramér–Wald method. It suffices to show that

$$\forall t_1, t_2 \quad \frac{1}{\sqrt{N}} \sum a(s_i, t_1, t_2)\xi_i \Longrightarrow \mathbf{\Phi}_1(0, \mathbf{E}a^2(s, t_1, t_2))$$

for

$$a(s, t_1, t_2) = t_1 c_1(s)w(s) + t_2 c_2(s)w(s).$$

This convergence follows from Lemma 4 and Remark 12. □

5.3. We now prove the main lemmas needed for derivation of Theorems 3 and 4.

Lemma 10. *If the conditions of Theorem* 3 *are satisfied then*

$$\mathbf{C}\eta/N \xrightarrow{p} \mathbf{0}, \quad \mathbf{C\Psi}^\top/N \xrightarrow{p} \mathbf{0}, \quad \mathbf{C\Lambda}^\top/N \to \mathbf{\Lambda}_0, \quad \mathbf{R} \to \mathbf{R}_0,$$

where $\mathbf{\Lambda}_0$ *and* $\mathbf{R}_0$ *are some nondegenerate matrices.*

Proof. The first two convergences in the statement of the lemma are immediate from Lemma 7 and the following two representations:

$$\mathbf{C}\eta = \sigma \begin{pmatrix} \sum c_1(s_i)w_i\xi_i \\ \sum c_2(s_i w_i\xi_i \end{pmatrix}, \quad \mathbf{C\Psi}^\top = \sigma \begin{pmatrix} \sum c_1(s_i)w_i k_i\xi_i & 0 \\ \sum c_2(s_i)w_i k_i \xi_i & 0 \end{pmatrix}.$$

Next, by Lemma 6

$$\mathbf{C\Lambda}^\top(\theta)/N = \begin{pmatrix} -V\frac{1}{N}\sum c_1(s_i)k_i & \frac{1}{N}\sum c_1(s_i) \\ -V\frac{1}{N}\sum c_2(s_i)k_i & \frac{1}{N}\sum c_2(s_i) \end{pmatrix} \to \begin{pmatrix} -V\mathbf{E}c_1(s)k(s) & \mathbf{E}c_1(s) \\ -V\mathbf{E}c_2(s)k(s) & \mathbf{E}c_2(s) \end{pmatrix} \equiv \mathbf{\Lambda}_0,$$

$$\mathbf{R} = \begin{pmatrix} 1 & -\sum c_1(s_i)/\sum c_2(s_i) \\ 0 & N/\sum c_2(s_i) \end{pmatrix} \to \begin{pmatrix} 1 & \mathbf{E}c_1(s)/\mathbf{E}c_2(s) \\ 0 & 1/\mathbf{E}c_2(s) \end{pmatrix} \equiv \mathbf{R}_0,$$

whereas the nondegeneracy of the limit matrices $\mathbf{\Lambda}_0$ and $\mathbf{R}_0$ is obvious in view of (4.5). □

Lemma 11. *If the conditions of Theorem* 4 *are satisfied then all claims of Lemma* 10 *hold and moreover*

$$\mathbf{U} \to \mathbf{U}_0 \quad \text{and} \quad \sigma^{-1}N^{-1/2}\mathbf{URC}\eta \Longrightarrow \Phi_2(0, \mathbf{I}), \tag{5.14}$$

where $\mathbf{U}_0$ *is some nondegenerate matrix.*

Proof. By (4.11), the entries of the matrix $\mathbf{U}$ are continuous functions of the entries of the covariance matrix of the random vector $N^{-1/2}\mathbf{RC}\eta$ which equals $\mathbf{RCT}_{D\eta}\mathbf{C}^\top\mathbf{R}^\top/N$. Hence, Lemmas 8 and 10 imply existence of the limit $\mathbf{U}_0$ of the matrices $\mathbf{U}$, and we can pass to the limit in (4.15):

$$\mathbf{I} \equiv \mathbf{URCT}_w^2\mathbf{C}^\top\mathbf{R}^\top\mathbf{U}^\top/N \to \sigma^{-2}\mathbf{U}_0\mathbf{R}_0\mathbf{V}_0\mathbf{R}_0^\top\mathbf{U}_0^\top = \mathbf{I}. \tag{5.15}$$

Furthermore, by Lemma 9 we have

$$\sigma^{-1}N^{-1/2}\mathbf{URC}\eta \Longrightarrow \Phi_2\big(0, \sigma^{-2}\mathbf{U}_0\mathbf{R}_0\mathbf{V}_0\mathbf{R}_0^\top\mathbf{U}_0^\top\big). \tag{5.16}$$

Comparing (5.16) and (5.15), we arrive at the second assertion in (5.14).

5.4. *Proof.* of Theorem 3. As follows from Theorem 1 of [4] (after changing the matrix $\mathbf{C}$ to $\mathbf{RC}$ in it), while proving Theorem 3 it suffices to verify the fulfillment of the conditions

$$(\mathbf{RC\Lambda})^{-1}\mathbf{RC}\eta \xrightarrow{p} \mathbf{0}, \quad (\mathbf{RC\Lambda})^{-1}\mathbf{RC\Psi}^\top \xrightarrow{p} \mathbf{0}. \tag{5.17}$$

We use Lemma 10. From the last assertion of this lemma we infer that the matrix $\mathbf{R}$ is nondegenerate from some number N onwards. Hence, the relations of (5.17) can be rewritten in the form

$$(\mathbf{C}\mathbf{\Lambda}^\top/N)^{-1}\mathbf{C}\eta/N \xrightarrow{p} \mathbf{\Lambda}_0^{-1}\mathbf{0} = \mathbf{0}, \quad (\mathbf{C}\mathbf{\Lambda}^\top/N)^{-1}\mathbf{C}\mathbf{\Psi}^\top/N \xrightarrow{p} \mathbf{\Lambda}_0^{-1}\mathbf{0} = 0.$$

□

Proof. of Theorem 4. In accordance with Corollary 1 and Remark 2 of [4], the conditions

$$\sigma^{-1}N^{-1/2}\mathbf{URC}\eta \Longrightarrow \Phi_2(0,\mathbf{I}), \quad \mathbf{URC}\mathbf{\Psi}^\top(\mathbf{RC}\mathbf{\Lambda}^\top)^{-1}\mathbf{U}^{-1} \xrightarrow{p} \mathbf{0} \tag{5.18}$$

guarantee (in our notations) the convergence of (4.13). It remains to note that validity of (5.18) is immediate from Lemmas 10 and 11 if we apply the assertions of these lemmas to the matrix products in (5.18).

5.5. We now establish several auxiliary assertions needed in derivation of Theorem 5. Given nonrandom t_1 and t_2, put

$$b(s) = t_1c_1(s) + t_2c_2(s), \quad b_i = b(s_i).$$

Lemma 12. *Suppose that condition* (A) *holds for*

$$A_0^2(s) = \big(c_1^2(s) + c_2^2(s)\big)(1 + w^2(s)).$$

Then

$$\frac{1}{N}\sum b_i^2w_i^2 \to \mathbf{E}b^2(s)w^2(s) \equiv b_0, \quad \frac{1}{N}\sum b_i^2k_i^2 \to \mathbf{E}b^2(s)k^2(s) < \infty, \tag{5.19}$$

$$\frac{1}{N}\sum b_i^2w_i^2\xi_i^2 \to b_0, \quad \frac{1}{N}\sum b_i^2k_i^2w_i^2\xi_i^2 \to \mathbf{E}b^2(s)k^2(s)w^2(s). \tag{5.20}$$

Proof. Both assertions in (5.19) ensue from Remark 12. From this fact and Lemma 5 with $a(\cdot) \equiv b^2(\cdot)w^2(\cdot)$ and $a(\cdot) \equiv b^2(\cdot)k^2(\cdot)w^2(\cdot)$ we find that

$$\frac{1}{N}\sum b_i^2w_i^2\big(\xi_i^2 - 1\big) \xrightarrow{p} 0, \quad \frac{1}{N}\sum b_i^2k_i^2w_i^2\big(\xi_i^2 - 1\big) \xrightarrow{p} 0. \tag{5.21}$$

From (5.21) and (5.19) we deduce (5.20). □

Lemma 13. *Under the assumptions of Theorem* 5*,*

$$\frac{1}{N}\sum b_i^2\eta_i^{*2} \xrightarrow{p} \sigma^2b_0.$$

Proof. We introduce the following notations:

$$\delta^* = \sqrt{\frac{1}{N}\sum b_i^2\eta_i^{*2}} - \sqrt{\frac{1}{N}\sum b_i^2\eta_i^2}, \quad \delta_c = \sqrt{\frac{1}{N}\sum b_i^2\eta_i^2} - \sigma\sqrt{b_0}.$$

Observe that by (2.8), (4.2), and (4.6)

$$\eta_i = \sigma w_i\xi_i, \quad \eta_i^* - \eta_i = (K^* - K)v_i/s_i - (V^* - V). \tag{5.22}$$

In particular, from the first equality in (5.22) and Lemma 12 we obtain

$$\delta_c \xrightarrow{p} 0. \tag{5.23}$$

We emphasize that the quantity $\sqrt{\frac{1}{N}a_i^2}$ possesses all properties of a norm for every sequence $\{a_1, \ldots, a_N\}$. From this fact and (5.22) we deduce the inequalities

$$|\delta^*| \le \sqrt{\frac{1}{N}\sum b_i^2(\eta_i^* - \eta_i)^2} \le |K^* - K|\sqrt{\frac{1}{N}\sum b_i^2v_i^2/s_i^2} + |V^* - V|, \tag{5.24}$$

$$\sqrt{\frac{1}{N}\sum b_i^2 v_i^2/s_i^2} \le \sqrt{\frac{1}{N}\sum b_i^2 k_i^2} + \sqrt{\frac{1}{N}\sum b_i^2 k_i^2 w_i^2 \xi_i^2}. \tag{5.25}$$

By Lemma 12, both summands on the right-hand side of (5.25) have finite limits. From this fact and (5.24) we find that

$$\delta^* \xrightarrow{p} 0, \tag{5.26}$$

for it merely suffices to recall that the estimators K^* and V^* are consistent because the assumptions of Theorem 5 are more stringent than those of Theorems 3 and 4.

Now, (5.23) and (5.26) imply the required conclusion of the lemma. □

Lemma 14. *Under the assumptions of Theorem* 5,

$$(\mathbf{C}\mathbf{T}_{\eta^{*2}}\mathbf{C}^\top)_{jk}/N \to (\mathbf{V}_0)_{jk}, \quad j,k = 1,2, \tag{5.27}$$

where $\mathbf{V}_0$ *is the matrix in Lemma* 8.

Proof. Setting $b(s) = c_j(s)$ in Lemma 13, we readily obtain the convergence in (5.27) for $k = j$. To prove the convergence for $j = 1$ and $k = 2$, we should use Lemma 13 with $b(s) = c_1(s) \pm c_2(s)$, observing that

$$c_1(s.)c_2(s.) = [(c_1(s.) + c_2(s.))^2 - (c_1(s.) - c_2(s.))^2]/4.$$

□

5.6. *Proof.* of Theorem 5. We introduce the matrix $\mathbf{U}^*$ that appears if in the definition (4.11) of the matrix $\mathbf{U}$ we replace the quantities b_K, b_{VV}, and ρ_{KV} with the respective quantities b_K^*, b_{VV}^*, and ρ_{KV}^*. Since the entries of the matrix $\mathbf{U}^*$ are continuous functions in the entries of the matrix $(\mathbf{C}\mathbf{T}_{\eta^{*2}}\mathbf{C}^\top)/N$, it is easy to see that $\mathbf{U}^* \xrightarrow{p} \sigma^{-1}\mathbf{U}_0$ by Lemma 14. From this fact and Lemma 11 we obtain

$$\sigma\mathbf{U}^*\mathbf{U}^{-1} \xrightarrow{p} \mathbf{U}_0\mathbf{U}_0^{-1} = \mathbf{I}. \tag{5.28}$$

Now, we note that the claim of Theorem 5 can be rephrased as

$$N^{-1/2}\mathbf{U}^*\mathbf{R}\mathbf{C}\mathbf{X}^\top(K^* - K, V^* - V)^\top \Rightarrow \mathbf{\Phi}_2(0,\mathbf{I}). \tag{5.29}$$

However, as follows from (3.7) in [4], the convergence in (5.29) amounts to

$$\sigma\mathbf{U}^*\mathbf{U}^{-1}(\sigma^{-1}N^{-1/2}\mathbf{U}\mathbf{R}C\eta) \Rightarrow \Phi_2(0,\mathbf{I}). \tag{5.30}$$

Relation (5.30) now follows from (5.28) and Lemma 11, as the assumptions of Theorem 5 are more stringent than those of Theorem 4. □

6. Proofs of Theorem 6 and Corollary 4

6.1. We begin with proving two auxiliary assertions.

Lemma 15. *Suppose that a sequence of numbers* $\{A_i\}$ *satisfies the conditions*

$$d_N^2 \equiv b_K/a_K^2 = \sum c_{Ki}^2 w_i^2 \Big/ \Big(\sum c_{Ki}k_i\Big)^2 \to 0, \tag{6.1}$$

$$d_N^2 \frac{1}{N}\sum M_i^2 A_i^2 w_i^2 \to 0. \tag{6.2}$$

Then

$$\tau_{1,N} \equiv \frac{1}{\sqrt{N}}\sum \delta_{ji}A_i w_i \xi_i \xrightarrow{p} 0, \quad j = 1,2.$$

Lemma 16. *Suppose that a sequence of numbers* $\{A_i\}$ *satisfies* (6.1) *and*

$$d_N \frac{1}{N} \sum M_i|A_i| \to 0. \tag{6.3}$$

Then

$$\tau_{2,N} \equiv \frac{1}{N} \sum \delta_{ji} A_i \xrightarrow{p} 0, \quad j = 1, 2.$$

Proof. of Lemmas 15 and 16 We will apply the results of [3], wherein a similar estimator K^* was studied. We bind the quantities in [3] (writing them on the left-hand sides) with those in this article as follows:

$$a_i = 1, \quad b_i = 1/s_i, \quad c_i = c_{Ki}, \quad \alpha_i = k_i, \quad \beta_i = \sigma w_i, \quad d_c = \sigma d_N,$$

$$X_i = v_i, \quad \theta = K, \quad \theta^* = K^*, \quad \gamma_i(\theta) = A_i \gamma(s_i, K), \quad K_i = A_i M_i.$$

Inspecting § 5 of [3], we can easily see that it is possible to define an estimator $\widetilde{K}$ analogous to $\widetilde{\theta}$ of [3] so that it possesses the following properties:

$$\widetilde{K} = K^*, \quad \text{provided that} \quad K/2 \le K^* \le 2K;$$

$$|A_i \delta_{ji}| \le |A_i| M_i \cdot |\widetilde{K} - K|, \quad \text{provided that } \widetilde{K} = K^*; \tag{6.4}$$

$$\mathbf{P}(\widetilde{K} \ne K^*) \to 0 \quad \text{as } d_N \to 0. \tag{6.5}$$

Moreover, in this case the assertions of Lemmas 8 and 12 of [3] can be restated as follows:

$$\mathbf{E}|\widetilde{K} - K|^2 \le 16\sigma^2 d_N^2, \tag{6.6}$$

$$\mathbf{E}\{|\tau_{1,N}| : \widetilde{K} = K^*\} \le 17\sigma d_N \sqrt{\frac{1}{N} \sum M_i^2 A_i^2 w_i^2}. \tag{6.7}$$

The claim of Lemma 15 is immediate from (6.5) and (6.6). To prove Lemma 16, we note that by (6.4)

$$|\tau_{2,N}| \le \frac{1}{N} \sum |A_i| M_i \cdot |\widetilde{K} - K|, \quad \text{provided that} \quad \widetilde{K} = K^*.$$

Using (6.6), whence we infer

$$\mathbf{E}\{|\tau_{2,N}| : \widetilde{K} = K^*\} \le \frac{1}{N} \sum |A_i| M_i \sqrt{\mathbf{E}|\widetilde{K} - K|^2} \le 4\sigma d_N \frac{1}{N} \sum |A_i| M_i.$$

The last inequality and (6.5) yield the conclusion of Lemma 16. □

6.2. We need several consequences of the above assertions.

Lemma 17. *If the conditions of Theorem* 6 *are satisfied then there exists a number* d_0 *such that*

$$N d_N^2 \equiv N b_K / a_K^2 \to d_0^2.$$

Proof. By (4.4), (4.5), and Lemma 6 we have

$$\frac{a_K}{N} = \frac{1}{N} \sum c_1(s_i) k_i - \frac{1}{N} \sum c_2(s_i) k_i \cdot \frac{\sum c_1(s_i)}{\sum c_2(s_i)} \to \mathbf{E}\frac{c_1(s)}{K+s} - \mathbf{E}\frac{c_2(s)}{K+s} \cdot \frac{\mathbf{E}c_1(s)}{\mathbf{E}c_2(s)} \ne 0. \tag{6.8}$$

Since $\sigma^2 b_K$ equals the variance of the first coordinate of the vector $\mathbf{R}C\eta$, it follows from Lemmas 8 and 10 that $\sigma^2 b_K / N \to (\mathbf{R}_0 \mathbf{V}_0 \mathbf{R}_0^\top)_{11}$. From this convergence and (6.8) we infer the conclusion of the lemma. □

Lemma 18. *If the conditions of Theorem* 6 *are satisfied then all conditions and conclusions of Lemmas* 15 *and* 16 *hold for $A_i \equiv 1$ and $A_i \equiv k_i$. In particular,*

$$\frac{1}{N}\sum \delta_{ji} v_i/s_i \xrightarrow{p} 0, \quad \frac{1}{\sqrt{N}}\sum \delta_{ji} w_i \xi_i \xrightarrow{p} 0,$$

$$\frac{1}{N}\sum \delta_{ji} \xrightarrow{p} 0, \quad \gamma_j^* \equiv \frac{1}{N}\sum \gamma_j(s_i, K^*) \xrightarrow{p} \mathbf{E}\gamma_j(s, K), \quad j = 1, 2.$$

Proof. The convergence $d_N^2 \to 0$ is immediate from the preceding lemma. Furthermore, by the conditions of Theorem 6 and Remark 12

$$\frac{1}{N}\sum M_i^2 w_i^2 k_i^2 \to \mathbf{E}M^2(s)w^2(s)k^2(s), \quad \frac{1}{N}\sum M_i^2 w_i^2 \to \mathbf{E}M^2(s)w^2(s),$$

$$\frac{1}{N}\sum M_i k_i \to \mathbf{E}M(s)k(s), \quad \frac{1}{N}\sum M_i \to \mathbf{E}M(s).$$

Hence, in view of (6.1) we have (6.2) and (6.3). To derive the last assertions of the lemma, it suffices to observe that

$$\sum \delta_{ji} v_i/s_i = V\sum \delta_{ji} k_i + \sigma \sum \delta_{ji} w_i k_i \xi_i,$$

$$\gamma_j^* \equiv \frac{1}{N}\sum \gamma_j(s_i, K^*) = \frac{1}{N}\sum \delta_{ji} + \frac{1}{N}\sum \gamma_j(s_i, K).$$

□

6.3. We now prove two lemmas that play a key role in derivation of Theorem 6.

Lemma 19. *Under the assumptions of Theorem* 6*,*

$$\mathbf{\Gamma}(K)\eta/N \xrightarrow{p} \mathbf{0}, \quad \mathbf{\Gamma}(K)\mathbf{\Psi}^\top/N \xrightarrow{p} \mathbf{0}, \quad \mathbf{\Gamma}(K)\mathbf{\Lambda}^\top/N \to \widetilde{\mathbf{\Lambda}}_0,$$

$$\mathbf{R}(K) \to \widetilde{\mathbf{R}}_0, \quad \widetilde{\mathbf{U}} \to \widetilde{\mathbf{U}}_0 \quad \text{and} \quad N^{-1/2}\widetilde{\mathbf{U}}\mathbf{R}(K)\mathbf{\Gamma}(K)\eta \Longrightarrow \Phi_2(0, \mathbf{I}),$$

where $\widetilde{\mathbf{U}}_0$, $\widetilde{\mathbf{\Lambda}}_0$, and $\widetilde{\mathbf{R}}_0$ are some nondegenerate matrices.

This statement is immediate from Lemmas 10 and 11 if we replace the functions $c_j(\cdot)$ in them with $\gamma_j(\cdot, K)$, $j = 1, 2$.

Lemma 20. *Under the assumptions of Theorem* 6*,*

$$N^{-1/2}\mathbf{\Delta}\eta \xrightarrow{p} \mathbf{0}, \quad N^{-1}\mathbf{\Delta}\mathbf{X}^\top \xrightarrow{p} \mathbf{0}, \quad \mathbf{R}(K^*) \xrightarrow{p} \widetilde{\mathbf{R}}_0.$$

This statement is immediate from Lemma 18 if we observe that

$$\mathbf{R}(K^*) = \begin{pmatrix} 1 & -\gamma_1^*/\gamma_2^* \\ 0 & \gamma_2^* \end{pmatrix}, \quad \mathbf{\Delta}\eta = \sigma \begin{pmatrix} \sum \delta_{1i} w_i \xi_i \\ \sum \delta_{2i} w_i \xi_i \end{pmatrix},$$

$$\mathbf{\Delta}X^\top = \begin{pmatrix} -\sum \delta_{1i} v_i/s_i & \sum \delta_{1i} \\ -\sum \delta_{2i} v_i/s_i & \sum \delta_{2i} \end{pmatrix}.$$

6.4. *Proof.* of Theorem 6. As shown in Theorem 3 of [4], it suffices to check the fulfillment of the following conditions:

$$\mathbf{W}_0 \equiv \sigma^{-1}N^{-1/2}\widetilde{\mathbf{U}}\mathbf{R}(K)\mathbf{\Gamma}(K)\eta \Longrightarrow \Phi_2(0, \mathbf{I}), \tag{6.9}$$

$$\mathbf{W}_1 \equiv \sigma^{-1}N^{-1/2}\widetilde{\mathbf{U}}(\mathbf{R}(K^*)\mathbf{\Gamma}(K^*) - \mathbf{R}(K)\mathbf{\Gamma}(K))\eta \xrightarrow{p} \mathbf{0}. \tag{6.10}$$

$$\mathbf{W}_2 \equiv \widetilde{\mathbf{U}}\mathbf{R}(K^*)\mathbf{\Gamma}(K^*)\mathbf{X}^\top(\mathbf{\Gamma}(K)\mathbf{\Lambda}^\top)^{-1}(\mathbf{R}(K))^{-1}\widetilde{\mathbf{U}}^{-1} \xrightarrow{p} \mathbf{I}, \tag{6.11}$$

Condition (6.9) is immediate from Lemma 19. This condition and Lemma 20 yield (6.10), since

$$\mathbf{W}_1 = \sigma^{-1}\widetilde{\mathbf{U}}\mathbf{R}(K^*)(N^{-1/2}\mathbf{\Delta}\eta) + \widetilde{\mathbf{U}}(\mathbf{R}(K^*)\mathbf{R}^{-1}(K) - \mathbf{I})\widetilde{\mathbf{U}}^{-1}\mathbf{W}_0.$$

To check the fulfillment of (6.11), it suffices to apply Lemmas 19 and 20 to the matrix product and observe that

$$\mathbf{\Gamma}(K^*)\mathbf{X}^\top/N = \mathbf{\Gamma}(K)\mathbf{\Lambda}^\top/N + \mathbf{\Gamma}(K)\mathbf{\Psi}^\top/N + \mathbf{\Delta}X^\top/N \xrightarrow{p} \widetilde{\mathbf{\Lambda}}_0 + \mathbf{0} + \mathbf{0}.$$

6.5. Corollary 4 is obvious from the following

Lemma 21. *Conditions* (4.8) *imply validity of* (4.5).

Proof. Put $f(x) = x^{-1}(K+x)^{-1}c(x) > 0$. Then in view of (4.7) the first condition in (4.5) can be rewritten in the form

$$\delta \equiv [\mathbf{E}(K+s)f(s)][\mathbf{E}sf(s)] - [\mathbf{E}s(K+s)f(s)][\mathbf{E}f(s)] \tag{1}$$

$$= [\mathbf{E}sf(s)]^2 - [\mathbf{E}s^2f(s)][\mathbf{E}f(s)] \tag{2}$$

$$= [\mathbf{E}(sf^{1/2}(s))f^{1/2}(s)]^2 - [\mathbf{E}(sf^{1/2}(s))^2][\mathbf{E}(f^{1/2}(s))^2] \neq 0. \tag{3}$$

Thus, the equality $\delta = 0$ may hold only if

$$\mathbf{P}(sf^{1/2}(s) = tf^{1/2}(s)) = 1$$

for some nonrandom t. But this contradicts (4.8).

If we put $c(x) = \gamma(x, K)$ in Lemma 21 then we come to the assertion mentioned in Remark 9. □

7. Concluding Remarks

Remark 13. It must be emphasized that, due to extensive practical use of the Michaelis–Menten equation, there are many publications devoted to the problem of estimating the parameters of this equation, and many methods were proposed for estimating the unknown parameters in this problem which is very important from a practical viewpoint (see, for instance, [1, 2, 5, 6, 10–14], and the references therein). Among the found estimators we distinguish the following (see, for instance, [2]):

$$\widetilde{V} = \frac{\sum v_i^2/s_i^2 \sum v_i^2 - (\sum v_i^2/s_i)^2}{\sum v_i^2/s_i^2 \sum v_i - \sum v_i^2/s_i \sum v_i/s_i}, \tag{7.1}$$

$$\widetilde{K} = \frac{\sum v_i^2 \sum v_i/s_i - \sum v_i^2/s_i \sum v_i}{\sum v_i^2/s_i^2 \sum v_i - \sum v_i^2/s_i \sum v_i/s_i}. \tag{7.2}$$

They are most close to the estimators that we propose in the present article.

However, as far as the authors know, none of these methods has a satisfactory theoretical justification. A detailed study of the methods and comparison with the estimators obtained by the method of this article will be carried out in a forthcoming article of the authors. For example, we will demonstrate that the estimators (7.1) and (7.2), known as Johansen–Lavry estimators, despite their explicit form, have an essential shortcoming in comparison to the explicit estimators that we constructed in the present article. Namely, the estimator $\widetilde{K}$ of the parameter K is unbiased only in one particular case, whereas the estimator $\widetilde{V}$ for the parameter V is always biased.

Remark 14. Throughout the article, K and V are unknown parameters. Moreover, the variances $\{\sigma_i^2\}$ may be unknown too. Thus, most conditions in all assertions of the article are restrictions on quantities which contain unknown parameters. It is clear that for practical application of these assertions we must verify all such conditions *for all possible values of all unknown parameters* (as this is done, for instance, in [9]).

References

[1] Michaelis L and Menten M. L., *Die Kinetik der Invertinwirkung.* Biochem. Z., 49 (1913), 333. (Translated from the Russian publication of 1989.)

[2] Cornish-Bowden A., *Fundamentals of Enzyme Kinetics.* Butterworth, London, 1979.

[3] Linke Yu. Yu. and Sakhanenko A. I. *Asymptotically normal estimation of a parameter in a linear-fractional regression problem.* Sibirsk. Mat. Zh., 41, iss.1 (2000), 150–163.

[4] Linke Yu. Yu. and Sakhanenko A. I. *Asymptotically normal estimation of a multidimensional parameter in the linear-fractional regression problem.* Sibirsk. Mat. Zh., 42,iss.2 (2001), 372–388.

[5] Zivin J. A. and Waud D. R. *How to analyze binding, enzyme and uptake data: The simplest case, a single phase.* Life Sciences, 30 (1982), 1407–1422.

[6] Currie D. J. *Estimating Michaelis–Menten parameters: Bias, variance and experimental design.* Biometrics, 38 (1982), 907–919.

[7] *The Mathematical Theory for Gliding Experiment [in Russian].* Nauka, Moscow, 1983.

[8] Borovkov A. A. *Mathematical Statistic [in Russian].* Nauka and Sobolev Institute Press, Moscow, 1997.

[9] Borovkov A. A. *Probability [in Russian].* Nauka, Moscow, 1986.

[10] Atkins G. L. and Nimmo I. A. *A comparison of seven methods for fitting the Michaelis–Menten equation.* Biochemical J., 149 (1975), 775–777.

[11] Cornish-Bowden A. and Eisenthal R. *Statistical consideration in the estimation of enzyme kinetic parameters by direct linear plot and other methods.* Biochemical J., 139 (1974), 721–730.

[12] Marquardt D. W. *An algorithm for least-squares estimation of nonlinear parameters.* J. Soc. Ind. Appl. Math., 2 (1963), 431–441.

[13] Raaijmakers J. *Statistical analysis of the Michaelis–Menten equation.* Biometrics, 43 (1987), 793–803.

[14] Wilkinson G. N. *Statistical estimations in enzyme kinetics.* Biochemical J., 30 (1961), 324–332.