MATHEMATICAL COMMUNICATIONS
Math. Commun. **22**(2017), 25–40

Generalization of Cramér-Rao and Bhattacharyya inequalities for the weighted covariance matrix

Mark Kelbert[1,2] and Pavel Mozgunov[1,3]*

[1] *Higher School of Economics, National Research University, 20 Myasnitskaya str., Moscow, 101 000, RF*
[2] *Department of Mathematics, Swansea University, Singleton Park, Swansea, SA2 8PP, UK*
[3] *Department of Mathematics and Statistics, Lancaster University, Lancaster, LA1 4YF, UK*

Received December 21, 2015; accepted May 20, 2016

Abstract. The paper considers a family of probability distributions depending on a parameter. The goal is to derive the generalized versions of Cramér-Rao and Bhattacharyya inequalities for the weighted covariance matrix and of the Kullback inequality for the weighted Kullback distance, which are important objects themselves [9, 23, 28]. The asymptotic forms of these inequalities for a particular family of probability distributions and for a particular class of continuous weight functions are given.

AMS subject classifications: 94A17, 62B10, 62C10

Key words: weighted covariance matrix, weighted Fisher information, Cramér-Rao inequality, Bhattacharyya inequality, Kullback inequality

1. Introduction

The concept of the discrete weighted mean can be extended to the concept of the weighted mean of continuous functions [16] which, for instance, plays an important role in the systems of weighted differential and integral calculus [13]. The corresponding weighted covariance matrix naturally arises in the problem of continuous weighted mean [23]. Generally, the need of the weight function in statistics appears when observations cannot be considered as equivalent or when the estimation of parameter is especially sensitive in a neighbourhood of some value. Regarding the former one, in financial studies it is common to judge the information from recent events as more valuable than from remote ones. This judgement can be taken into account by the Pearson weighted correlation matrix [23]. On the other hand, the problem of sensitive estimation is very common in statistics as well, i.e., in medical dose finding studies it is necessary to find the dose which has the closest probability of toxicity to the specific value γ. Thus, costs of the wrong estimation are greater in a neighbourhood of this value of a specific interest γ, which usually lies in the interval $(0.2, 0.33)$ (see [17, 6, 2]). The weighted covariance matrix naturally arises

*Corresponding author. *Email addresses:* `mark.kelbert@gmail.com` (M. Kelbert), `pmozgunov@gmail.com` (P. Mozgunov)

in both cases [23, 21] as well as in a variety of other fields, i.e., in the weighted principal component analysis [9], in the comparison of two multiple regressions by means [24], etc. Regarding the discussion above, the goal of this paper is to derive a weighted version of the Bhattacharyya [3] and Cramér-Rao [4] inequalities for the weighted covariance matrix.

It is natural to assume that a bound of the weighted covariance matrix will involve the recently studied weighted version of the Fisher information matrix [21] which is connected to the weighted Kullback-Leibler divergence [28] (in a similar way the unweighted Fisher information is connected to the unweighted Kullback-Leibler divergence [8]). Moreover, it is shown in [20] that weighted versions of measures of information which have already attracted considerable attention in the literature [7, 10, 14, 25, 26, 27] arise in the problems of sensitive estimation. Consider the following example of the dose-escalation study as an illustration of sensitive estimation. A statistician has a range of doses of the drug and after a series of toxic and non-toxic realizations he needs to estimate the probability of toxicity of this dose. The goal of this study is to find the dose whose probability is the closest to the target probability γ. Therefore, if he wrongly declares that the probability of a toxicity for particular dose is in a specified small neighbourhood of particular value γ, the penalty for this error should be more severe than for a similar mistake far from the sensitive area. This penalty arises as the result of the declared dose is investigated in further studies and if the true probability is greater than target γ, one will observe a lot of toxic responses and otherwise, a lot of patients will be undertreated which causes additional costs. It was shown in [21] that weighted measures of information naturally arise in this framework.

Consider the family of RVs $\mathbf{Z}_\theta \in \mathbb{R}^d$ with PDF f_θ, where $\theta \in \Theta \subset \mathbb{R}^m$ is the vector of parameters of PDF f_θ and $\mathbf{z} = [z_1, \dots, z_d]^{\mathrm{T}}$. Let $T(\mathbf{Z}) = T(Z_1, \dots, Z_d)$ be an arbitrary estimator of $\tau(\theta)$, where $\tau(\theta)$ is a preassigned function of parameter θ. Let $\phi(\mathbf{z}, \theta, \gamma)$ be a continuous and positive weight function such that

$$\int_{\mathbb{R}^d} f_\theta(\mathbf{z})\phi(\mathbf{z}, \theta, \gamma)\mathrm{d}\mathbf{z} = 1. \tag{1}$$

In this paper, we consider the class of weight functions which can be represented in the following form

$$\phi(\mathbf{z}, \theta, \gamma) = \frac{1}{\kappa(\theta, \gamma)}\tilde{\phi}(\mathbf{z}, \gamma). \tag{2}$$

Here $\kappa(\theta, \gamma) \in \mathbf{C}^k$, where $\mathbf{C}^k$ is the family of functions with continuous derivatives up to order k (k will be specified below), and $\kappa(\theta, \gamma)$ is found from the normalizing condition (1). Note that condition (1) can be rewritten in the following form

$$\int_{\mathbb{R}^d} \tilde{\phi}(\mathbf{z}, \gamma) f_\theta(\mathbf{z})\mathrm{d}\mathbf{z} = \kappa(\theta, \gamma) \tag{3}$$

where $\tilde{\phi}(\mathbf{z}, \gamma)$ is a function that does not depend on θ. Let $\mathbb{E}_\theta^\phi(T(\mathbf{Z}))$ be the weighted expectation of $T(\mathbf{Z})$

$$g(\theta) \equiv \mathbb{E}_\theta^\phi(T(\mathbf{Z})) = \int_{\mathbb{R}^d} T(\mathbf{z}) f_\theta(\mathbf{z})\phi(\mathbf{z})\mathrm{d}\mathbf{z}, \tag{4}$$

and let $\mathbb{E}_\theta(T(\mathbf{Z}))$ be the classic expectation of $T(\mathbf{Z})$

$$e(\theta) \equiv \mathbb{E}_\theta(T(\mathbf{Z})) = \int_{\mathbb{R}^d} T(\mathbf{z}) f_\theta(\mathbf{z}) \mathrm{d}\mathbf{z}. \tag{5}$$

Recall the definition of the weighted version of a quantitative measure of information gain, [1, 10, 14, 18, 21]. The weighted $(m \times m)$ Fisher information matrix

$$\mathrm{I}^\phi(\theta) = \mathbb{E}_\theta^\phi \left[(\nabla \log f_\theta(\mathbf{Z})) (\nabla \log f_\theta(\mathbf{Z}))^{\mathrm{T}} \right], \tag{6}$$

where ∇ is the notation for the gradient (the vector $\nabla \log f_\theta(\mathbf{Z})$ is the *score*), and the weighted Kullback-Leibler divergence of g from f [28]

$$\mathbb{D}^\phi(f||g) = \int_{\mathbb{R}^d} \phi(\mathbf{z}) f(\mathbf{z}) \log \frac{f(\mathbf{z})}{g(\mathbf{z})} \mathrm{d}\mathbf{z}. \tag{7}$$

For simplicity, we assume that the inverse Fisher matrix exists. However, in a general case, we understand under inverse the Moore-Penrose pseudoinverse. It is shown that it is more convenient to study the calibrated Kullback-Leibler divergence defined in [28]:

$$K^\phi(f||g) = \int_{\mathbb{R}^d} \tilde{\phi}(\mathbf{z}) \frac{f(\mathbf{z})}{\kappa_f} \log \frac{f(\mathbf{z})\kappa_g}{g(\mathbf{z})\kappa_f} \mathrm{d}\mathbf{z} = \mathbb{D}(\tilde{f}||\tilde{g}), \tag{8}$$

where κ_f and κ_g are normalizing constants for the weight function $\tilde{\phi}$ and PDFs f and g, respectively, $\tilde{f} = \tilde{\phi}(\mathbf{z}) f(\mathbf{z}) \kappa_f^{-1}$ and $\mathbb{D}(f||g)$ is the standard Kullback-Leibler divergence of g from f

$$\mathbb{D}(f||g) = \int_{\mathbb{R}^d} f(\mathbf{z}) \log \frac{f(\mathbf{z})}{g(\mathbf{z})} \mathrm{d}\mathbf{z}. \tag{9}$$

The main goal of this paper is to derive the generalization of the Bhattacharyya and Cramér-Rao inequalities for the weighted covariance matrix of $T(\mathbf{Z})$

$$\mathbb{V}_\theta^\phi(T(\mathbf{Z})) \equiv \mathbb{E}_\theta^\phi \left[(T(\mathbf{Z}) - e(\theta))(T(\mathbf{Z}) - e(\theta))^{\mathrm{T}} \right] \tag{10}$$

and consider the weighted Kullback-Leibler divergence derived [28] for the class of two close distributions.

We will use the Beta and Dirichlet distributions as examples throughout the paper as these distributions are quite common in problems of sensitive estimation in medical statistics, i.e., see [5, 11]. As a unidimensional example, let us consider the family of RVs $Z_\alpha^{(n)}$ with PDF $f_\alpha^{(n)}$

$$f_\alpha^{(n)}(p) = \frac{1}{\mathbb{B}(\alpha n + 1, n - \alpha n + 1)} p^{\alpha n} (1-p)^{(1-\alpha)n}, \quad 0 \le p \le 1, \tag{11}$$

where α is the parameter of distribution. Note that as $n \to \infty$, $\mathbb{E}(Z_\alpha^{(n)}) = \alpha + O\left(\frac{1}{n}\right)$ and $\mathbb{V}(Z_\alpha^{(n)}) = O\left(\frac{1}{n}\right)$, where $\mathbb{E}(Z)$ and $\mathbb{V}(Z)$ are expectation and variance of RV Z, respectively. We would like to find the explicit asymptotic expansions (as $n \to \infty$) for the lower bound of $V^{\phi}(Z_{\alpha}^{(n)})$ in the case of the following weight function

$$\phi^{(n)}(p) = \frac{1}{\kappa(\alpha,\gamma)} p^{\gamma\sqrt{n}}(1-p)^{(1-\gamma)\sqrt{n}}, \tag{12}$$

where $\kappa(\alpha,\gamma)$, is found from condition (1). As a multidimensional example, let us consider the family of Dirichlet RVs $Z_{\beta_1,\beta_2}^{(n)}$ with PDF $f_{\beta_1,\beta_2}^{(n)}$

$$f_{\beta_1,\beta_2}^{(n)} = \frac{1}{\mathcal{B}(\beta_1,\beta_2)} p_1^{\beta_1 n} p_2^{\beta_2 n}(1-p_1-p_2)^{(1-\beta_1-\beta_2)n}, \quad 0 \le p_1, p_2 \le 1, \tag{13}$$

where β_1 and β_2 are parameters of the distribution and

$$\mathcal{B}(\beta_1,\beta_2) = \frac{\Gamma(\beta_1 n+1)\Gamma(\beta_2 n+1)\Gamma(n-\beta_1 n-\beta_2 n+1)}{\Gamma(n+3)}.$$

As before, we would like to find the explicit asymptotic expansions for the lower bound of $V^{\phi}(T(\mathbf{Z}_{\beta_1,\beta_2}^{(n)}))$ in the case of the following weight function

$$\hat{\phi}^{(n)}(p) = \frac{1}{\hat{\kappa}(\beta_1,\beta_2,\gamma_1,\gamma_2)} p_1^{\gamma_1\sqrt{n}} p_2^{\gamma_2\sqrt{n}}(1-p_1-p_2)^{(1-\gamma_1-\gamma_2)\sqrt{n}},$$

where $\kappa(\beta_1,\beta_2,\gamma_1,\gamma_2)$ is defined in (50).

2. Main results

Recall that $\mathbf{Z}_\theta \in \mathbb{R}^d$ is the family of RVs with PDF f_θ, where $\theta \in \Theta \subset \mathbb{R}^m$ is the vector of parameters of PDF f_θ. Let $\phi(\mathbf{z},\theta,\gamma)$ be the continuous positive weight function defined in (2), $\mathrm{I}^{\phi}(\theta)$ the weighted Fisher information $(m\times m)$ matrix given in (6), $g(\theta)$ the weighted expectation given in (4) and $\mathbb{V}_\theta^{\phi}(T(\mathbf{Z}))$ the weighted covariance matrix of $T(\mathbf{Z})$ given in (10). We assume that in (4) and (1) differentiation w.r.t. the parameters up to order to be considered under the sign of integration is valid. A sufficient condition for this is that the integrand after the operation of differentiation $\eta(\theta)$ is bounded by an integrable function χ which does not depend on θ

$$|\eta(\theta)| \le \chi,$$

i.e., the integral converges uniformly in θ. Also, let us denote the partial derivative of order j

$$f^{(j)} = \frac{\partial^j f}{\partial \theta^j}.$$

Theorem 1 (Weighted Bhattacharyya inequality, uniparametric case)**.** *Let θ be a scalar parameter, and $\tau(\theta)$ a preassigned scalar function of parameter θ. An unbiased estimator of $\tau(\theta)$ is a scalar function $T(\boldsymbol{Z})$ such that*

$$e(\theta) = \mathbb{E}_\theta[T(\boldsymbol{Z})] = \tau(\theta).$$

Consider the weight function that satisfies condition (1). Recall

$$g(\theta) \equiv \int_{\mathbb{R}^d} T(\boldsymbol{z})\phi(\boldsymbol{z},\theta,\gamma) f_\theta(\boldsymbol{z})\mathrm{d}\boldsymbol{z}. \tag{14}$$

Assume that integrands in (14) and (1) converge uniformly in θ after operation of differentiation up to order ν. Then the following inequality for the weighted variance of $T(\boldsymbol{Z})$ holds

$$\mathbb{V}_\theta^\phi(T(\boldsymbol{Z})) \geq \sum_{i,j=1}^{\nu} \left(g^{(i)}(\theta) - \mathrm{Q}_1^i + \tau \mathrm{Q}_2^i\right)\left(g^{(j)}(\theta) - \mathrm{Q}_1^j + \tau \mathrm{Q}_2^j\right) \mathrm{J}_{ij}^\phi, \tag{15}$$

where Q_i^j, $i = 1, 2$ are given in (35) and (37), respectively, and J_{ij}^ϕ are the elements of the matrix $\mathbb{J}^\phi$ defined in (33).

Remark 1. *Note that this inequality includes the weighted version of the Cramér-Rao inequality. One can obtain the Cramér-Rao inequality when $\nu = 1$. In this particular case*

$$\left(J_{11}^\phi\right)^{-1} \equiv \mathrm{I}^\phi(\theta) = \int_{\mathbb{R}^d} (f_\theta')^2 f_\theta^{-1} \phi \mathrm{d}\boldsymbol{z},$$
$$\int_{\mathbb{R}^d} \phi f_\theta^{(j)} \mathrm{d}\boldsymbol{z} = \frac{\kappa'(\theta,\gamma)}{\kappa(\theta,\gamma)}$$

and

$$\int_{\mathbb{R}^d} T(\boldsymbol{z})\phi f_\theta^{(1)} \mathrm{d}\boldsymbol{z} = g'(\theta) + \frac{\kappa'(\theta,\gamma)}{\kappa(\theta,\gamma)} g(\theta).$$

Thus, we obtain the following inequality

$$\begin{aligned} \mathbb{V}_\theta^\phi(T(\boldsymbol{Z})) \geq & \left(\frac{\partial g(\theta)}{\partial\theta} - \frac{\kappa'(\theta,\gamma)}{\kappa(\theta,\gamma)}\left(e(\theta) - g(\theta)\right)\right) \mathrm{I}^\phi(\theta)^{-1} \\ & \times \left(\frac{\partial g(\theta)}{\partial\theta} - \frac{\kappa'(\theta,\gamma)}{\kappa(\theta,\gamma)}\left(e(\theta) - g(\theta)\right)\right)^{\mathrm{T}}. \end{aligned} \tag{16}$$

It is easy to see that if $\phi(\boldsymbol{z},\theta,\gamma) \equiv 1$, then inequality (16) gives the well-known Cramér-Rao lower bound for the unweighted variance.

Example 1. *Consider RV $Z_\alpha^{(n)}$ with PDF $f_\alpha^{(n)}$ given in (11) with $x = \alpha n$, where $0 < \alpha < 1$. When $\nu = 1$, $\theta = \alpha$, $T(\boldsymbol{Z}) = \boldsymbol{Z}_\alpha^{(n)}$ for the weight function $\phi^{(n)}(p)$, inequality (15) takes the following form*

$$\mathbb{V}^\phi(T(\boldsymbol{Z}_\alpha^{(n)})) \geq \frac{\alpha(1-\alpha) + (\alpha-\gamma)^2}{n} + \frac{-2\alpha + \alpha^2 + \gamma + 2\alpha\gamma - 2\gamma^2}{n^{3/2}} + O\left(\frac{1}{n^2}\right). \tag{17}$$

Due to cumbersome computations, all details are given in Section 3.2.

Theorem 2 (Weighted Bhattacharyya inequality, multiparametric case)**.** *Let $\theta \in \Theta \subset R^m$ be a vector of parameters, $\tau(\theta) = (\tau_1(\theta), \ldots, \tau_l(\theta))^{\mathrm{T}} \in \mathbb{R}^l$ the preassigned vector function of parameter θ and $T(\boldsymbol{Z})$ an unbiased estimate of $\tau(\theta)$:*

$$e(\theta) = \mathbb{E}_\theta(T) = \int_{R^d} T(\boldsymbol{z}) f_\theta(\boldsymbol{z}) \mathrm{d}\boldsymbol{z} = \tau(\theta).$$

Consider the weight function $\phi(\boldsymbol{z}, \theta, \gamma)$ such that condition (1) holds. Assume that the following positively definite matrix exists

$$I^\phi = \mathbb{E}_\theta^\phi [\beta\beta^{\mathrm{T}}], \tag{18}$$

where

$$\beta = (\beta_1(\theta), \ldots, \beta_r(\theta))^{\mathrm{T}}$$

is r-dimensional RV, components of which are all possible expressions of the following form

$$\frac{1}{f_\theta(\boldsymbol{Z})} \frac{\partial^{i_1,\ldots,i_m}}{\partial\theta_1^{i_1}, \ldots \partial\theta_m^{i_m}} f_\theta(\boldsymbol{Z}), \tag{19}$$

where $(1 \leq i_1 + \ldots + i_m \leq s)$ and r is the total number of all these expressions.

Let $\mathbb{F}^\phi$ be the $(r \times l)$ matrix whose rows have the following form

$$\int_{\mathbb{R}^d} (T(\boldsymbol{z}) - \tau(\theta)) \, \phi(\boldsymbol{z}, \theta, \gamma) \frac{\partial^{i_1,\ldots,i_m}}{\partial\theta_1^{i_1}, \ldots \partial\theta_m^{i_m}} f_\theta(\boldsymbol{z}) \mathrm{d}\boldsymbol{z} \tag{20}$$

numbered in the same order as expressions (19). Assume that integrands in (20) and (1) converge uniformly in θ after the operation of differentiation. Then the following inequality for a weighted variance of T holds

$$\mathbb{V}_\theta^\phi(T) \geq (\mathbb{F}^\phi)^{\mathrm{T}} I^\phi(\theta)^{-1} \mathbb{F}^\phi. \tag{21}$$

Remark 2. *Here and below for $(d \times d)$ matrices of the same dimension d, $\mathbb{A}$ and $\mathbb{B}$, the inequality*

$$\mathbb{A} \geq \mathbb{B}$$

means that

$$\mathbb{C} = \mathbb{A} - \mathbb{B}$$

is a non-negative definite matrix.

Example 2. *Consider a random vector $\boldsymbol{Z}^{(n)}_{\beta_1,\beta_2} = (Z_1^{(n)}, Z_2^{(n)})^{\mathrm{T}}$ with PDF $f^{(n)}_{\beta_1,\beta_2}$ given in (13), where $0 < \beta_1, \beta_2 < 1$ and $\tau = \beta_1/\beta_2$. If*

$$T(\boldsymbol{Z}^{(n)}_{\beta_1,\beta_2}) = \frac{Z_1^{(n)}}{Z_2^{(n)}}$$

for the weight function $\phi^{(n)}_{\gamma_1,\gamma_2}$ the asymptotics of the lower bound of the weighted variance of this estimator has the form

$$\mathbb{V}^\phi(T(\boldsymbol{Z}^{(n)}_{\beta_1,\beta_2})) \geq \beta_1^2(1 - \beta_1 - \beta_2)^2 \frac{C(\beta_1, \beta_2, \gamma_1, \gamma_2)}{n} + O\left(\frac{1}{n^{3/2}}\right), \tag{22}$$

where $C(\beta_1, \beta_2, \gamma_1, \gamma_2)$ is a constant that depends on the parameters only and is explicitly given in (54).

Theorem 3 (Weighted Kullback inequality).

(a) For given PDFs f,g

$$K^{\phi}(f||g) \geq \Psi^{*}_{\tilde{g}}(\mu_{\phi}(\tilde{f})) = \sup_{t}\left[\langle t, \mu_{\phi}(f)\rangle - \log\bar{M}_{g}(t)\right], \tag{23}$$

where

$$\bar{M}_{g}(t) = \int_{\mathbb{R}^d} \phi(\boldsymbol{z}) e^{\langle t, \boldsymbol{z}\rangle} g(\boldsymbol{z}) \mathrm{d}\boldsymbol{z} \tag{24}$$

is a weighted moment generating function, $t \in \mathbb{R}^d$, and

$$\mu_{\phi}(f) = \frac{\mathbb{E}_f[\boldsymbol{Z}\phi(\boldsymbol{Z})]}{\mathbb{E}_f[\phi(\boldsymbol{Z})]} \in \mathbb{R}^d$$

is the classical expectation of $\tilde{f}$.

(b) Let $Z^{(n)}_{\alpha}$ and $Z^{(n)}_{\rho}$ be Beta RVs with PDF $f^{(n)}_{\alpha}$ given in (11) with $x = \alpha n$ and with PDF $f^{(n)}_{\rho}$ given in (11) with $x = \rho n$, respectively, where $0 < \alpha, \rho < 1$ and the weight function $\phi^{(n)}(p)$. Denote $\epsilon = \alpha - \rho$. Then

$$K^{\phi}(f^{(n)}_{\alpha}||f^{(n)}_{\rho}) \geq \frac{\epsilon^2\left(1+\sqrt{n}-n\right)^2}{2(1-\alpha)\alpha n} + O\left(1\right).$$

As $\epsilon \to 0$,

$$\exists \lim_{\epsilon\to 0}\frac{1}{\epsilon^2}K^{\phi}(f^{(n)}_{\alpha}||f^{(n)}_{\rho}) = \frac{1}{2}I(\tilde{f}_{\alpha}) \geq \frac{n}{2\alpha(1-\alpha)} - \frac{\sqrt{n}}{\alpha(1-\alpha)} + O\left(1\right), \tag{25}$$

where $I(\tilde{f}_{\alpha})$ is the standard Fisher information.

3. Proofs

Let us denote by $\psi^{(0)}(x) = \psi(x)$ and by $\psi^{(1)}(x)$ the digamma function and its first derivative, respectively,

$$\psi^{(n)}(x) = \frac{\mathrm{d}^{\mathrm{n}+1}}{\mathrm{d}x^{n+1}}\log\left(\Gamma(x)\right),$$

where $\Gamma(x)$ is the Gamma-function. In further calculations, the asymptotics of these functions for $x \to \infty$ will be used [12, #8.362.2]

$$\psi(x) = \log(x) - \frac{1}{2x} + O\left(\frac{1}{x^2}\right) \text{ as } x\to\infty, \tag{26}$$

$$\psi^{(1)}(x) = \frac{1}{x} + \frac{1}{2x^2} + O\left(\frac{1}{x^3}\right) \text{ as } x\to\infty. \tag{27}$$

3.1. Proof of Theorem 1

Consider $R_\nu(\mathbf{Z},\theta)$:

$$R_\nu(\mathbf{Z};\theta) = T(\mathbf{Z}) - \tau(\theta) - \sum_{i=1}^{\nu} \lambda_i f_\theta^{(i)} f_\theta^{-1}, \tag{28}$$

where λ_i are undefined parameters. It is easy to note that

$$\mathbb{E}[R_\nu(\mathbf{Z};\theta)] = 0. \tag{29}$$

Consider the weighted variance given in (2) of R_ν. Because of (29), it can be written in the following form

$$\mathbb{V}_\theta^\phi(R_\nu) = \int_{\mathbb{R}^d} \left(T(\mathbf{z}) - \tau(\theta) - \sum_{i=1}^{\nu} \lambda_i f_\theta^{(i)} f_\theta^{-1} \right)^2 \phi(\mathbf{z},\theta,\gamma) f_\theta \mathrm{d}\mathbf{z}. \tag{30}$$

By conditions of the Theorem, the differentiation is justified and leads to the following condition:

$$\int_{\mathbb{R}^d} \left(T(\mathbf{z}) - \tau(\theta) - \sum_{i=1}^{\nu} \lambda_i^\star f_\theta^{(i)} f_\theta^{-1} \right) \phi f_\theta^{(j)} \mathrm{d}\mathbf{z} = 0. \tag{31}$$

It can be rewritten as

$$\sum_{i=1}^{\nu} \lambda_i^\star \int_{\mathbb{R}^d} f_\theta^{(i)} f_\theta^{-1} f_\theta^{(j)} \phi \mathrm{d}\mathbf{z} = \int_{\mathbb{R}^d} T(\mathbf{z}) \phi f_\theta^{(j)} \mathrm{d}\mathbf{z} - \tau(\theta) \int_{\mathbb{R}^d} \phi f_\theta^{(j)} \mathrm{d}\mathbf{z}. \tag{32}$$

Let $\mathbb{I}_\theta^\phi$ be the $\nu \times \nu$ matrix whose elements are

$$I_{ij}^\phi = \int_{\mathbb{R}^d} f_\theta^{(i)} f_\theta^{-1} f_\theta^{(j)} \phi \, \mathrm{d}\mathbf{z},$$

$i, j \leq \nu$. Let

$$\mathbb{J}_\theta^\phi = \left(\mathbb{I}_\theta^\phi \right)^{-1} \tag{33}$$

be the inverse $\nu \times \nu$ matrix and elements of this matrix are J_{ij}^ϕ. Note that in the case $i = j = 1$, I_{11}^ϕ equals to the weighted Fisher information given in (6).

Consider integrals in RHS of (32) separately. Firstly,

$$\int_{\mathbb{R}^d} T(\mathbf{z}) \tilde{\phi} \left(\frac{1}{\kappa(\theta,\gamma)} f_\theta \right)^{(j)} \mathrm{d}\mathbf{z} = g^{(j)}(\theta),$$

$$\int_{\mathbb{R}^d} T(\mathbf{z}) \tilde{\phi} \left[\sum_{k=0}^{j-1} \binom{j}{k} \left(\frac{1}{\kappa(\theta,\gamma)} \right)^{(j-k)} f_\theta^{(k)} \right] \mathrm{d}\mathbf{z} + \int_{\mathbb{R}^d} T(\mathbf{z}) \phi f_\theta^{(j)} \mathrm{d}\mathbf{z} = g^{(j)}(\theta).$$

Thus,

$$\int_{\mathbb{R}^d} T(\mathbf{z}) \phi f_\theta^{(j)} \mathrm{d}\mathbf{z} = g^{(j)}(\theta) - \mathrm{Q}_1^j, \tag{34}$$

where

$$\mathrm{Q}_1^j = \int_{\mathbb{R}^d} T(\mathbf{z})\tilde{\phi}\left[\sum_{k=0}^{j-1}\binom{j}{k}\left(\frac{1}{\kappa(\theta,\gamma)}\right)^{(j-k)} f_\theta^{(k)}\right]\mathrm{d}\mathbf{z}. \tag{35}$$

Similarly to condition (1), the following equality can be derived:

$$\int_{\mathbb{R}^d} \phi f_\theta^{(j)}\mathrm{d}\mathbf{z} = -\mathrm{Q}_2^j, \tag{36}$$

where

$$\mathrm{Q}_2^j = \int_{\mathbb{R}^d} \tilde{\phi}\left[\sum_{k=0}^{j-1}\binom{j}{k}\left(\frac{1}{\kappa(\theta,\gamma)}\right)^{(j-k)} f_\theta^{(k)}\right]\mathrm{d}\mathbf{z}. \tag{37}$$

So, (32) takes the form

$$g^{(j)}(\theta) = \sum_{i=1}^{\nu}\lambda_i^\star I_{ij}^\phi + \mathrm{Q}_1^j - \tau\mathrm{Q}_2^j \tag{38}$$

and

$$\lambda_i^\star = \sum_{j=1}^{\nu}\left(g^{(j)}(\theta) - \mathrm{Q}_1^j + \tau\mathrm{Q}_2^j\right) J_{ij}^\phi. \tag{39}$$

Thus, we obtain the following equality

$$\mathbb{V}(R_\nu^*) = \mathbb{V}_\theta^\phi(T) - \sum_{i,j=1}^{\nu}\left(g^{(i)}(\theta) - \mathrm{Q}_1^i + \tau\mathrm{Q}_2^i\right)\left(g^{(j)}(\theta) - \mathrm{Q}_1^j + \tau\mathrm{Q}_2^j\right) J_{ij}^\phi.$$

The non-negativity of variance implies the lower bound for a weighted variance of T given in (15).

It is easy to see that this inequality includes the weighted version of the Cramér-Rao inequality. It appears when $i = j = \nu = 1$. In this particular case

$$I_{11}^\phi = I^\phi(\theta) = \int_{\mathbb{R}^d}(f_\theta')^2 f_\theta^{-1}\phi\mathrm{d}\mathbf{z},$$

$$\int_{\mathbb{R}^d}\phi f_\theta^{(j)}\mathrm{d}\mathbf{z} = \frac{\kappa'(\theta,\gamma)}{\kappa(\theta,\gamma)}$$

and

$$\int_{\mathbb{R}^d} T(\mathbf{z})\phi f_\theta^{(1)}\mathrm{d}\mathbf{z} = g'(\theta) + \frac{\kappa'(\theta,\gamma)}{\kappa(\theta,\gamma)}g(\theta).$$

Thus, we obtain inequality (16). The weighted version of the Cramér-Rao inequality was initially proposed in [28], where a slightly different definition of the weighted variance was used. According to our definition of the weighted variance (10), the standard (unweighted) mean should be subtracted while in setting [28] the weighted mean is subtracted.

3.2. Example 1 revised

Consider the weight function $\phi^{(n)}(p)$ with $\kappa(\alpha,\gamma)$ is found from the normalizing condition (1),

$$\frac{1}{\kappa(\alpha,\gamma)} = \frac{\Gamma(x+1)\Gamma(n-x+1)\Gamma(n+2+\sqrt{n})}{\Gamma(x+\gamma\sqrt{n}+1)\Gamma(n-x+1+\sqrt{n}-\gamma\sqrt{n})\Gamma(n+2)}.$$

For a given weight function (12) the Fisher information equals

$$\begin{aligned} I^{\phi}(f_{\alpha}^{(n)}) = {} & n^2\left(\psi^{(1)}(x+z+1)+\psi^{(1)}(n-x+1+\sqrt{n}-z)\right) \\ & +n^2\Big[\big(\psi(x+z+1)-\psi(x+1)\big)^2+\big(\psi(n-x+1+\sqrt{n}-z)-\psi(n-x+1)\big)^2\Big] \\ & +2n^2\big[\big(\psi(n-x+1)-\psi(n-x+\sqrt{n}-z+1)\big)\big(\psi(x+z+1)-\psi(x+1)\big)\big], \end{aligned} \tag{40}$$

where $z=\gamma\sqrt{n}$.

For the weight function (12)

$$\int_0^1 p\phi^{(n)} f_{\alpha}^{(n)} \mathrm{d}p = \frac{\Gamma(n+\sqrt{n}+2)\Gamma(x+\gamma\sqrt{n}+2)}{\Gamma(n+\sqrt{n}+3)\Gamma(x+\gamma\sqrt{n}+1)} \equiv g(\alpha). \tag{41}$$

Then

$$\frac{\partial g(\alpha)}{\partial\alpha} = n\frac{\Gamma(n+\sqrt{n}+2)\Gamma(x+\gamma\sqrt{n}+2)}{\Gamma(n+\sqrt{n}+3)\Gamma(x+\gamma\sqrt{n}+1)}\left(\psi(x+\gamma\sqrt{n}+2)-\psi(x+\gamma\sqrt{n}+1)\right). \tag{42}$$

Differentiating $\kappa(\alpha,\gamma)$ we obtain

$$\frac{\kappa'(\alpha,\gamma)}{\kappa(\alpha,\gamma)} = n\left(\psi(n-x+1)-\psi(n-x+1-\gamma\sqrt{n}+1)+\psi(x+\gamma\sqrt{n}+1)-\psi(x+1)\right). \tag{43}$$

Inserting in (40),(41),(42),(43) in (16) we get

$$\mathbb{V}^{\phi}(Z_{\alpha}^{(n)}) \geq \frac{\alpha(1-\alpha)+(\alpha-\gamma)^2}{n} + \frac{-2\alpha+\alpha^2+\gamma+2\alpha\gamma-2\gamma^2}{n^{3/2}} + O\left(\frac{1}{n^2}\right).$$

First of all, it might be of interest to consider the cases when the obtained bound is achieved. In the considered case, the weighted variance can be computed explicitly and one can conclude that the lower bound is achieved asymptotically.

Moreover, one can find the family of distributions for which the Cramér-Rao bound is attained. For this purpose, let us consider the proof of the weighted Cramér-Rao inequality itself. Consider the following integral

$$g(\theta) \equiv \int_{\mathbb{R}^d} T(\mathbf{z})\phi(\mathbf{z},\theta,\gamma)f_{\theta}(\mathbf{z})\mathrm{d}\mathbf{z}. \tag{44}$$

Differentiating both sides in (44) and in (3) w.r.t. θ and multiplying the latter by $e(\theta)$ defined in (5)

$$\int_0^1 T(\mathbf{z})\phi(\mathbf{z},\theta,\gamma)\frac{\partial f_{\theta}}{\partial\theta}\mathrm{d}\mathbf{z} - \frac{\kappa'(\theta,\gamma)}{\kappa^2(\theta,\gamma)}\int_0^1 T(\mathbf{z})\tilde{\phi}(\mathbf{z},\gamma)f_{\theta}(\mathbf{z})\mathrm{d}\mathbf{z} = \frac{\partial g(\theta)}{\partial\theta}, \tag{45}$$

$$e(\theta)\int_0^1 \phi(\mathbf{z},\theta,\gamma)\frac{\partial f_{\theta}}{\partial\theta}\mathrm{d}\mathbf{z} = \frac{\kappa'(\theta,\gamma)}{\kappa(\theta,\gamma)}e(\theta). \tag{46}$$

Subtracting (46) from (45),

$$\int_0^1 (T(\mathbf{z}) - e(\theta))\phi(\mathbf{z},\theta,\gamma)\frac{\partial f_\theta}{\partial\theta}\mathrm{d}\mathbf{z} = \frac{\partial g(\theta)}{\partial\theta} - \frac{\kappa'(\theta,\gamma)}{\kappa(\theta,\gamma)}\left(e(\theta) - g(\theta)\right).$$

Multiplying and dividing by $\sqrt{f_\theta}$, multiplying by the conjugate vector and by applying the Cauchy-Schwarz inequality we obtain inequality(16). Thus, the bound is attained when the equality holds for the Cauchy-Schwarz inequality that is true when

$$c(T(\mathbf{z}) - e(\theta))\sqrt{\phi(\mathbf{z},\theta,\gamma)}\sqrt{f_\theta} = \sqrt{\phi(\mathbf{z},\theta,\gamma)}\frac{\partial f_\theta}{\partial\theta}\frac{1}{\sqrt{f_\theta}}.$$

The weight is cancelled out and we obtain a special form of the exponential family which can be expressed through $T(\mathbf{z})$

$$f(\mathbf{z},\theta) = h(\mathbf{z})e^{c\theta(T(\mathbf{z})-e(\theta))}.$$

Note that conditions when the standard Bhattachrayya inequality is attained are not trivial (see [29]). In the weighted case this problem needs special consideration as well.

It is important to say that by using a relation between the weighted variance for the RV $Z_\alpha^{(n)}$ with PDF f and the standard variance of the RV $\tilde{Z}_\alpha^{(n)}$ with PDF $\tilde{f} \equiv \frac{\tilde{\phi}f}{\kappa}$

$$\mathbb{V}^\phi(Z_\alpha^{(n)}) = \mathbb{V}(\tilde{Z}_\alpha^{(n)}) + \left(g(\alpha) - e(\alpha)\right)^2$$

one can find the lower bound for the weighted variance applying the standard inequality for $\mathbb{V}(\tilde{Z}_\alpha^{(n)})$ and it is of interest which of the bounds is better. By computing the standard bound routinely we deduce that two leading terms of asymptotics coincide in both cases in the context of Example 1. The bounds are equivalent when $\alpha = 0.5, 1 - 2\gamma - 2\gamma^2 \neq 0$ or $\alpha = \gamma$. It appears that the bound obtained with the standard inequality for PDF $\tilde{f}$ is better in other cases. However, it does not generally hold. Indeed, using the relation between the weighted Fisher Information and the standard Fisher Information

$$I(\tilde{f}) \equiv \int \frac{\tilde{\phi}f}{\kappa}\left([\ln\frac{\tilde{\phi}f}{\kappa}]'\right)^2 = I^\phi(f) - \frac{(\kappa')^2}{\kappa^2} \tag{47}$$

one can find the condition when the weighted bound is more tight. By inserting the relation for the weighted Fisher Information, the standard Fisher Information and the weighted variance, the standard variance in the weighted Cramér-Rao inequality and the standard Cramér-Rao inequality one can obtain that the weighted lower bound is more tight if

$$\frac{\left(g^{(1)}(\theta)\kappa(\theta) + (g(\theta) - e(\theta))\kappa^{(1)}(\theta)\right)^2}{\kappa(\theta)^2 I^\phi} + \frac{(g^{(1)}(\theta)\kappa)^2}{\left(\kappa^{(1)}(\theta)\right)^2 - \kappa^2(\theta)I^\phi} > \left(e(\theta) - g(\theta)\right)^2. \tag{48}$$

3.3. Proof of Theorem 2

Note that elements of matrix $\mathbb{F}^\phi$ can be found from condition (1).

Consider a one-dimensional RV

$$\delta = [(T-\tau) - \beta^\star (I^\phi)^{-1}\mathbb{F}^\phi]y,$$

where $y^{\mathrm{T}} = (y_1, \dots, y_l) \in \mathbb{R}^l$ is a non-random vector. It is easy to see that $\mathbb{E}_\theta(\delta) = 0$. Taking the weighted expectation of both sides in the equality

$$\delta^2 = y^{\mathrm{T}} \left[(T-\tau)(T-\tau)^\star - 2(T-\tau)\beta^\star (I^\phi)^{-1}\mathbb{F}^\phi + (\mathbb{F}^\phi)^\star (I^\phi)^{-1}\beta\beta^* (I^\phi)^{-1}\mathbb{F}^\phi\right] y, \quad (49)$$

for any y we obtain

$$\mathbb{E}^\phi_\theta(\delta^2) = y^{\mathrm{T}} \left[\mathbb{V}^\phi_\theta(T) - (\mathbb{F}^\phi)^{\mathrm{T}} (I^\phi)^{-1}\mathbb{F}^\phi\right] y.$$

The non-negativity of variance implies a multi-parametric version of the Bhattacharyya inequality given in (21). One can easily see that in a uni-parametric and a 1D case this inequality is equivalent to the weighted Cramér-Rao inequality.

3.4. Example 2 revised

Let us consider the random vector $\mathbf{Z}^{(n)}_{\beta_1,\beta_2} = (Z^{(n)}_1, Z^{(n)}_2)^{\mathrm{T}}$ that has a Dirichlet distribution given by (13) with parameters $\theta = (\beta_1, \beta_2)^{\mathrm{T}}$. Assume that one would like to estimate the ratio $\tau(\beta_1, \beta_2) = \beta_1/\beta_2$ and to study the lower bound of the variance of the estimator $T(\mathbf{Z}^{(n)}_{\beta_1,\beta_2}) = Z^{(n)}_1/Z^{(n)}_2$ given the point of a particular interest γ_1/γ_2. The weight function $\hat{\phi}^{(n)}(p)$ of the form (2) with $\hat{\kappa} \equiv \kappa(\beta_1, \beta_2, \gamma_1, \gamma_2)$ could be found from the normalizing condition (1) and has the form

$$\frac{1}{\hat{\kappa}} = \frac{\Gamma(n+\sqrt{n}+3)\Gamma(x_1+1)\Gamma(x_2+1)\Gamma(n-x_1-x_2+1)}{\Gamma(n+3)\Gamma(x_1+z_1+1)\Gamma(x_2+z_2+1)\Gamma(n+\sqrt{n}-x_1-x_2-z_1-z_2+1)}, \quad (50)$$

where $x_1 = \beta_1 n$, $x_2 = \beta_2 n$, $z_1 = \gamma_1\sqrt{n}$ and $z_2 = \gamma_2\sqrt{n}$. Therefore, one can obtain that

$$g(\theta) = \frac{x_1+z_1+1}{x_2+z_2}$$

and

$$e(\theta) = \frac{x_1+1}{x_2}.$$

Computing the gradients one can obtain that

$$\frac{\partial}{\partial\theta} g(\theta) = (g_1, g_2)^{\mathrm{T}} \text{ and } \frac{1}{\kappa}\frac{\partial}{\partial\theta}\kappa(\theta) = (\kappa_1, \kappa_2)^{\mathrm{T}},$$

where

$$\begin{aligned}
g_1 &= \frac{n}{x_2+z_2},\\
g_2 &= -\frac{n(x_1+z_1+1)}{(x_2+z_2)^2},\\
\kappa_1 &= n\,\big(\psi(x_1+z_1+1) - \psi(x_1+1) + \psi(n-x_1-x_2+1)\\
&\quad -\psi(n+\sqrt{n}-x_1-x_2-z_1-z_2+1)\big)
\end{aligned}$$

and

$$\begin{aligned}\kappa_2 = n\,&(\psi(x_2+z_2+1)-\psi(x_2+1)+\psi(n-x_1-x_2+1)\\&-\psi(n+\sqrt{n}-x_1-x_2-z_1-z_2+1)\big)\,.\end{aligned}$$

The weighted Fisher Information matrix takes the form

$$I^{\phi} = \begin{pmatrix} I_{11} & I_{12} \\ I_{12} & I_{22} \end{pmatrix}, \tag{51}$$

where the elements of the weighted Fisher Information matrix (51) I_{11},I_{12} and I_{22} have the following form. For $i = 1, 2$,

$$\begin{aligned}n^2 I_{ii} = \psi^{(1)}&(x_i+z_i+1)+\psi^{(1)}(n+\sqrt{n}-x_1-x_2-z_1-z_2+1)\\&+(\psi(x_i+1)-\psi(n-x_1-x_2)-\psi(x_i+z_i+1)\\&+\psi(n+\sqrt{n}-x_1-x_2-z_1-z_2+1)\big)^2\end{aligned}$$

and

$$\begin{aligned}n^2 I_{12} = \psi^{(1)}&(n+\sqrt{n}-x_1-x_2-z_1-z_2+1)\\&+\prod_{i=1,2}(\psi(n-x_1-x_2+1)-\psi(x_i+1)+\psi(x_i+z_i+1)\\&-\psi(n+\sqrt{n}-x_1-x_2-z_1-z_2+1)\big)\,.\end{aligned} \tag{52}$$

Let us denote

$$\begin{aligned}c_i &= \frac{\gamma_i}{\beta_i} - \frac{1-\gamma_1-\gamma_2}{1-\beta_1-\beta_2},\\ t &= \beta_1\gamma_2-\beta_2\gamma_1\end{aligned}$$

and

$$s = \frac{(\gamma_1-\beta_2\gamma_1-\beta_1(1-\gamma_2))(\beta_2-\beta_2\gamma_1-\gamma_2(1-\beta_1)))}{\beta_1\beta_2}$$

Inserting all the terms and applying the asymptotics of the digamma function we can obtain that

$$\mathbb{V}^{\phi}(T(\mathbf{Z}^{(n)}_{\beta_1,\beta_2})) \geq \beta_1^2(1-\beta_1-\beta_2)^2\frac{C(\beta_1,\beta_2,\gamma_1,\gamma_2)}{n} + O\left(\frac{1}{n^{3/2}}\right), \tag{53}$$

where

$$\begin{aligned}C(\beta_1,\beta_2,\gamma_1,\gamma_2) = \sum_{i=1,2}\Bigg(&(\beta_i+tc_i)\left(\frac{(\beta_1+\beta_2-1+s)(\beta_{-i}-tc_i)}{(1-\beta_1-\beta_2)^2}\right)\\&+\left(\frac{1}{\beta_1}+\frac{1}{1-\beta_1-\beta_2}+c_i^2\right)(\beta_i-tc_{-i})\Bigg)\end{aligned}$$

and c_{-i} is a notation for the element different from c_i (say, for $i = 1$, $c_{-i} = c_2$).

3.5. Proof of Theorem 3

(a): The inequality similar to inequality (23) without assumption of the normalization (1) is proved in [28]. The result is the direct consequence of this Theorem.
(b): The weighted generating function of RV $Z_\rho^{(n)}$ with PDF $f_\rho^{(n)}$ equals:

$$\begin{aligned}\bar{M}_{f_\rho^{(n)}}(t) &= \int_0^1 \phi^{(n)} e^{tp} f_\rho^{(n)} \mathrm{d}p = {}_1F_1(\rho n + \gamma\sqrt{n} + 1, n + \sqrt{n} + 2; t)\\ &= 1 + \sum_{k=1}^{\infty} \frac{t^k}{k!} \prod_{j=0}^{k-1} \frac{\rho n + \gamma\sqrt{n} + 1 + j}{n + \sqrt{n} + 2 + j},\end{aligned}$$

where ${}_1F_1(x, y; z)$ is the confluent hypergeometric function.

For large n, the expression for a weighted generating function can be written in the following way [15, formula 12]:

$$\begin{aligned}\bar{M}_{f_\rho^{(n)}}(t) &= 1 + \sum_{k=1}^{\infty} \frac{t^k}{k!} \prod_{j=0}^{k-1} \frac{\rho n + \gamma\sqrt{n} + 1 + j}{n + \sqrt{n} + 2 + j}\\ &= \sum_{k=0}^{\infty} \frac{t^k}{k!} \left(\rho^k - k(\rho^k - \rho^{k-1}\gamma) \frac{1}{\sqrt{n}} \right.\\ &\quad \left. + \frac{\rho^{k-2} k(\rho - 2\rho^2 - \gamma^2 + \rho k - 2\rho\gamma k + \gamma^2 k)}{2n} + O\left(\frac{1}{n^{3/2}}\right) \right)\\ &= e^{\rho t} \left(1 - (\rho - \gamma) t \frac{1}{\sqrt{n}} + \frac{2(1 - \rho - \gamma) t + (\rho - 2\rho\gamma + \gamma^2) t^2}{2n} + O\left(\frac{1}{n^{3/2}}\right) \right).\end{aligned}$$

Thus, we have that

$$\begin{aligned}\log \bar{M}_{f_\rho^{(n)}}(t) &= \rho t + \log \left(1 - (\rho - \gamma) t \frac{1}{\sqrt{n}} \right.\\ &\quad \left. + \frac{2(1 - \rho - \gamma) t + (\rho - 2\rho\gamma + \gamma^2) t^2}{2n} + O\left(\frac{1}{n^{3/2}}\right) \right)\\ &= \rho t - (\rho - \gamma) t \frac{1}{\sqrt{n}} + \frac{(1 - \rho - \gamma) t + \frac{\rho t^2}{2}(1 - \rho)}{n} + O\left(\frac{1}{n^{3/2}}\right).\end{aligned}$$

The first term in (23) for PDF $f_\alpha^{(n)}$ and weight function $\phi^{(n)}$ takes the following form

$$\mu_\phi(f_\alpha^{(n)}) = \frac{\alpha n + \gamma\sqrt{n} + 2}{n + \sqrt{n} + 2} = \alpha + (\gamma - \alpha) \frac{1}{\sqrt{n}} + \frac{1 - \alpha - \gamma}{n} + O\left(\frac{1}{n^{3/2}}\right).$$

Then

$$\Psi_{f_\rho}^*(\mu_\phi(\tilde{f}_\alpha)) = \sup_t \left[(\alpha - \rho) t - (\alpha - \rho) \frac{t}{\sqrt{n}} + \frac{\rho - \alpha}{n} t - \frac{(1 - \rho)\rho}{2n} t^2 + O\left(\frac{1}{n^{3/2}}\right) \right]. \quad (54)$$

Finding the supremum of the expression above, we obtain

$$\tau = \frac{(\alpha-\rho)\,(n-1-\sqrt{n})}{(1-\alpha)\alpha} + O\left(\frac{1}{n^{1/2}}\right).$$

So,

$$\Psi^*_{f_\rho}(\mu_\phi(\tilde{f}_\alpha)) = \frac{(\alpha-\rho)^2\,(1+\sqrt{n}-n)^2}{2(1-\alpha)\alpha n} + O\,(1)\,.$$

Denote $\epsilon = \alpha - \rho$. When $\epsilon \to 0$, we obtain

$$\frac{1}{\epsilon^2}\Psi^*_{f_\rho}(\mu_\phi(\tilde{f}_\alpha)) = \frac{n}{2\alpha(1-\alpha)} - \frac{\sqrt{n}}{\alpha(1-\alpha)} + O\,(1)\,.$$

Thus,

$$\exists \lim_{\epsilon\to 0} \frac{1}{\epsilon^2} K^\phi(f_\alpha^{(n)}||f_\rho^{(n)}) = \frac{1}{2} I(\tilde{f}_\alpha) \geq \frac{n}{2\alpha(1-\alpha)} - \frac{\sqrt{n}}{\alpha(1-\alpha)} + O\,(1)$$

which completes the proof of Theorem 3.

Acknowledgement

We would like to thank anonymous referees for useful comments and remarks that helped to improve the paper significantly. The article was partly prepared within the framework of a subsidy granted to the HSE by the Government of the Russian Federation for the implementation of the Global Competitiveness Program. This project has been partly funded by the European Unions Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 633567.

References

[1] M. Belis, S. Guiasu, *A quantitative and qualitative measure of information in cybernetic systems*, IEEE Trans. Inf. Th. **14**(1968), 593–594.

[2] S. Berry, B. Carlin, J. Lee, P. Muller, *Bayesian adaptive methods for clinical trials*, CRC Press, Boca Raton, 2011.

[3] A. Bhattacharyya, *On some analogues of the amount of information and their use in statistical estimation*, Sankhya **8**(1946), 1.

[4] L. Bolshev, *A refinement of Rao-Cramér inequality*, Th. Prob. Appl. **6**(1961), 319–326.

[5] J. Chen, R. Kodell, R. Howe, D. Gaylor, *Analysis of trinomial responses from reproductive and developmental toxicity experiments*, Biometrics **47**(1991), 1049–58.

[6] S. Chevret, *Statistical methods for dose-finding experiments*, Wiley, New Jersey, 2006.

[7] A. Clim, *Weighted entropy with application*, Analele Universitatii Bucurestica Matematica **LVII**(2008), 223–231.

[8] A. G. Dabak, D. H. Johnson, *Relations between Kullback-Leibler distance and Fisher information*, 2002, available at `http://www.ece.rice.edu/ dhj/distance.pdf`

[9] L. Delchambre, *Weighted principal component analysis: a weighted covariance eigendecomposition approach*, Math. Notes Royal Astronom. Soc. **446**(2014), 3545–3555.
[10] G. Dial, I. Taneja, *On weighted entropy of type (α, β) and its generalizations*, Appl.Math. **26**(1981), 418–425.
[11] M. Gasparini, J. Eisele, *A curve-free method for phase I clinical trials*, Biometrics **56**(2000), 609–615.
[12] I. S. Gradshteyn, I. M. Ryzhik, *Table of integrals, series, and product*, Elsevier, Amsterdam, 2007.
[13] J. Grossman, M. Grossman, R. Katz, *The first systems of weighted differential and integral calculus*, Rockport, MA: Archimedes Foundation, 1980.
[14] S. Guiasu, *Weighted entropy*, Report Math. Physics **2**(1971), 165–179.
[15] M. Hapaev, *Asymptotic expansions of hypergeometric and confluent hypergeometric functions*, Izv. Vyssh. Uchebn. Zaved. Mat. (1961), 98–101.
[16] G. H. Hardy, J. E. Littlewood, G. Polya, *Inequalities*, Cambridge University Press, Cambridge, 1988.
[17] T. Jaki, S. Clive, C. Weir, *Principles of dose finding studies in cancer: a comparison of trial designs*, Cancer Chemotherapy and Pharmacology **71**(2013), 1107–1114.
[18] J. Kapur, *Measures of Information and Their Applications*, Wiley Eastern Limited, New Delhi, 1994.
[19] M. Kelbert, Yu. Suhov, *Information theory and coding by example*, Cambridge University Press, Cambridge, 2013.
[20] M. Kelbert, P. Mozgunov, *Shannon's differential entropy asymptotic analysis in a Bayesian problem*, Math. Commun. **20**(2015), 219–228.
[21] M. Kelbert, P. Mozgunov, *Asymptotic behaviour of the weighted Renyi, Tsallis and Fisher entropies in a Bayesian problem*, Eurasian Math. J. **6**(2015), 6–17.
[22] M. Kelbert, P. Mozgunov, *Asymptotic behaviour of weighted differential entropies in a Bayesian problem*, arXiv 1504.01612, 2015.
[23] F. Pozzi, T. Di Matteo, T. Aste, *Exponential smoothing weighted correlations*, The Eur. Phys. J. **85**(2012), 175.
[24] I. Salama, D. Quade, *A non-parametric comparison of two multiple regressions by means of a weighted measure of correlation*, Comm. Statist. Theory Methods **11**(1982), 1185-1195.
[25] R. Singh, J. Bhardwaj, *On parametric weighted information improvement*, Inf. Sci. **59**(1992), 149–163.
[26] A. Sreevally, S. Varma, *Generating measure of cross entropy by using measure of weighted entropy*, Soochow J. Math. **30**(2004), 237–243.
[27] A. Srivastava, *Some new bounds of weighted entropy measures*, Cybern. Inf. Technol. **11**(2011), 60–65.
[28] M. Kelbert, Y. Suhov, I. Stuhl, S. Y. Sekeh, *Basic inequalities for weighted entropies*, Aequationes Mathematicae (2016), 1–32.
[29] H. Tanaka, M. Akahira, *On a family of distributions attaining the Bhattacharyya bound*, Ann. Inst. Statist. Math. Vol. **55**(2003), 309–317.