Ann Inst Stat Math (2011) 63:689–716
DOI 10.1007/s10463-009-0254-0

Large deviation theory for non-regular location shift family

Masahito Hayashi

Received: 9 July 2008 / Revised: 7 January 2009 / Published online: 28 July 2009

Abstract We apply non-regular extensions of the large deviation theory to non-regular location shift families. Our calculation contains the location shift families generated by Beta distribution, Weibull distribution, and Gamma distribution. We point out the optimal estimator depends on the choice of our criterion in the non-regular case. The limits of relative Rényi entropies play an important role in our derivation.

Keywords Non-regular location shift family · Beta distribution · Weibull distribution · Gamma distribution · Large deviation · Relative Rényi entropy

1 Introduction

In statistics, we often treat Beta distribution, Weibull distribution, and Gamma distribution. However, these distributions have a limited support. That is, if we consider the location shift family generated by Beta distribution, Weibull distribution, or Gamma distribution, the support depends on the location. In the conventional estimation theory, we assume that the support does not depend on the true parameter and evaluate the accuracy of the asymptotically best estimator based on Fisher information. When the location shift family is generated by a probability density function with a limited support, Fisher information cannot be defined in general. So, we cannot apply the conventional estimation theory. However, the fact that the support depends on the true parameter gives an advantage for increasing the accuracy of our estimator. Therefore, we can expect to overcome the accuracy of the estimator in the conventional case.

M. Hayashi (✉)
Graduate School of Information Sciences, Tohoku University,
Aoba-ku, Sendai 980-8579, Japan
e-mail: hayashi@math.is.tohoku.ac.jp

The main points of this problem are finding an asymptotic optimum estimator and evaluating the accuracy of the estimator in the non-regular location shift family. This problem has been discussed by Akahira (1996) and Akahira et al. (1995) from the viewpoint of information loss. Recently, using Bahadur (1960, 1967, 1971) large deviation theory, Hayashi (2008) formulated a general estimation theory of non-regular family. Hayashi's formulation has two types of criteria $\alpha_1(\theta)$ and $\alpha_2(\theta)$ in large deviation theory, in which these two criteria coincide in the regular case. As is mentioned in Hayashi (2008), the first criterion $\alpha_1(\theta)$ corresponds to the interval estimation, and the second criterion $\alpha_2(\theta)$ does to the point estimation.

Although the previous paper (Hayashi 2008) treated how to avoid the super efficiency of Bahadur (1960, 1967, 1971) large deviation theory in the non-regular case pointed by Ibragimov and Has'minskii Ibragimov et al. (1981), it did not apply its formulation to concrete examples. Further, while it derived the upper bounds of $\alpha_1(\theta)$ and $\alpha_2(\theta)$, the attainability of these upper bounds were not investigated sufficiently. In order to clarify the effectiveness of this formulation and the obtained upper bound, it is required to apply it to several specific statistical models and discuss the following issues. (1) Comparison and calculation of two types of criteria $\alpha_1(\theta)$ and $\alpha_2(\theta)$ on several statistical models. (2) Checking whether their upper bounds derived in Hayashi (2008) can be attained on several statistical models.

The main purpose of the present paper is to apply the formulation in Hayashi (2008) to several realistic examples and check the attainability of upper bound given in Hayashi (2008). For this purpose, the present paper focuses on a specific non-regular location shift family generated by the following support-limited probability density function. When the support is a closed interval $[a, b]$, the probability density function behaves as $A_1(x-a)^{\kappa_1-1}$ and $A_2(x-b)^{\kappa_2-1}$ at the neighborhood of the boundary. Our analysis depends on the two shape parameters κ_1 and κ_2, and contains the half-line support case by substituting 0 into A_2. That is, it covers location shift families generated by Beta distribution, Weibull distribution, and Gamma distribution. As is summarized in Tables 1, 2 and 3, we calculate both bounds and discuss the attainability of both upper bounds. Through this type discussion, we clarify that the optimum estimator depends on the shape parameters and the choice of the criterion. The maximum likelihood estimator (MLE) and the convex combination of extremal statistics give good performances in the second criterion (point estimation), while the likelihood ratio estimator (LRE) and a modification of extremal statistics give good performances in the first criterion (interval estimation). Further, we discuss the relation among their performances, optimal values, and the shape parameters.

The present paper is organized as follows. In Sect. 2, a general large deviation theory for non-regular families is reviewed as Hayashi (2008) extension of Bahadur (1960, 1967, 1971) large deviation theory. In Sect. 3, we summarize the limits of relative Rényi entropies, which are calculated in Hayashi (2002b) In Sect. 4, we review the existing results concerning the calculation of the performances of respective estimators. In Sect. 5, which is the main part, we calculate the performances of respective estimators and the optimum performances $\alpha_1(\theta)$ and $\alpha_2(\theta)$ in the location shift families depending on the shape parameters. In Sect. 6, we apply the main result to the location shift families generated by Gamma distribution, Weibull distribution, and Beta distribution.

2 General theory

Bahadur (1960, 1967, 1971) proposed a large deviation theory, which is fundamental for a general large deviation theory for non-regular families by Hayashi (2008) and is summarized as follows. Given n-i.i.d. data $\omega_1, \ldots, \omega_n$, the error probability $p_\theta^n\{|T_n - \theta| > \epsilon\}$ goes to 0 exponentially. That is, the rate of the quantity $-\log p_\theta^n\{|T_n - \theta| > \epsilon\}$ is in order n. Hence, for the estimator $\mathbf{T} = \{T_n\}$, we focus on the following exponential decreasing rate of the error probability with a fixed error bar Bahadur (1960, 1967, 1971):

$$\beta(\mathbf{T}, \theta, \epsilon) := \liminf \frac{-1}{n} \log p_\theta^n\{|T_n - \theta| > \epsilon\}. \tag{1}$$

This exponential rate $\beta(\mathbf{T}, \theta, \epsilon)$ can be written by the estimator $\mathbf{T} = \{T_n\}$ is written by the exponential rates of half-side error probabilities as

$$\beta(\mathbf{T}, \theta, \epsilon) = \min\{\beta^+(\mathbf{T}, \theta, \epsilon), \beta^-(\mathbf{T}, \theta, \epsilon)\},$$

where the exponential rates of half-side error probabilities are given by

$$\beta^+(\mathbf{T}, \theta, \epsilon) := \liminf \frac{-1}{n} \log p_\theta^n\{T_n > \theta + \epsilon\}$$
$$\beta^-(\mathbf{T}, \theta, \epsilon) := \liminf \frac{-1}{n} \log p_\theta^n\{T_n < \theta - \epsilon\}.$$

When an estimator $\mathbf{T} = \{T_n\}$ satisfies the weak consistency

$$p_\theta^n\{|T_n - \theta| > \epsilon\} \to 0 \quad \forall \epsilon > 0, \quad \forall \theta \in \Theta,$$

using the monotonicity of KL-divergence, we can prove the inequality

$$\beta(\mathbf{T}, \theta, \epsilon) \leq \min\{D(p_{\theta+\epsilon}\|p_\theta), D(p_{\theta-\epsilon}\|p_\theta)\}. \tag{2}$$

Note that if, and only if, the family is exponential, there exists an estimator attaining the equality (2) at $\forall \theta \in \Theta$, $\forall \epsilon > 0$. Therefore, for a general family, it is difficult to optimize the exponential rate $\beta(\mathbf{T}, \theta, \epsilon)$.

Instead of the exponential rate $\beta(\mathbf{T}, \theta, \epsilon)$, We usually consider the slope of the exponential rate:

$$\alpha(\mathbf{T}, \theta) := \lim_{\epsilon \to +0} \frac{1}{\epsilon^2} \beta(\mathbf{T}, \theta, \epsilon). \tag{3}$$

In this case, when the Fisher information J_θ satisfies the condition

$$J_\theta := \int l_\theta(\omega)^2 p_\theta(\mathrm{d}\omega) = \lim_{\epsilon \to 0} \frac{2}{\epsilon^2} D(p_{\theta+\epsilon}\|p_\theta), \quad l_\theta(\omega) := \frac{\mathrm{d}}{\mathrm{d}\theta} \log p_\theta(\omega), \tag{4}$$

the inequality

$$\alpha(\mathbf{T}, \theta) \leq \frac{1}{2} J_\theta \tag{5}$$

holds.

However, when the support depends on the true parameter, the KL-divergence $D(p_{\theta+\epsilon}\|p_\theta)$ does not have a finite value. In order to resolve this problem, we use the relative Rényi entropies $I^s(p_\theta\|p_{\theta+\epsilon}) := -\log \int p^s\theta(\omega)p^{1-s}\theta + \epsilon(\omega)d\omega$ $(0 < s < 1)$. In general, the order of $I^s(p_\theta\|p_{\theta+\epsilon})$ is not necessarily ϵ^2 at the limit $\epsilon \to 0$. However, its order is independent of the parameter s, as is guaranteed by the inequalities

$$2\min\{s, 1-s\} I^{\frac{1}{2}}(p_\theta\|p_{\theta+\epsilon}) \leq I^s(p_\theta\|p_{\theta+\epsilon}) \leq 2\max\{s, 1-s\} I^{\frac{1}{2}}(p_\theta\|p_{\theta+\epsilon}),$$

which are proven in Hayashi (2008). In several cases, the order of the exponential rate $\beta(\mathbf{T}, \theta, \epsilon)$ coincides with the order of the relative Rényi entropies $I^s(p_\theta\|p_{\theta+\epsilon})$. In the following, we use a strictly monotonically decreasing function $g(x)$ such that $I^s(p_\theta\|p_{\theta+\epsilon}) \cong O(g(\epsilon))$ and $g(0) = 0$.

In the present paper, as the criteria of the asymptotic accuracy of the estimator, we focus on the two quantities

$$\alpha_1(\theta) := \limsup_{\epsilon\to+0} \frac{1}{g(\epsilon)} \sup_{\mathbf{T}} \inf_{\theta-\epsilon\leq\theta'\leq\theta+\epsilon} \beta(\mathbf{T}, \theta', \epsilon) \tag{6}$$

$$\alpha_2(\theta) := \sup_{\mathbf{T}} \alpha_2(\mathbf{T}, \theta), \tag{7}$$

which are defined as two extensions of Bahadur's bound (slope) $\alpha(\theta)$, where

$$\alpha_2(\mathbf{T}, \theta) := \liminf_{\epsilon\to+0} \frac{1}{g(\epsilon)} \inf_{\theta-\epsilon\leq\theta'\leq\theta+\epsilon} \beta(\mathbf{T}, \theta', \epsilon). \tag{8}$$

These two bounds correspond to the point estimation and interval estimation, as is mentioned in Hayashi (2008).

The following propositions hold concerning these two quantities:

Proposition 1 (Hayashi 2008) *When the convergence* $\lim_{\epsilon\to0} \frac{I^s(p_{\theta-\epsilon/2}\|p_{\theta+\epsilon/2})}{g(\epsilon)}$ *is uniform for* $0 < s < 1$*, the inequality*

$$\alpha_1(\theta) \leq \overline{\alpha}_1(\theta) := 2^\kappa \sup_{0<s<1} I^s_{g,\theta} \tag{9}$$

holds, where κ *and* $I^s_{g,\theta}$ *are defined by*

$$I^s_{g,\theta} := \lim_{\epsilon\to+0} \frac{I^s(p_{\theta-\epsilon/2}\|p_{\theta+\epsilon/2})}{g(\epsilon)} \quad 1 \geq s \geq 0 \tag{10}$$

$$x^\kappa = \lim_{\epsilon\to+0} \frac{g(x\epsilon)}{g(\epsilon)}. \tag{11}$$

Note that the function $s \to I_g^s$ is concave and continuous because the function $s \to I^s(p_{\theta-\frac{1}{2}\epsilon}\|p_{\theta+\frac{1}{2}\epsilon})$ is concave and continuous. Therefore, when $I_{g,\theta}^s = I_{g,\theta}^{1-s}$, we have

$$\overline{\alpha}_1(\theta) = 2^\kappa I_{g,\theta}^{\frac{1}{2}}. \tag{12}$$

Proposition 2 (Hayashi 2008) *If the convergence* $\lim_{\epsilon\to 0} \frac{I^s(p_{\theta-\epsilon/2}\|p_{\theta+\epsilon/2})}{g(\epsilon)}$ *is uniform for* $s \in (0, 1)$ *and* $\theta \in K$ *for any compact set* $K \subset \mathbb{R}$*, the inequality*

$$\alpha_2(\theta) \le \overline{\alpha}_2(\theta) := \begin{cases} \sup_{0<s<1} \frac{I_{g,\theta}^s}{s(1-s)} \left(s^{\frac{1}{\kappa-1}} + (1-s)^{\frac{1}{\kappa-1}}\right)^{\kappa-1} & \text{if } \kappa < 1 \\ 2I_{g,\theta}^{\frac{1}{2}} & \text{if } \kappa = 1 \\ \inf_{0<s<1} \frac{I_{g,\theta}^s}{s(1-s)} \left(s^{\frac{1}{\kappa-1}} + (1-s)^{\frac{1}{\kappa-1}}\right)^{\kappa-1} & \text{if } \kappa > 1 \end{cases} \tag{13}$$

holds.

When we choose $g(x) = x^2$, we have $\kappa = 2$, $I_{g,\theta}^s = \frac{1}{2} J_\theta s(1-s)$, and the relations

$$\overline{\alpha}_1(\theta) = 4 \max_{0\le s\le 1} I_{g,\theta}^s = \frac{1}{2} J_\theta$$
$$\overline{\alpha}_2(\theta) = \min_{0\le s\le 1} \frac{I_{g,\theta}^s}{s(1-s)} = \frac{1}{2} J_\theta$$

hold. In particular, if the distribution family satisfies the concavity of the logarithmic derivative $l_\theta(\omega)$ for θ and some other conditions, the bound $\frac{1}{2} J_\theta$ is attained by the MLE. Thus, the relations $\alpha_1(\theta) = \alpha_2(\theta) = \overline{\alpha}_1(\theta) = \overline{\alpha}_2(\theta) = \frac{1}{2} J_\theta$ hold.

As a relation between two bounds $\overline{\alpha}_1(\theta)$ and $\overline{\alpha}_2(\theta)$, we can prove the following theorem.

Proposition 3 (Hayashi 2008) *The inequality*

$$\overline{\alpha}_1(\theta) \ge \overline{\alpha}_2(\theta) \tag{14}$$

holds, and (14) *holds as an equality if and only if the equations*

$$\overline{\alpha}_1(\theta) = 2^\kappa I_{g,\theta}^{\frac{1}{2}} \tag{15}$$
$$2^\kappa I_{g,\theta}^{\frac{1}{2}} = \overline{\alpha}_2(\theta) \tag{16}$$

hold. When $\kappa \le 1$, (14) *holds as an equality if and only if Eq.* (15) *holds.*

When $I_{g,\theta}^s$ is differentiable, condition (15) is equivalent to $\frac{\mathrm{d}}{\mathrm{d}s} I_{g,\theta}^s|_{s=\frac{1}{2}} = 0$.

3 Limits of relative Rényi entropies

In order to calculate the upper bounds $\overline{\alpha}_1(\theta)$ and $\overline{\alpha}_2(\theta)$, we need to calculate the limits of relative Rényi entropies. In this section, we summarize the calculation in the case of location shift family, which is given in Hayashi (2002b)

3.1 Interval support case

In this subsection, we discuss the location shift family generated by a C^3 continuous probability density function f whose support is an open interval $(a, b) \subset \mathbb{R}$. We assume conditions (17) and (18) for f:

$$f_1(x) := f(a+x) \cong A_1 x^{\kappa_1 - 1} \quad \text{as } x \to +0 \tag{17}$$
$$f_2(x) := f(b-x) \cong A_2 x^{\kappa_2 - 1} \quad \text{as } x \to +0, \tag{18}$$

where $\kappa_1, \kappa_2 > 0$. In addition, if $\kappa_i \neq 1$, we assume the following conditions:

$$f_i'(x) \cong A_i(\kappa_i - 1)x^{\kappa_i - 2} \quad \text{as } x \to +0 \tag{19}$$
$$f_i''(x) \cong A_i(\kappa_i - 1)(\kappa_i - 2)x^{\kappa_i - 3} \quad \text{as } x \to +0 \text{ if } \kappa_i \neq 2 \tag{20}$$
$$x f_i''(x) \to 0 \quad \text{as } x \to +0 \text{ if } \kappa_i = 2. \tag{21}$$

If $\kappa_i = 1$, we assume the existence of the limits $\lim_{x\to+0} f_i'(x)$ and $\lim_{x\to+0} f_i''(x)$. If $\kappa_i > 2$, we assume that

$$J_f := \int_a^b f^{-1}(x)(f')^2(x)\mathrm{d}x < \infty. \tag{22}$$

For example, when f is the beta distribution $f(x) = \frac{1}{B(\alpha,\beta)}x^{\alpha-1}(1-x)^{\beta-1}$ whose support is $(0, 1)$, the above conditions are satisfied and we have

$$\kappa_1 = \alpha, \quad \kappa_2 = \beta, \quad A_1 = A_2 = \frac{1}{B(\alpha, \beta)}. \tag{23}$$

In this paper, we denote the beta function by $B(x, y)$. Then, we have the following theorem:

Proposition 4 (Hayashi 2002b) *Assume that $\kappa := \kappa_1 = \kappa_2$, we obtain the following relations:*

$$\lim_{\epsilon\to+0}\frac{I^s(f_\theta\|f_{\theta+\epsilon})}{\epsilon^\kappa}=\begin{cases}\frac{1-\kappa}{\kappa}\big(A_1 sB(s+\kappa(1-s),1-\kappa) \\ \quad +A_2(1-s)B(1-s+\kappa s,1-\kappa)\big) & 0<\kappa<1\\ A_1 s+A_2(1-s) & \kappa=1\\ \frac{A_1 s(1-s(\kappa-1))B(s+\kappa(1-s),2-\kappa)}{\kappa} & 1<\kappa<2\\ \quad +\frac{A_2(1-s)(1-(1-s)(\kappa-1))B(1-s+\kappa s,2-\kappa)}{\kappa} & \end{cases} \tag{24}$$

$$\lim_{\epsilon\to+0}\frac{I^s(f_\theta\|f_{\theta+\epsilon})}{-\epsilon^2\log\epsilon}=\frac{(A_1+A_2)s(1-s)}{2}\quad \kappa=2$$

$$\lim_{\epsilon\to+0}\frac{I^s(f_\theta\|f_{\theta+\epsilon})}{\epsilon^2}=\frac{s(1-s)}{2}J_f\quad 2<\kappa,$$

where $f_\theta(x):=f(x-\theta)$. *These convergences are uniform for* $0<s<1$. *If* $\kappa_1<\kappa_2$, *substituting* $\kappa:=\kappa_1$, $A_2:=0$, *we obtain the above equations.*

The cases $\kappa>2$ and $\kappa=1$ were proved by Akahira (1996). However, he did not treat the uniformity of $0<s<1$, which is essential for the discussion in Propositions 1 and 2. The case $\kappa>2$ is an example where relation $\lim_{\epsilon\to+0}\frac{I^s(p_\theta\|p_{\theta+\epsilon})}{\epsilon^2}=\frac{J_\theta s(1-s)}{2}$ holds, but relation (4) does not. Further, when $f_i(x)$ goes to zero faster than x, i.e., $f_i(x)/x\to 0$ as $x\to 0$, similarly, the relation $\lim_{\epsilon\to+0}\frac{I^s(p_\theta\|p_{\theta+\epsilon})}{\epsilon^2}=\frac{J_\theta s(1-s)}{2}$ holds, but relation (4) does not. Note that when $0<\kappa<2$, in general, the equation $\lim_{\epsilon\to+0}\frac{I^s(f_\theta\|f_{\theta+\epsilon})}{\epsilon^\kappa}=\lim_{\epsilon\to-0}\frac{I^s(f_\theta\|f_{\theta+\epsilon})}{|\epsilon|^\kappa}$ does not hold.

Proof Since $I^s(f_\theta\|f_{\theta+\epsilon})=I^s(f_{-\epsilon}\|f_0)$, Lemma 1 yields Eq. (24). □

Next, we introduce two quantities for any $c\in(a,b)$ as

$$I_s^-(c,f,\epsilon):=\int_a^c f^{1-s}(x)f^s(x+\epsilon)\mathrm{d}x-\int_a^c f(x)\mathrm{d}x-f(c)s\epsilon-\frac{s}{2}f'(c)\epsilon^2,$$

$$I_s^+(c,f,\epsilon):=\int_c^{b-\epsilon} f^{1-s}(x)f^s(x+\epsilon)\mathrm{d}x-\int_c^b f(x)\mathrm{d}x+f(c)s\epsilon+\frac{s}{2}f'(c)\epsilon^2.$$

Lemma 1 (Hayashi 2002b) *We obtain the following relations:*

$$\lim_{\epsilon\to+0}\frac{I_s^-(c,f,\epsilon)}{\epsilon^{\kappa_1}}=\begin{cases}-\frac{1-\kappa_1}{\kappa_1}A_1 sB(s+\kappa_1(1-s),1-\kappa_1) & 0<\kappa_1<1\\ -A_1 s & \kappa_1=1\\ -\frac{A_1 s(1-s(\kappa_1-1))B(s+\kappa_1(1-s),2-\kappa_1)}{\kappa_1} & 1<\kappa_1<2\end{cases} \tag{25}$$

$$\lim_{\epsilon\to+0}\frac{I_s^-(c,f,\epsilon)}{-\epsilon^2\log\epsilon}=-\frac{A_1 s(1-s)}{2}\quad \kappa_1=2$$

$$\lim_{\epsilon\to+0}\frac{I_s^-(c,f,\epsilon)}{\epsilon^2}=-\frac{s(1-s)}{2}J^-_{f,c}\quad 2<\kappa_1$$

and

$$\lim_{\epsilon \to +0} \frac{I_s^+(c, f, \epsilon)}{\epsilon^{\kappa_2}}$$
$$= \begin{cases} \frac{1-\kappa_2}{\kappa_2} - A_2(1-s)B(1-s+\kappa_2 s, 1-\kappa_2)) & 0 < \kappa_2 < 1 \\ -A_2(1-s) & \kappa_2 = 1 \\ -\frac{A_2(1-s)(1-(1-s)(\kappa_2-1))B(1-s+\kappa_2 s, 2-\kappa_2)}{\kappa_2} & 1 < \kappa_2 < 2 \end{cases} \tag{26}$$

$$\lim_{\epsilon \to +0} \frac{I_s^+(c, f, \epsilon)}{-\epsilon^2 \log \epsilon} = -\frac{A_2 s(1-s)}{2} \quad \kappa_2 = 2$$
$$\lim_{\epsilon \to +0} \frac{I_s^+(c, f, \epsilon)}{\epsilon^2} = -\frac{s(1-s)}{2} J_{f,c}^+ \quad 2 < \kappa_2,$$

where J_f^- *and* J_f^+ *are defined as*

$$J_{f,c}^- := \int_a^c f^{-1}(x)(f'(x))^2 \mathrm{d}x, \quad J_{f,c}^+ := \int_c^b f^{-1}(x)(f'(x))^2 \mathrm{d}x.$$

These convergences are uniform for $0 < s < 1$.

3.2 Half-line support case

In this subsection, we discuss the case where the support is the half-line $(0, \infty)$ and the probability density function f is C^3 continuous. Similar to (17) and (18), we assume that

$$f(x) \cong A x^{\kappa-1} \quad \text{as } x \to 0. \tag{27}$$

When $\kappa \neq 1$, we assume the following conditions:

$$f'(x) \cong A_i(\kappa-1)x^{\kappa-2} \quad \text{as } x \to +0 \tag{28}$$

$$f''(x) \cong A_i(\kappa-1)(\kappa-2)x^{\kappa-3} \quad \text{as } x \to +0 \text{ if } \kappa \neq 2 \tag{29}$$

$$x f''(x) \to 0 \quad \text{as } x \to +0 \text{ if } \kappa = 2. \tag{30}$$

When $\kappa = 1$, we assume the existence of the limits $\lim_{x\to+0} f'(x)$ and $\lim_{x\to+0} f''(x)$. In addition, we assume that there exist real numbers $c > 0$ and $\epsilon > 0$ such that

$$\int_c^\infty f^{-1}(x)(f'(x))^2 \mathrm{d}x < \infty \tag{31}$$

$$\int_c^\infty \sup_{0\le t_1\le\epsilon} f(x+t_1) \sup_{0\le t_2\le\epsilon} |f^{-3}(x+t_2)(f')^3(x+t_2)|\mathrm{d}x < \infty \tag{32}$$

$$\int_c^\infty \sup_{0\le t_1\le\epsilon} f(x+t_1) \sup_{0\le t_2\le\epsilon} |f^{-2}(x+t_2)f'(x+t_2)f''(x+t_2)|\mathrm{d}x < \infty \tag{33}$$

$$\int_c^\infty \sup_{0\le t_1\le\epsilon} f(x+t_1) \sup_{0\le t_2\le\epsilon} |f^{-1}(x+t_2)f'''(x+t_2)|\mathrm{d}x < \infty. \tag{34}$$

For example, when f is Weibull distribution $\alpha\beta x^{\alpha-1}e^{\beta x^\alpha}$, the above conditions are satisfied and we have

$$\kappa = \alpha, \quad A = \alpha\beta. \tag{35}$$

When f is gamma distribution $f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{\beta x}$, the above conditions are satisfied and

$$\kappa = \alpha, \quad A = \frac{\beta^\alpha}{\Gamma(\alpha)}. \tag{36}$$

Now, we obtain the following theorem:

Proposition 5 (Hayashi 2002b) *We obtain*

$$\lim_{\epsilon\to+0} \frac{I^s(f_\theta\|f_{\theta+\epsilon})}{\epsilon^\kappa} = \begin{cases} \dfrac{1-\kappa}{\kappa}(AsB(s+\kappa(1-s),1-\kappa) & 0<\kappa<1 \\ As & \kappa=1 \\ \dfrac{As(1-s(\kappa-1))B(s+\kappa(1-s),2-\kappa)}{\kappa} & 1<\kappa<2 \end{cases} \tag{37}$$

$$\lim_{\epsilon\to+0} \frac{I^s(f_\theta\|f_{\theta+\epsilon})}{-\epsilon^2\log\epsilon} = \frac{As(1-s)}{2} \quad \kappa=2$$

$$\lim_{\epsilon\to+0} \frac{I^s(f_\theta\|f_{\theta+\epsilon})}{\epsilon^2} = \frac{s(1-s)}{2}J_f \quad 2<\kappa,$$

where

$$J_f := \int_0^\infty f^{-1}(x)(f')^2(x)\mathrm{d}x. \tag{38}$$

These convergences are uniform for $0<s<1$.

For a real number $c>0$ satisfying (31)–(34), we define

$$I_s^+(c,f,\epsilon) := \int_c^\infty f^{1-s}(x)f^s(x+\epsilon)\mathrm{d}x - \int_c^b f(x)\mathrm{d}x + f(c)s\epsilon + \frac{s}{2}f'(c)\epsilon^2.$$

Similar to Propositions 4 and 5 is proven from Lemmas 1 and 2.

Lemma 2 *We obtain*

$$\lim_{\epsilon \to +0} \frac{I_s^+(c, f, \epsilon)}{\epsilon^2} = -\frac{s(1-s)}{2} J_{f,c}^+ \tag{39}$$

where

$$J_{f,c}^+ := \int_c^\infty f^{-1}(x)(f'(x))^2 \mathrm{d}x$$

and the convergence of (39) *is uniform for* $0 < s < 1$.

4 Large deviation type performances of useful estimators

In this section, in order to discuss the bounds $\alpha_1(\theta)$ and $\alpha_2(\theta)$ in the case of location shift family, we focus on the following estimators:

For the first criterion $\alpha_1(\theta)$, we focus on the likelihood ratio estimator $\boldsymbol{\theta}_{LR,\epsilon} := \{\theta_{LR,\epsilon,n}(x_1, \dots, x_n)\}$ and the estimator $\underline{\boldsymbol{\theta}}_\epsilon := \{\underline{\theta}_{\epsilon,n} := \underline{\theta}_n - \epsilon\}$. For the second criterion $\alpha_2(\theta)$, we focus on the maximum likelihood estimator $\boldsymbol{\theta}_{ML} := \{\theta_{ML,n}\}$, and the two extremal estimators $\overline{\theta}_n := \max\{x_1, \dots, x_n\} - b$ and $\underline{\theta}_n := \min\{x_1, \dots, x_n\} - a$. In order to strike a balance between two exponential rates $\beta^+(\mathbf{T}, \theta, \epsilon)$ and $\beta^-(\mathbf{T}, \theta, \epsilon)$, we use the convex combination (CC) estimator $\boldsymbol{\theta}_{CC,\lambda} := \{\theta_{CC,\lambda,n} := \lambda\underline{\theta}_n + (1-\lambda)\overline{\theta}_n\}$ with the ratio $\lambda : 1-\lambda$ of the two estimators $\underline{\boldsymbol{\theta}}$ and $\overline{\boldsymbol{\theta}}$, where $0 < \lambda < 1$. The following formulas are obtained in Sect. 5 of Hayashi (2008):

First, we discuss the performances related to the first criterion $\alpha_1(\theta)$. The modified extremal estimator $\underline{\boldsymbol{\theta}}_\epsilon$ satisfies

$$\beta(\underline{\boldsymbol{\theta}}_\epsilon, \theta, \epsilon) \cong A_1 \frac{2^{\kappa_1}}{\kappa_1} \epsilon^{\kappa_1} + o(\epsilon^{\kappa_1}). \tag{40}$$

Further, when $f(x)$ is monotonically decreasing, it also satisfies

$$\lim_{\epsilon \to +0} \frac{\beta(\underline{\boldsymbol{\theta}}_\epsilon, \theta, \epsilon)}{g(\epsilon)} = \overline{\alpha}_1(\theta). \tag{41}$$

When the function $\log f(x)$ is concave, the likelihood ratio estimator $\boldsymbol{\theta}_{LR,\epsilon}$ satisfies the equation

$$\lim_{\epsilon \to +0} \frac{\beta(\boldsymbol{\theta}_{LR,\epsilon}, \theta, \epsilon)}{g(\epsilon)} = \overline{\alpha}_1(\theta). \tag{42}$$

Next, we discuss the performances related to the second criterion $\alpha(\theta)$. Two extremal estimators $\underline{\boldsymbol{\theta}}$ and $\overline{\boldsymbol{\theta}}$ satisfy

$$\beta(\underline{\boldsymbol{\theta}}, \theta, \epsilon) \cong \frac{A_1}{\kappa_1} \epsilon^{\kappa_1} + o(\epsilon^{\kappa_1}), \quad \beta(\overline{\boldsymbol{\theta}}, \theta, \epsilon) \cong \frac{A_2}{\kappa_2} \epsilon^{\kappa_2} + o(\epsilon^{\kappa_2}). \tag{43}$$

When $\kappa_1 = \kappa_2 = \kappa$, the convex combination estimator $\boldsymbol{\theta}_{CC,n}$ satisfies the equations

$$\beta^+(\boldsymbol{\theta}_{CC,\lambda}, \theta, \epsilon) \cong \frac{A_1}{\kappa\lambda}\epsilon^\kappa + o(\epsilon^\kappa), \quad \beta^-(\boldsymbol{\theta}_{CC,\lambda}, \theta, \epsilon) \cong \frac{A_2}{\kappa(1-\lambda)}\epsilon^\kappa + o(\epsilon^\kappa).$$

When $\lambda_0 := \frac{A_1^{\frac{1}{\kappa}}}{A_1^{\frac{1}{\kappa}} + A_2^{\frac{1}{\kappa}}} = \operatorname{argmax}_{0 \le \lambda \le 1} \min\{\frac{A_1}{\kappa\lambda}, \frac{A_2}{\kappa(1-\lambda)}\}$, the relations

$$\beta^+(\boldsymbol{\theta}_{CC,\lambda_0}, \theta, \epsilon) \cong \beta^-(\boldsymbol{\theta}_{CC,\lambda_0}, \theta, \epsilon) \cong \frac{(A_1^{\frac{1}{\kappa}} + A_2^{\frac{1}{\kappa}})^\kappa}{\kappa}\epsilon^\kappa + o(\epsilon^\kappa) \tag{44}$$

hold. When the function $x \mapsto \log f(x)$ is concave, the MLE $\boldsymbol{\theta}_{ML}$ satisfies

$$\alpha(\boldsymbol{\theta}_{ML}, \theta) \ge \frac{\overline{\alpha}_1(\theta)}{2^\kappa}. \tag{45}$$

Under the same assumption, Fu (1973) essentially proved

$$\beta^+(\boldsymbol{\theta}_{ML}, \theta, \epsilon) = \sup_{t \ge 0} - \log \int_a^{b-\epsilon} \exp\left(-t\frac{f'(x)}{f(x)}\right) f(x+\epsilon)\mathrm{d}x. \tag{46}$$

Further, when $f(x)$ is monotonically decreasing, the MLE $\boldsymbol{\theta}_{ML}$ satisfies

$$\beta(\boldsymbol{\theta}_{ML}, \theta, \epsilon) \cong \frac{A_1}{\kappa_1}\epsilon^{\kappa_1} + o(\epsilon^{\kappa_1}). \tag{47}$$

5 Large deviation bounds depending on the shape parameter κ_1 and κ_2

In this section, we calculate the bounds $\alpha_1(\theta)$, $\overline{\alpha}_1(\theta)$ $\alpha_2(\theta)$, and $\overline{\alpha}_2(\theta)$ in the case of location shift family, depending on the shape parameter κ_1 and κ_2. Using facts in Sect. 4, we can calculate $\lim_{\epsilon \to +0} \frac{\beta(\boldsymbol{\theta}_{LR,\epsilon}, \theta, \epsilon)}{g(\epsilon)}$, $\lim_{\epsilon \to +0} \frac{\beta(\underline{\boldsymbol{\theta}}_\epsilon, \theta, \epsilon)}{g(\epsilon)}$, $\max_\lambda \alpha(\boldsymbol{\theta}_{CC,\lambda}, \theta)$, and $\alpha(\boldsymbol{\theta}_{ML}, \theta)$. Before proceeding to the discussion of this section, we summarize these calculations and the main results concerning the bounds $\alpha_1(\theta)$, $\overline{\alpha}_1(\theta)$ $\alpha_2(\theta)$, and $\overline{\alpha}_2(\theta)$ by Table 1.

In Tables 1, 2 and 3, (C) means that the value is calculated only when $\log f(x)$ is concave. (D) does that the value is calculated only when $f(x)$ is monotonically decreasing. In the above table, we consider the case of $A_1 \ge A_2 > 0$. In this case, there is no example such that f is monotonically decreasing and $\kappa \ne 1$. Also, there is no example such that $\log f(x)$ is concave and $1 > \kappa > 0$.

Table 1 is simplified in the case $A_1 = A_2 = A$ as Table 2.

Further, when $A_2 = 0$ or the support of f is the half line $(0, \infty)$ and when $A_1 = A$, the table is simplified as Table 3.

Table 1 Slopes in the general case ($\kappa_1 = \kappa_2 = \kappa$)

κ	$\kappa > 2$	$\kappa = 2$	$2 > \kappa > 1$	$\kappa = 1$	$1 > \kappa > 0$
$g(\epsilon)$	ϵ^2	$-\epsilon^2 \log \epsilon$	ϵ^κ	ϵ	ϵ^κ
$\overline{\alpha}_1(\theta)$	J_θ	$\frac{A_1+A_2}{2}$	?	$2\max\{A_1, A_2\}$	?
$\alpha_1(\theta)$	J_θ (C)	$\frac{A_1+A_2}{2}$ (C)	$\overline{\alpha}_1(\theta)$ (C)	$2\max\{A_1, A_2\}$	?
$\lim_{\epsilon\to+0} \frac{\beta(\boldsymbol{\theta}_{LR,\epsilon},\theta,\epsilon)}{g(\epsilon)}$	J_θ (C)	$\frac{A_1+A_2}{2}$ (C)	$\overline{\alpha}_1(\theta)$ (C)	$2\max\{A_1, A_2\}$(C)	?
$\lim_{\epsilon\to+0} \frac{\beta(\underline{\boldsymbol{\theta}}_{\epsilon},\theta,\epsilon)}{g(\epsilon)}$	0	0	$A_1 \frac{2^\kappa}{\kappa}$	$2A_1$	$A_1 \frac{2^\kappa}{\kappa}$
$\overline{\alpha}_2(\theta)$	J_θ	$\frac{A_1+A_2}{2}$	?	$A_1 + A_2$	?
$\alpha_2(\theta)$	J_θ (C)	$\frac{A_1+A_2}{2}$ (C)	?	$A_1 + A_2$	?
$\alpha(\boldsymbol{\theta}_{ML}, \theta)$	J_θ (C)	$\frac{A_1+A_2}{2}$ (C)	$\geq \frac{\overline{\alpha}_1(\theta)}{2^\kappa}$ (C)	A_1 (D)	?
$\max_\lambda \alpha(\boldsymbol{\theta}_{CC,\lambda}, \theta)$	0	0	$\frac{(A_1^{\frac{1}{\kappa}}+A_2^{\frac{1}{\kappa}})^\kappa}{\kappa}$	$A_1 + A_2$	$\frac{(A_1^{\frac{1}{\kappa}}+A_2^{\frac{1}{\kappa}})^\kappa}{\kappa}$

Table 2 Slopes when $A_1 = A_2 = A$

κ	$\kappa > 2$	$\kappa = 2$	$2 > \kappa > 1$	$\kappa = 1$	$1 > \kappa > 0$
$g(\epsilon)$	ϵ^2	$-\epsilon^2 \log \epsilon$	ϵ^κ	ϵ	ϵ^κ
$\overline{\alpha}_1(\theta) = \overline{\alpha}_2(\theta)$	J_θ	A	$\frac{A2^{\kappa-1}(3-\kappa)B(\frac{1+\kappa}{2},2-\kappa)}{\kappa}$	$2A$	$\frac{A2^{\kappa}(1-\kappa)B(\frac{1+\kappa}{2},1-\kappa)}{\kappa}$
$\alpha_1(\theta)$	J_θ (C)	A (C)	$\overline{\alpha}_1(\theta)$ (C)	$2A$	?
$\lim_{\epsilon\to+0} \frac{\beta(\boldsymbol{\theta}_{LR,\epsilon},\theta,\epsilon)}{g(\epsilon)}$	J_θ (C)	A (C)	$\overline{\alpha}_1(\theta)$ (C)	$2A$ (C)	?
$\lim_{\epsilon\to+0} \frac{\beta(\underline{\boldsymbol{\theta}}_{\epsilon},\theta,\epsilon)}{g(\epsilon)}$	0	0	$A \frac{2^\kappa}{\kappa}$	$2A$	$A \frac{2^\kappa}{\kappa}$
$\alpha_2(\theta)$	J_θ (C)	A (C)	?	$2A$	?
$\alpha(\boldsymbol{\theta}_{ML}, \theta)$	J_θ (C)	A (C)	$\geq \frac{\overline{\alpha}_1(\theta)}{2^\kappa}$ (C)	A (D)	?
$\alpha(\boldsymbol{\theta}_{CC,1/2}, \theta)$	0	0	$\frac{2^\kappa A}{\kappa}$	$2A$	$\frac{2^\kappa A}{\kappa}$

Table 3 Slopes when $A_1 = A$ and $A_2 = 0$ or the half-line case. t_0 will be given in (63)

κ	$\kappa > 2$	$\kappa = 2$	$2 > \kappa > 1$	$\kappa = 1$	$1 > \kappa > 0$
$g(\epsilon)$	ϵ^2	$-\epsilon^2 \log \epsilon$	ϵ^κ	ϵ	ϵ^κ
$\overline{\alpha}_1(\theta)$	J_θ	$\frac{A}{2}$	$\frac{A2^\kappa}{\kappa}$ ($\kappa \leq 2 - t_0$)	$2A$	$\frac{A2^\kappa}{\kappa}$
$\alpha_1(\theta)$	J_θ (C)	$\frac{A}{2}$ (C)	$\overline{\alpha}_1(\theta)$ (C or $\kappa \leq 2 - t_0$)	$2A$	$\frac{A2^\kappa}{\kappa}$
$\lim_{\epsilon\to+0} \frac{\beta(\boldsymbol{\theta}_{LR,\epsilon},\theta,\epsilon)}{g(\epsilon)}$	J_θ (C)	$\frac{A}{2}$ (C)	$\overline{\alpha}_1(\theta)$ (C)	$2A$(C)	?
$\lim_{\epsilon\to+0} \frac{\beta(\underline{\boldsymbol{\theta}}_{\epsilon},\theta,\epsilon)}{g(\epsilon)}$	0	0	$A \frac{2^\kappa}{\kappa}$	$2A$	$A \frac{2^\kappa}{\kappa}$
$\overline{\alpha}_2(\theta)$	J_θ	$\frac{A}{2}$	$\leq \frac{A\pi(\kappa-1)}{\kappa \sin \pi(\kappa-1)}$	A	$\frac{A}{\kappa}$
$\alpha_2(\theta)$	J_θ (C)	$\frac{A}{2}$ (C)	?	A	$\frac{A}{\kappa}$
$\alpha(\boldsymbol{\theta}_{ML}, \theta)$	J_θ (C)	$\frac{A}{2}$ (C)	$\geq \frac{A}{\kappa}$ (C)	A (C) or (D)	$\frac{A}{\kappa}$ (D)
$\alpha(\underline{\boldsymbol{\theta}}, \theta)$	0	0	$\frac{A}{\kappa}$	A	$\frac{A}{\kappa}$

5.1 Semi-regular case

From Proposition 4, when $\kappa_1, \kappa_2 > 2$, the relation

$$\lim_{\epsilon \to +0} \frac{I^s(f_\theta \| f_{\theta+\epsilon})}{\epsilon^2} = \frac{s(1-s)}{2} J_f \tag{48}$$

holds, where this convergence is uniform for s and J_f is defined by

$$J_f := \int_a^b \left(\frac{\mathrm{d} f(x)}{\mathrm{d}x} \right)^2 f(x)^{-1} \mathrm{d}x.$$

Hence, it is suitable to choose $g(x) = x^2$.

Theorem 1 *When $x \mapsto \log f(x)$ is concave, we obtain $\kappa = 2$ and the relation*

$$\overline{\alpha}_1(\theta) = \alpha_1(\theta) = \overline{\alpha}_2(\theta) = \alpha_2(\theta) = \frac{1}{2} J_\theta. \tag{49}$$

Proof Using (48), we have

$$\begin{aligned}
\overline{\alpha}_1(\theta) &= 4 \max_{0 \le s \le 1} I^s_{g,\theta} = \frac{1}{2} J_\theta \\
\overline{\alpha}_2(\theta) &= \min_{0 \le s \le 1} \frac{I^s_{g,\theta}}{s(1-s)} = \frac{1}{2} J_\theta.
\end{aligned}$$

Since the function $\log f(x)$ is concave, relation (46) implies

$$\begin{aligned}
&\beta^+(\boldsymbol{\theta}_{ML}, \theta, \epsilon) \\
&\ge -\log \int_a^{b-\epsilon} \exp\left(-\epsilon \frac{f'(x)}{f(x)} \right) f(x+\epsilon) \mathrm{d}x \\
&\cong -\log \left(\int_a^{b-\epsilon} \left(1 - \epsilon \frac{f'(x)}{f(x)} + \frac{\epsilon^2}{2} \left(\frac{f'(x)}{f(x)} \right)^2 \right) \right. \\
&\quad \left. \times \left(f(x) + \epsilon f'(x) + \frac{\epsilon^2}{2} f''(x) \right) \mathrm{d}x + o(\epsilon^2) \right) \\
&\cong -\log \left(\int_a^{b-\epsilon} f(x) - \frac{\epsilon^2}{2} \left(\frac{f'(x)^2}{f(x)} + f''(x) \right) \mathrm{d}x + o(\epsilon^2) \right) \\
&\cong -\log \left(1 - \int_{b-\epsilon}^b f(x) \mathrm{d}x - \frac{\epsilon^2}{2} \int_a^{b-\epsilon} \frac{f'(x)^2}{f(x)} \mathrm{d}x + o(\epsilon^2) \right) \\
&\cong -\log \left(1 - \frac{\epsilon^2}{2} \left(\int_a^{b-\epsilon} \frac{f'(x)^2}{f(x)} \mathrm{d}x - f'(b-\epsilon) \right) + o(\epsilon^2) \right).
\end{aligned}$$

Thus, we obtain

$$\lim_{\epsilon\to 0}\frac{1}{\epsilon^2}\beta^+(\boldsymbol{\theta}_{ML},\theta,\epsilon)\geq\frac{1}{2}\lim_{\epsilon\to 0}\left(\int_a^{b-\epsilon}\frac{f'(x)^2}{f(x)}\mathrm{d}x-f'(b-\epsilon)\right)=\frac{1}{2}J_f.$$

Similarly, we can prove that

$$\lim_{\epsilon\to 0}\frac{1}{\epsilon^2}\beta^-(\boldsymbol{\theta}_{ML},\theta,\epsilon)\geq\frac{1}{2}J_f.$$

Hence, the relation $\alpha_2(\boldsymbol{\theta}_{ML},\theta)\leq\frac{1}{2}J_f$ yields the equation

$$\alpha_2(\boldsymbol{\theta}_{ML},\theta)=\frac{1}{2}J_f,$$

which implies (49). □

5.2 The case that $\kappa_1=\kappa_2=1a$

From Proposition 4, when $\kappa_1=\kappa_2=1$, the equation

$$\lim_{\epsilon\to+0}\frac{I^s(f_\theta\|f_{\theta+\epsilon})}{\epsilon}=A_1s+A_2(1-s)$$

holds, where this convergence is uniform for $s\in(0,1)$. Letting $g(x)=|x|$, we have $\kappa=1$.

Theorem 2 *The relations*

$$\alpha_1(\theta)=\overline{\alpha}_1(\theta)=2\sup_{0<s<1}I^s_{g,\theta}=2\max\{A_1,A_2\}\tag{50}$$

$$\alpha_2(\theta)=\overline{\alpha}_2(\theta)=2I^{\frac{1}{2}}_{g,\theta}=A_1+A_2\tag{51}$$

hold. Therefore, $\alpha_1(\theta)=\alpha_2(\theta)$ *if and only if* $A_1=A_2$.

Proof The second and third equations of (50) and (51) follow from the formula $I^s_{g,\theta}=A_1s+A_2(1-s)$. In the following, we prove the first equations of (50) and (51). Since (44) implies

$$\max_{0\leq\lambda\leq 1}\alpha_2(\boldsymbol{\theta}_{CC,\lambda},\theta)=A_1+A_2,$$

we obtain the first equation of (51).

Next, we prove the first equation of (50) in the case where $A_1\geq A_2$. From (40), the estimator $\underline{\boldsymbol{\theta}}_\epsilon$ satisfies

$$\lim_{\epsilon\to+0}\frac{1}{\epsilon}\beta(\underline{\boldsymbol{\theta}}_\epsilon,\theta,\epsilon)=2A_1,$$

which implies the first equation of (50). When $A_2 \geq A_1$, we can similarly prove it. □

5.3 The case that $\kappa_1 = \kappa_2 = 2$

From Proposition 4, when $\kappa_1 = \kappa_2 = 2$, the equation

$$\lim_{\epsilon \to +0} \frac{I^s(f_\theta \| f_{\theta+\epsilon})}{-\epsilon^2 \log |\epsilon|} = \frac{(A_1 + A_2)s(1-s)}{2} \tag{52}$$

holds, where this convergence is uniform for $s \in (0, 1)$. When we choose $g(x) = -x^2 \log x$, the parameter κ defined in (11) is equal to 2. Then, Eq. (52) implies

$$\begin{aligned} \overline{\alpha}_1(\theta) &= 4 \sup_{0<s<1} I^s_{g,\theta} = \frac{A_1 + A_2}{2} \\ \overline{\alpha}_2(\theta) &= \inf_{0<s<1} \frac{I^s_{g,\theta}}{s(1-s)} = \frac{A_1 + A_2}{2}. \end{aligned} \tag{53}$$

Theorem 3 *When the function $x \mapsto \log f(x)$ is concave, the relations*

$$\alpha_1(\theta) = \overline{\alpha}_1(\theta) = \alpha_2(\theta) = \overline{\alpha}_2(\theta) = \frac{A_1 + A_2}{2} \tag{54}$$

hold.

Proof When $a + \delta < x < b - \delta$, we can approximate that

$$\exp\left(-\epsilon \frac{f'(x)}{f(x)}\right) f(x+\epsilon) \cong f(x) + \frac{\epsilon^2}{2}\left(f''(x) - \frac{(f'(x))^2}{f(x)}\right). \tag{55}$$

Therefore, from relation (46), we can evaluate

$$\begin{aligned} &\beta^+(\boldsymbol{\theta}_{ML}, \theta, \epsilon) \\ &\geq -\log\left(\int_a^{b-\epsilon} \exp\left(-\epsilon \frac{f'(x)}{f(x)}\right) f(x+\epsilon)\mathrm{d}x\right) \\ &\cong -\log\left(\int_{a+\delta}^{b-\delta} \exp\left(-\epsilon \frac{f'(x)}{f(x)}\right) f(x+\epsilon)\mathrm{d}x + \int_0^\delta \exp\left(-\epsilon \frac{A_1}{A_1 x}\right) A_1(x+\epsilon)\mathrm{d}x\right. \\ &\quad \left. + \int_\epsilon^\delta \exp\left(+\epsilon \frac{A_2}{A_2 x}\right) A_2(x-\epsilon)\mathrm{d}x\right) \\ &\cong -\log\left(\int_{a+\delta}^{b-\delta} f(x)\mathrm{d}x + \int_a^{a+\delta} f(x)\mathrm{d}x + \int_{b-\delta}^b f(x)\mathrm{d}x\right. \\ &\quad + A_1 \int_0^\delta \left(\exp\left(-\epsilon \frac{1}{x}\right)(x+\epsilon) - x\right)\mathrm{d}x \\ &\quad \left. + A_2 \int_\epsilon^\delta \left(\exp\left(+\epsilon \frac{1}{x}\right)(x-\epsilon) - x\right)\mathrm{d}x + o(-\epsilon^2 \log \epsilon)\right) \end{aligned}$$

$$
\begin{aligned}
&\cong -\log\left(1+\frac{A_1+A_2}{2}\epsilon^2\log\epsilon+o(-\epsilon^2\log\epsilon)\right)\\
&\cong -\frac{A_1+A_2}{2}\epsilon^2\log\epsilon+o(-\epsilon^2\log\epsilon), \qquad (56)
\end{aligned}
$$

where the relation (56) follows from Lemma 6 in Appendix A. Thus,

$$
\alpha_2(\boldsymbol{\theta}_{ML},\theta)=\frac{A_1+A_2}{2}. \qquad (57)
$$

Using (53) and (57), we obtain (54). □

5.4 The case that $1<\kappa_1=\kappa_2<2$

From Proposition 4, when $1<\kappa_1=\kappa_2<2$, the equation

$$
\lim_{\epsilon\to+0}\frac{1}{\epsilon^{\kappa_1}}I^s(f_\theta\|f_{\theta+\epsilon})=\frac{A_1 s(1-s(\kappa_1-1))B(s+\kappa_1(1-s),2-\kappa_1)}{\kappa_1}+\frac{A_2(1-s)(1-(1-s)(\kappa_1-1))B(1-s+\kappa_1 s,2-\kappa_1)}{\kappa_1}
$$

holds, where this convergence is uniform for $s\in(0,1)$, and $B(x,y)$ is a beta function. Letting $g(x)=|x|^{\kappa_1}$, we have $\kappa=\kappa_1$.

$$
\begin{aligned}
\overline{\alpha}_1(\theta)&=2^\kappa\sup_{0<s<1}I^s_{g,\theta}=\frac{2^\kappa}{\kappa}\max_{0\le s\le 1}[A_1 s(1-s(\kappa-1))B(s+\kappa(1-s),2-\kappa)\\
&\quad+A_2(1-s)(1-(1-s)(\kappa-1))B(1-s+\kappa s,2-\kappa)]\\
\overline{\alpha}_2(\theta)&=\inf_{0<s<1}\frac{I^s_{g,\theta}}{s(1-s)}\left(s^{\frac{1}{\kappa-1}}+(1-s)^{\frac{1}{\kappa-1}}\right)^{\kappa-1}=\inf_{0<s<1}\Bigg[(A_1 s(1-s(\kappa-1))B(s+\kappa(1-s),2-\kappa)\\
&\quad+A_2(1-s)(1-(1-s)(\kappa-1))B(1-s+\kappa s,2-\kappa))\\
&\quad\times\frac{\left(s^{\frac{1}{\kappa-1}}+(1-s)^{\frac{1}{\kappa-1}}\right)^{\kappa-1}}{\kappa s(1-s)}\Bigg]\\
&\le\frac{1}{\kappa}(A_1+A_2)2^{\kappa-2}(1-\kappa)B\left(\frac{1+\kappa}{2},2-\kappa\right), \qquad (58)
\end{aligned}
$$

where the last inequality is obtained by substituting $s=\frac{1}{2}$.

Theorem 4 *If, and only if,* $A_1=A_2$*, the equality*

$$
\overline{\alpha}_1(\theta)=\overline{\alpha}_2(\theta)
$$

holds. In this case,

$$\overline{\alpha}_1(\theta) = \overline{\alpha}_2(\theta) = \frac{A_1 2^{\kappa-1}(3-\kappa)B\left(\frac{1+\kappa}{2}, 2-\kappa\right)}{\kappa}. \tag{59}$$

Proof When $A_1 \neq A_2$, (73) and (74) in Lemma 3 guarantee that

$$\frac{\mathrm{d}}{\mathrm{d}s} I^s_{g,\theta}\bigg|_{s=\frac{1}{2}} = (A_1 - A_2)\frac{(\kappa-1)(3-\kappa)}{4}\pi \tan\frac{2-\kappa}{2}\pi B\left(\frac{1+\kappa}{2}, 2-\kappa\right) \neq 0.$$

From the concavity and the continuity of the maximized function, we have $\overline{\alpha}_1(\theta) > 2^\kappa I^{\frac{1}{2}}_{g,\theta}$. When $A_1 = A_2$, we have $I^s_{g,\theta} = I^{1-s}_{g,\theta}$. The relations

$$\overline{\alpha}_1(\theta) = 2^\kappa I^{\frac{1}{2}}_{g,\theta} = \frac{A_1 2^{\kappa-1}(3-\kappa)B\left(\frac{1+\kappa}{2}, 2-\kappa\right)}{\kappa} \tag{60}$$

follow from the concavity. Since the minimums

$$\min_{0\le s\le 1}\left(s^{\frac{1}{\kappa-1}} + (1-s)^{\frac{1}{\kappa-1}}\right)$$

and

$$\min_{0\le s\le 1}\left(\frac{(1-s(\kappa-1))B(s+\kappa(1-s), 2-\kappa)}{1-s} + \frac{(1-(1-s)(\kappa-1))B((1-s)+\kappa s, 2-\kappa)}{s}\right)$$

are achieved at the same point $s = \frac{1}{2}$ (See Lemma 5), the relation

$$\begin{aligned}
&\min_{0\le s\le 1}\left(s^{\frac{1}{\kappa-1}} + (1-s)^{\frac{1}{\kappa-1}}\right)^{\kappa-1}\left(\frac{(1-s(\kappa-1))B(s+\kappa(1-s), 2-\kappa)}{1-s}\right. \\
&\quad \left. + \frac{(1-(1-s)(\kappa-1))B((1-s)+\kappa s, 2-\kappa)}{s}\right) \\
&= 2^{\kappa-1}(3-\kappa)B\left(\frac{1+\kappa}{2}, 2-\kappa\right)
\end{aligned} \tag{61}$$

holds. Thus, Eq. (59) follows from (60) and (61). □

Next, we consider the case $A_2 = 0$. Substituting 0 into s, we obtain

$$\begin{aligned}
\overline{\alpha}_2(\theta) &= \inf_{0<s<1}\left[\frac{A_1}{\kappa}\frac{(1-s(\kappa-1))}{1-s}B(\kappa-(\kappa-1)s, 2-\kappa)\left(s^{\frac{1}{\kappa-1}} + (1-s)^{\frac{1}{\kappa-1}}\right)^{\kappa-1}\right] \\
&\ge \frac{A_1}{\kappa}B(\kappa, 2-\kappa).
\end{aligned}$$

Concerning $\overline{\alpha}_1(\theta)$, we have the following proposition.

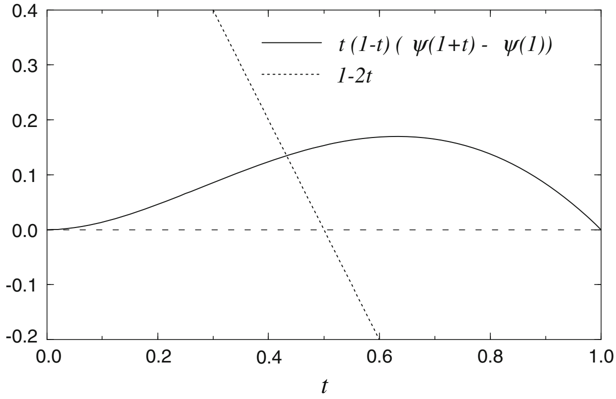

Fig. 1 Functions $t(1-t)\,(\psi(1+t)-\psi(1))$ and $1-2t$

Theorem 5 *The inequality*

$$\overline{\alpha}_1(\theta) \geq A_1 \frac{2^\kappa}{\kappa} \tag{62}$$

holds. When $1 < \kappa < 2 - t_0$*, the above equality holds, where the real number* $t_0 \in (0, \frac{1}{2})$ *is uniquely defined by* (*see Lemma* 4 *and Fig.* 1)

$$2t_0 + t_0(1-t_0)(\psi(1+t_0) - \psi(1)) = 1, \tag{63}$$

where $\psi(x)$ *is the D-psi function defined by* $\psi(x) := \frac{\mathrm{d}}{\mathrm{d}x} \log \Gamma(x)$.

The number t_0 is enumerated by $t_0 \cong 0.432646$, as is checked by the following graph.

Further, when $x \mapsto \log f(x)$ is concave, from (40), we have

$$\lim_{\epsilon \to +0} \frac{\alpha(\boldsymbol{\theta}_{LR,\epsilon}, \theta, \epsilon)}{\epsilon^\kappa} = \frac{A_1 2^\kappa}{\kappa}. \tag{64}$$

Hence, when $1 < \kappa < 2 - t_0$,

$$\alpha_1(\theta) = \overline{\alpha}_1(\theta) = \frac{A_1 2^\kappa}{\kappa}. \tag{65}$$

Proof Inequality follows by substituting 1 into s. In this case, since the function $s \mapsto s(1 - s(\kappa - 1))B(s + \kappa(1-s), 2-\kappa)$ is concave, Lemma 4 guarantees that

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}s}\, s(1 - s(\kappa - 1))B(s + \kappa(1-s), 2-\kappa)) \\ &\quad \geq \frac{\mathrm{d}}{\mathrm{d}s} s(1 - s(\kappa - 1))B(s + \kappa(1-s), 2-\kappa)|_{s=1} \\ &\quad = (3 - 2\kappa) + (2-\kappa)(\kappa - 1)(\psi(3-\kappa) - \psi(1)) \geq 0 \end{aligned}$$

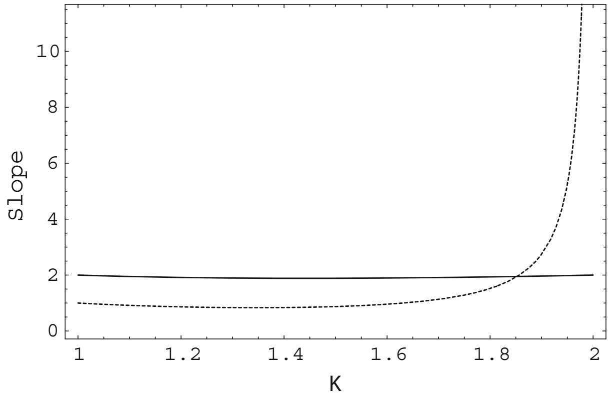

Fig. 2 Slopes when $A_1 = A_2 = A = 1$. *solid curve* $\alpha(\boldsymbol{\theta}_{CC,1/2}, \theta) = \frac{A2^\kappa}{\kappa}$, *dots curve* $\frac{A(3-\kappa)B(\frac{1+\kappa}{2}, 2-\kappa)}{2\kappa}$

for $s \in (0, 1)$. Thus,

$$\begin{aligned}\overline{\alpha}_1(\theta) &= \frac{A_1 2^\kappa}{\kappa} \max_{0 \le s \le 1} s(1 - s(\kappa - 1))B(s + \kappa(1-s), 2-\kappa) \\ &= \frac{A_1 2^\kappa}{\kappa} 1(1 - 1(\kappa - 1))B(1 + \kappa(1-1), 2-\kappa) = \frac{A_1 2^\kappa}{\kappa},\end{aligned}$$

which implies the equality. □

Next, we compare $\alpha(\boldsymbol{\theta}_{CC,1/2}, \theta) = \frac{A2^\kappa}{\kappa}$ and $\alpha(\boldsymbol{\theta}_{ML}, \theta)$ in the case of $A_1 = A_2$. When $\log f(x)$ is concave, $\alpha(\boldsymbol{\theta}_{ML}, \theta)$ has the lower bound $\frac{A(3-\kappa)B(\frac{1+\kappa}{2}, 2-\kappa)}{2\kappa}$. As is illustrated by Fig. 2,

$$\frac{A(3-\kappa)B\left(\frac{1+\kappa}{2}, 2-\kappa\right)}{2\kappa} \ge \frac{A2^\kappa}{\kappa} \tag{66}$$

if and only if $\kappa \ge \kappa_0 = 1.85238799$. That is, for $\kappa \ge \kappa_0$, the maximum likelihood estimator $\boldsymbol{\theta}_{ML}$ is better than the convex combination estimator $\boldsymbol{\theta}_{CC,1/2}$.

5.5 The case that $0 < \kappa_1 = \kappa_2 < 1$

From Proposition 4, when $0 < \kappa_1 = \kappa_2 < 1$, the equation

$$\begin{aligned}\lim_{\epsilon \to +0} \frac{I^s(f_\theta \| f_{\theta+\epsilon})}{\epsilon^{\kappa_1}} = \frac{1-\kappa_1}{\kappa_1} & (A_1 s B(s + \kappa_1(1-s), 1-\kappa_1) \\ & + A_2(1-s)B(1 - s + \kappa_1 s, 1 - \kappa_1))\end{aligned}$$

holds, where this convergence is uniform for $s \in (0, 1)$. Letting $g(x) = x_1^\kappa$, we have $\kappa = \kappa_1$.

$$\begin{aligned}
\overline{\alpha}_1(\theta) &= 2^\kappa \max_{0\le s\le 1} I^s_{g,\theta} \\
&= 2^\kappa \max_{0\le s\le 1} (1-\kappa) \\
&\quad \times \frac{A_1 s B(s+\kappa(1-s), 1-\kappa) + A_2(1-s)B(1-s+\kappa s, 1-\kappa)}{\kappa} \\
\overline{\alpha}_2(\theta) &= \sup_{0<s<1} \frac{I^s_{g,\theta}}{s(1-s)} \left(s^{\frac{1}{\kappa-1}} + (1-s)^{\frac{1}{\kappa-1}}\right)^{\kappa-1} \\
&= \sup_{0<s<1} (1-\kappa) \frac{A_1 s B(s+\kappa(1-s), 1-\kappa) + A_2(1-s)B(1-s+\kappa s, 1-\kappa)}{\kappa s(1-s)} \\
&\quad \times \left(s^{\frac{1}{\kappa-1}} + (1-s)^{\frac{1}{\kappa-1}}\right)^{\kappa-1}.
\end{aligned}$$

Theorem 6 *If, and only if,* $A_1 = A_2$, *the equality*

$$\overline{\alpha}_1(\theta) = \overline{\alpha}_2(\theta) \tag{67}$$

holds. In this case, the equation

$$\overline{\alpha}_1(\theta) = \overline{\alpha}_2(\theta) = \frac{1}{\kappa} A_1 2^\kappa (1-\kappa) B\left(\frac{1+\kappa}{2}, 1-\kappa\right) \tag{68}$$

holds.

Proof Using (75) and (76) of Lemma 3, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}s} I^s_{g,\theta}\bigg|_{s=\frac{1}{2}} = (A_1 - A_2)\frac{1-\kappa}{2}\pi \cot \frac{1-\kappa}{2}\pi B\left(\frac{1+\kappa}{2}, 1-\kappa\right).$$

Since $\frac{1-\kappa}{2}\pi \cot \frac{1-\kappa}{2}\pi B(\frac{1+\kappa}{2}, 1-\kappa) > 0$, Proposition 3 yields this sufficient and necessary condition for (67). Equation (15) implies (68). □

However, since the function $x \mapsto (\kappa - 1)\log x$ is convex on $(0, \infty)$, the function $x \mapsto \log f(x)$ is not concave on (a, b). There does not exist an example in which (42) can be applied. Thus, it is an open problem whether there exists an example such that

$$\overline{\alpha}_1(\theta) = \alpha_1(\theta)$$

in this case, except for the case $A_1 A_2 = 0$.

In the following, we consider the case when $A_2 = 0$.

Theorem 7 *We have*

$$\alpha_1(\theta) = \overline{\alpha}_1(\theta) = \frac{A_1 2^\kappa}{\kappa} \tag{69}$$

$$\alpha_2(\theta) = \overline{\alpha}_2(\theta) = \frac{A_1}{\kappa}. \tag{70}$$

Proof Since the function $s \mapsto sB(s+\kappa(1-s), 1-\kappa)$ is concave, we have

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}s} sB(s+\kappa(1-s), 1-\kappa) &\geq \frac{\mathrm{d}}{\mathrm{d}s} sB(s+\kappa(1-s), 1-\kappa)|_{s=1} \\
&= 1 + (1-\kappa)(\psi(1) - \psi(2-\kappa)) \\
&\geq 1 + (1-\kappa)(\psi(1) - \psi(2)) \qquad (71)\\
&= 1 - (1-\kappa) = \kappa > 0, \qquad (72)
\end{aligned}$$

where inequality (71) holds because $\psi(x)$ is monotonically increasing in $x \in (0, \infty)$, and the first equation of (72) follows from the formula $\psi(x+1) = \psi(x) + \frac{1}{x}$. Thus,

$$\begin{aligned}
\overline{\alpha}_1(\theta) &= \frac{A_1 2^\kappa (1-\kappa)}{\kappa} \max_{0 \leq s \leq 1} sB(s+\kappa(1-s), 1-\kappa) \\
&= \frac{A_1 2^\kappa (1-\kappa)}{\kappa} B(1, 1-\kappa) = \frac{A_1 2^\kappa}{\kappa}.
\end{aligned}$$

From (40), we can check that the estimators $\{\underline{\theta}_\epsilon\}$ achieve the bound $\frac{A_1 2^\kappa}{\kappa}$. Hence, Eq. (69) hold.

The other upper bound $\overline{\alpha}_2(\theta)$ is calculated as

$$\overline{\alpha}_2(\theta) = \frac{A_1(1-\kappa)}{\kappa} \max_{0 \leq s \leq 1} B(s+\kappa(1-s), 1-\kappa) \left(\left(\frac{1-s}{s} \right)^{\frac{1}{1-\kappa}} + 1 \right)^{-(1-\kappa)}.$$

Note that the beta function $B(x, y)$ is monotonically decreasing for $x, y > 0$. Since both $\max_{0 \leq s \leq 1} B(s+\kappa(1-s), 1-\kappa)$ and $\max_{0 \leq s \leq 1} ((\frac{1-s}{s})^{\frac{1}{1-\kappa}} + 1)^{-(1-\kappa)}$ are achieved at the same point, $s = 1$, we have

$$\overline{\alpha}_2(\theta) = \frac{A_1(1-\kappa)}{\kappa} B(1, 1-\kappa) = \frac{A_1}{\kappa}.$$

This bound is achieved by the estimator $\underline{\boldsymbol{\theta}}$ because of (43). Therefore, we have (70). □

6 Examples and graphs

In the previous section, we treat our problem separating our cases by the shape parameters κ_1 and κ_2. In this section, we treat it separating our cases into two cases, i.e., the case $A_1 = A_2$ and the case $A_1 = A$, $A_2 = 0$. That is, we fixed the parameters A_1 and

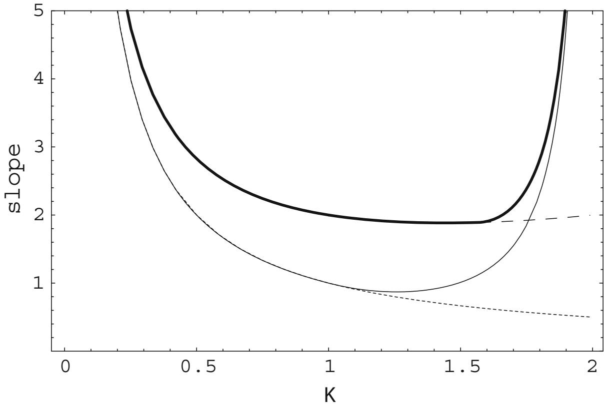

Fig. 3 Slopes when $A_1 = 1$ and $A_2 = 0$, *thick curve* $\overline{\alpha}_1(\theta)$, *solid curve* $\overline{\alpha}_2(\theta)$, *dash curve* $\lim_{\epsilon \to +0} \frac{\beta(\underline{\boldsymbol{\theta}}_\epsilon, \theta, \epsilon)}{\epsilon^\kappa}$, *dots curve* $\alpha(\underline{\boldsymbol{\theta}}, \theta)$

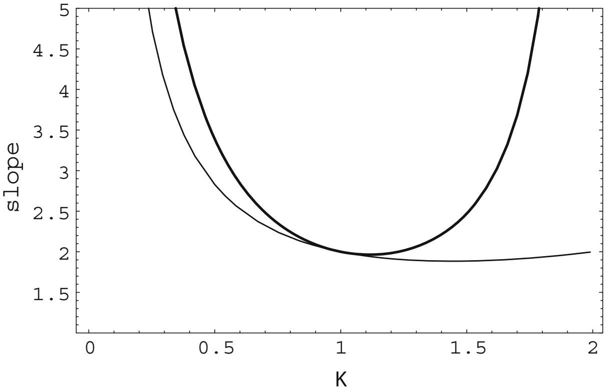

Fig. 4 Slopes when $A_1 = A_2 = 1$, *thick curve* $\overline{\alpha}_2(\theta) = \overline{\alpha}_1(\theta)$, *solid curve* $\alpha(\boldsymbol{\theta}_{CC,1/2}, \theta)$

A_2, and focus on the dependence of the slopes on the shape parameters κ_1 and κ_2. When $A_1 = A = 1$ and $A_2 = 0$, in the latter case, the slopes $\overline{\alpha}_1(\theta), \overline{\alpha}_2(\theta), \alpha(\underline{\boldsymbol{\theta}}, \theta)$, and the limit $\lim_{\epsilon \to +0} \frac{\beta(\underline{\boldsymbol{\theta}}_\epsilon, \theta, \epsilon)}{\epsilon^\kappa}$ are illustrated as the function of $\kappa \in (0, 2)$ in Fig. 3. This case is essentially equivalent with the half-line-support case. When $A_1 = A_2 = 1$, the slopes $\overline{\alpha}_1(\theta), \overline{\alpha}_2(\theta)$, and $\alpha(\boldsymbol{\theta}_{CC,1/2}, \theta)$ are illustrated as the function of $\kappa \in (0, 2)$ in Fig. 4.

Next, we treat the same problem in three cases, the Gamma distribution, Weibull distribution, and Beta distribution, respectively.

Example 1 (*Gamma distribution*) Consider the case where the pdf f is the Gamma distribution function $f^{\Gamma}_{\alpha,\beta}(x) := \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}$. The case of $\alpha = 1$ is called an exponential distribution, which the time of nuclear decay obeys. In this case, the parameter κ is equal to α, and the parameter A is equal to $\frac{\beta^\alpha}{\Gamma(\alpha)}$. The uniformity in Proposition 5 for s is also satisfied.

If $\alpha \le 1$, $f^{\Gamma}_{\alpha,\beta}(x)$ is monotonically decreasing. Then, we can use (41), i.e., the modified external estimator $\underline{\boldsymbol{\theta}}_{\epsilon}$ attains $\alpha_1(\theta) = \overline{\alpha}_1(\theta) = \frac{(2\beta)^\alpha}{\alpha\Gamma(\alpha)}$. The maximum likelihood estimator $\boldsymbol{\theta}_{ML}$ has the same performance as the extremal estimator $\underline{\boldsymbol{\theta}}$. If $\alpha \ge 1$, the function $\log f^{\Gamma}_{\alpha,\beta}(x) = \log \frac{\beta^\alpha}{\Gamma(\alpha)} + (\alpha - 1)\log x - \beta x$ is concave. Then, we can use (42), i.e., the likelihood ratio estimator $\boldsymbol{\theta}_{LR,\epsilon}$ attains $\alpha_1(\theta) = \overline{\alpha}_1(\theta)$. When $\alpha > 2$, the Fisher information J_θ is $\frac{\beta^2}{\alpha-2}$. Hence, we can calculate many values in Table 3.

Example 2 (*Weibull distribution*) Consider the case where the pdf f is the Weibull distribution function $f^{W}_{\alpha,\beta}(x) := \alpha\beta x^{\alpha-1}e^{-\beta x^\alpha}$. The case of $\alpha = 2$ is called a Rayleigh distribution. This distribution describes the fading in radio wave propagation. Also it is treated in acoustics. The parameter κ is equal to α, and the parameter A is equal to $\alpha\beta$. The uniformity in Proposition 5 for s is also satisfied.

If $\alpha \le 1$, $f^{W}_{\alpha,\beta}(x)$ is monotonically decreasing. Then, we can use (41), i.e., the modified external estimator $\underline{\boldsymbol{\theta}}_{\epsilon}$ attains $\alpha_1(\theta) = \overline{\alpha}_1(\theta) = \beta 2^\alpha$. The maximum likelihood estimator $\boldsymbol{\theta}_{ML}$ has the same performance as the extremal estimator $\underline{\boldsymbol{\theta}}$. If $\alpha \ge 1$, the function $\log f^{W}_{\alpha,\beta}(x) = \log \alpha\beta + (\alpha - 1)\log x - \beta x^\alpha$ is concave. Then, we can use (42), i.e., the likelihood ratio estimator $\boldsymbol{\theta}_{LR,\epsilon}$ attains $\alpha_1(\theta) = \overline{\alpha}_1(\theta)$. When $\alpha > 2$, the Fisher information J_θ is $(\alpha-1)^2\Gamma(1-\frac{2}{\alpha})\beta^{\frac{2}{\alpha}} - 2(\alpha-1)\Gamma(2-\frac{1}{\alpha})\beta^{\frac{1}{\alpha}} + 2$. Hence, we can calculate many values in Table 3.

Example 3 (Beta distribution) Consider the case where the pdf f is the Beta distribution function $f^{B}_{\alpha,\beta}(x) := \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)}$ with the support $(0, 1)$.

The case of $\alpha = \beta = 1$ is called a uniform distribution. The parameter α corresponds to κ_1 and the other parameter β does to κ_2. The parameter A is equal to $\frac{1}{B(\alpha,\beta)}$. The uniformity in Proposition 4 for s is also satisfied.

When $\alpha = \beta > 1$, the function $\log f^{B}_{\alpha,\beta}(x) = (\alpha - 1)\log x + (\beta - 1)\log(1 - x)$ is concave. Then, we can use (42), i.e., the likelihood ratio estimator $\boldsymbol{\theta}_{LR,\epsilon}$ attains $\alpha_1(\theta) = \overline{\alpha}_1(\theta) = \frac{2^{\alpha-1}(3-\alpha)B(\frac{1+\alpha}{2},2-\alpha)}{\alpha B(\alpha,\alpha)}$. When $\alpha \ge \kappa_0$, the maximum likelihood estimator $\boldsymbol{\theta}_{ML}$ is better than the extremal estimator $\underline{\boldsymbol{\theta}}$. However, when $\alpha = 1$, the maximum likelihood estimator $\boldsymbol{\theta}_{ML}$ is better than the extremal estimator $\underline{\boldsymbol{\theta}}$. It is not clear whether $\overline{\alpha}_2(\theta)$ can be attained. When $\alpha = \beta < 1$, the function $f^{B}_{\alpha,\beta}(x)$ is monotonically decreasing. Then, we can use (41), i.e., the modified external estimator $\underline{\boldsymbol{\theta}}_{\epsilon}$ attains $\alpha_1(\theta) = \overline{\alpha}_1(\theta) = \frac{2^{\alpha}(1-\alpha)B(\frac{1+\alpha}{2},1-\alpha)}{\alpha B(\alpha,\alpha)}$. It also is not clear whether $\overline{\alpha}_2(\theta)$ can be attained. When $\alpha = \beta = 1$, The bound $\overline{\alpha}_2(\theta) = \overline{\alpha}_1(\theta) = 2$ can be attained by the convex combination estimator $\boldsymbol{\theta}_{CC,1/2}$.

7 Conclusion

Applying general large deviation theory for non-regular family (Hayashi 2008), we have calculated the two criteria $\alpha_1(\theta)$ and $\alpha_2(\theta)$ depending on the parameters A_1 and A_2 and the shape parameters κ_1 and κ_2. This analysis essentially depends on the limiting behaviors of the relative Rényi entropies, which were calculated by Hayashi (2002b). As mentioned in Hayashi (2008), the first criterion $\alpha_1(\theta)$ corresponds to the interval estimation, and the second criterion $\alpha_2(\theta)$ corresponds to the point estimation.

In the regular case, the maximum likelihood estimator has the asymptotically best performance. Even in the non-regular case, it has a good performance in the second criterion $\alpha_2(\theta)$. However, it does not always have the asymptotically best performance. The likelihood ratio estimator has good performance in the first criterion $\alpha_1(\theta)$. The estimators based on the extremal statistics have a good performance in both the criteria.

Throughout the obtained result, we have found that the two criteria $\alpha_1(\theta)$ and $\alpha_2(\theta)$ do not coincide when $A_1 \neq A_2$ and $\kappa_1, \kappa_2 < 2$. This fact indicates that we should be careful for the choice of our criterion. This type difference appears in quantum setting (Nagaoka 1994, 1992; Hayashi 2002a). Further, we have found that estimators suitable for the first criteria $\alpha_1(\theta)$ are different from that suitable for the second criteria $\alpha_2(\theta)$. This result suggests that we should choose our estimator depending on the choice of our criterion. Even if the criterion is fixed, the optimal estimator depends on the shape parameter κ_1 and κ_2. Therefore, even in location shift families, it seems difficult to construct an asymptotic optimal estimator universally.

However, we could not calculate $\alpha_1(\theta)$ and $\alpha_2(\theta)$ perfectly. Hence, it is required as a future study to calculate the value that is not calculated.

The extremal statistics are given by one of order statistics while they play an important role in the both criteria. Hence, it is suitable to treat nth order statistic for arbitrary n for improvement. As another research direction, we can discuss the relation with Finsler geometry. Indeed, Amari (1984) pointed out the relation between the geometry of non-regular location shift family and Finsler geometry. Our result seems to indicate that a non-regular family gives a new geometrical structure.

Acknowledgments The author thanks the referees for helpful comments concerning this manuscript.

Appendix A: Lemmas concerning beta and D-psi functions

In this section, we prove some formulas concerning the beta and D-psi functions used in Sect. 5.

Lemma 3 *When* $1 < \kappa < 2$*, we have*

$$\frac{\mathrm{d}}{\mathrm{d}s} s(1-s(\kappa-1))B(s+\kappa(1-s), 2-\kappa)\Big|_{s=1/2} = \frac{(\kappa-1)(3-\kappa)}{4}\pi \tan\frac{2-\kappa}{2}\pi B\left(\frac{1+\kappa}{2}, 2-\kappa\right) > 0 \tag{73}$$

$$\frac{\mathrm{d}}{\mathrm{d}s}(1-s)(1-(1-s)(\kappa-1))B((1-s)+\kappa s, 2-\kappa)\Big|_{s=1/2} = -\frac{(\kappa-1)(3-\kappa)}{4}\pi \tan\frac{2-\kappa}{2}\pi B\left(\frac{1+\kappa}{2}, 2-\kappa\right) < 0 \tag{74}$$

When $0 < \kappa < 1$*, the equations*

$$\frac{\mathrm{d}}{\mathrm{d}s} sB(s+\kappa(1-s), 1-\kappa)\Big|_{s=1/2} = \frac{1-\kappa}{2}\pi \cot \frac{1-\kappa}{2}\pi B\left(\frac{1+\kappa}{2}, 1-\kappa\right) > 0 \tag{75}$$

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}s}(1-s)B((1-s)+\kappa s, 1-\kappa)\Big|_{s=1/2} \\ &\quad = -\frac{1-\kappa}{2}\pi \cot \frac{1-\kappa}{2}\pi B\left(\frac{1+\kappa}{2}, 1-\kappa\right) < 0 \end{aligned} \tag{76}$$

hold.

Proof Using the function $\psi(x) := \frac{\mathrm{d}}{\mathrm{d}x}\log \Gamma(x)$, we can calculate

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}s} s(1-s(\kappa-1))B(s+\kappa(1-s), 2-\kappa)\Big|_{s=1/2} \\ &= \left((2-\kappa) + \frac{(3-\kappa)(1-\kappa)}{4}\left(\psi\left(\frac{1+\kappa}{2}\right) - \psi\left(\frac{5-\kappa}{2}\right)\right)\right) B\left(\frac{1+\kappa}{2}, 2-\kappa\right) \\ &= \left((2-\kappa) + \frac{(3-\kappa)(1-\kappa)}{4}\left(\psi\left(\frac{\kappa-1}{2}\right) + \frac{2}{\kappa-1} - \psi\left(\frac{3-\kappa}{2}\right) - \frac{2}{3-\kappa}\right)\right) \\ &\quad \times B\left(\frac{1+\kappa}{2}, 2-\kappa\right) &(77)\\ &= \left((2-\kappa) + \frac{(3-\kappa)(1-\kappa)}{4}\left(\pi \cot \pi \frac{3-\kappa}{2} + \frac{8-4\kappa}{(\kappa-1)(3-\kappa)}\right)\right) \\ &\quad \times B\left(\frac{1+\kappa}{2}, 2-\kappa\right) &(78)\\ &= \frac{(\kappa-1)(3-\kappa)}{4}\pi \tan \frac{2-\kappa}{2}\pi B\left(\frac{1+\kappa}{2}, 2-\kappa\right) &(79) \end{aligned}$$

where we use the formula $\psi(x+1) = \frac{1}{x} + \psi(x)$ in (77), the formula $\psi(1-x) - \psi(x) = \pi \cot \pi x$ in (78), and the formula $\cot(\frac{\pi}{2} + x) = -\tan x$ in (79). We obtain (73). Similarly, we can prove (74).

Next, we prove (75). We can calculate

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}s} sB(s+\kappa(1-s), 1-\kappa)\Big|_{s=1/2} \\ &= \left(1 + \frac{1-\kappa}{2}\left(\psi\left(\frac{1+\kappa}{2}\right) - \psi\left(\frac{3-\kappa}{2}\right)\right)\right) B\left(\frac{1+\kappa}{2}, 1-\kappa\right) \\ &= \left(1 + \frac{1-\kappa}{2}\left(\psi\left(\frac{1+\kappa}{2}\right) - \psi\left(\frac{3-\kappa}{2}\right) - \frac{2}{1-\kappa}\right)\right) B\left(\frac{1+\kappa}{2}, 1-\kappa\right) &(80)\\ &= \frac{1-\kappa}{2}\pi \cot \frac{1-\kappa}{2}\pi B\left(\frac{1+\kappa}{2}, 1-\kappa\right), &(81) \end{aligned}$$

where (80) follows from the formula $\psi(x+1) = \frac{1}{x} + \psi(x)$ and (81) follows from the formula $\psi(1-x) - \psi(x) = \pi \cot \pi x$. Similarly, we obtain (76). □

Lemma 4 *Assume that* $1 < \kappa < 2$. *There uniquely exists* $t_0 \in (0, \frac{1}{2})$ *satisfying* (63), *and the inequality*

$$(3 - 2\kappa) + (2 - \kappa)(\kappa - 1)(\psi(3 - \kappa) - \psi(1)) \geq 0 \tag{82}$$

holds if, and only if, $1 < \kappa \leq 2 - t_0$.

Proof In the following, this lemma is proven by replacing κ with $2 - t$. Define the function $h(t) := 2t - 1 + t(1-t)(\psi(1+t) - \psi(1))$. When $t < \frac{1}{2}$, we have

$$h'(t) = 2 + (1 - 2t)(\psi(1+t) - \psi(1)) + (t - 2t^2)\psi'(1+t) > 0$$

because $\psi'(x) \geq 0$ for $x > 0$. Therefore, $h(t)$ is strictly monotonically increasing in $(0, \frac{1}{2})$. Since $h(0) = -1 < 0$, $h(\frac{1}{2}) = \frac{1}{4}(\psi(\frac{3}{2}) - \psi(1)) > 0$, there uniquely exists the number $t_0 \in (0, \frac{1}{2})$ satisfying (63). Also, in this case, the inequality $h(t) \geq 0$ holds if $t \geq t_0$.

Next, we consider case $t \geq \frac{1}{2}$. Given the relations,

$$2t - 1 \geq 0, \quad t(1-t) \geq 0, \quad \psi(1+t) - \psi(1) \geq 0,$$

$h(t) \geq 0$. □

Lemma 5 *The minimum*

$$\min_{0 \leq s \leq 1} \left(\frac{(1 - s(\kappa - 1))B(s + \kappa(1-s), 2-\kappa)}{1-s} + \frac{(1 - (1-s)(\kappa-1))B((1-s) + \kappa s, 2 - \kappa)}{s} \right)$$

is attained at $s = \frac{1}{2}$.

Proof Since the minimized function is invariant for the replacement $s \mapsto 1 - s$, it is sufficient to show its concavity. Since $\frac{d^2}{dx^2} e^{h(x)} = (h''(x) + (h('x))^2)e^{h(x)}$, we can show its concavity by proving the concavity of the function $s \mapsto \log \frac{1 - s(\kappa-1)}{1-s} B(s + \kappa(1-s), 2-\kappa)$. We can evaluate

$$\begin{aligned} &\frac{d^2}{ds^2} \log B(s + \kappa(1-s), 2-\kappa) \\ &\quad = (1-\kappa)^2 \left(\psi'(s + \kappa(1-s)) - \psi'(s + \kappa(1-s) + 2 - \kappa)\right) > 0 \end{aligned}$$

because $\psi'(x)$ is monotonically decreasing for $x > 0$. Also, we have

$$\frac{d^2}{ds^2} \log \frac{1 - s(\kappa-1)}{1-s} = \frac{(2-\kappa)(\kappa - 2(\kappa-1)s)}{(1 + \kappa s + (\kappa-1)s^2)^2} > 0$$

because $\kappa > 2(\kappa - 1)$. The proof is now complete. □

Appendix B: A lemma used in Sect. 5.3

Lemma 6 *For any $\delta > 0$, we have*

$$\lim_{\epsilon\to 0} \frac{1}{\epsilon^2 \log \epsilon} \int_0^\delta \exp\left(-\epsilon \frac{1}{x}\right)(x+\epsilon) - x \mathrm{d}x = \frac{1}{2} \tag{83}$$

$$\lim_{\epsilon\to 0} \frac{1}{\epsilon^2 \log \epsilon} \int_\epsilon^\delta \exp\left(+\epsilon \frac{1}{x}\right)(x-\epsilon) - x \mathrm{d}x = \frac{1}{2}. \tag{84}$$

Proof we can calculate

$$\begin{aligned}\int_\epsilon^\delta \exp\left(+\epsilon \frac{1}{x}\right)(x-\epsilon) - x \mathrm{d}x &= \int_\epsilon^\delta \sum_{n=1}^\infty - \frac{n\epsilon^{n+1}}{(n+1)!x^n} \mathrm{d}x \\ &= -\frac{\epsilon^2}{2}(\log\delta - \log\epsilon) \\ &\quad + \sum_{n=2}^\infty \frac{n(n-1)}{(n+1)!}\left(\frac{\epsilon^{n+1}}{\delta^{n-1}} - \epsilon^2\right).\end{aligned}$$

Since $\frac{\epsilon^2}{\epsilon^2 \log \epsilon} \to 0$, we obtain (84). Similarly, we can calculate

$$\begin{aligned}\int_\epsilon^\delta \exp\left(-\epsilon \frac{1}{x}\right)(x+\epsilon) - x \mathrm{d}x &= -\frac{\epsilon^2}{2}(\log\delta - \log\epsilon) \\ &\quad + \sum_{n=2}^\infty \frac{n(n-1)(-1)^n}{(n+1)!}\left(\frac{\epsilon^{n+1}}{\delta^{n-1}} - \epsilon^2\right).\end{aligned}$$

Since

$$\begin{aligned}0 \le \int_0^\epsilon \exp\left(-\epsilon \frac{1}{x}\right)(x+\epsilon)\mathrm{d}x &\le \int_0^\epsilon (x+\epsilon)\mathrm{d}x = \frac{3}{2}\epsilon^2 \\ \int_0^\epsilon x \mathrm{d}x &= \frac{1}{2}\epsilon^2,\end{aligned}$$

we obtain (83). □

References

Akahira, M. (1996). Loss of information of a statistic for a family of non-regular distributions, *Annals of Institute of Statistical Mathematics, 48(2)*, 349–364.

Akahira, M., Takeuchi, K. (1995). Non-regular statistical estimation. *Lecture Notes in Statistics* (Vol. 107). New York: Springer.

Amari, S. (1984). Non-regular probability family and Finsler geometry (in Japanese), *RIMS koukyuroku, 6*, 27.

Bahadur, R. R. (1960). Asymptotic efficiency of tests and estimates. *Sankhyā, 22*, 229–252.

Bahadur, R. R. (1967). Rates of convergence of estimates and test statistics. *Annals of Mathematical Statistics, 38*, 303–324

Bahadur, R. R. (1971). Some limit theorems in statistics. In *Regional conference series in applied mathematics* (Vol. 4). Philadelphia: SIAM.

Fu, J. C. (1973). On a theorem of Bahadur on the rate of convergence of point estimator. *Annals of Statistics, 1*, 745–749.

Hayashi, M. (2002a). Two quantum analogues of Fisher information from a large deviation viewpoint of quantum estimation. *Journal of Physics. A: Mathematical and General, 35*, 7689–7727. Also appeared as Chap. 28 of *Asymptotic theory of quantum statistical inference,* M. Hayashi (ed.), 2005.

Hayashi, M. (2002b). Limiting behavior of relative Rényi entropy in a non-regular location shift family. *Annals of Institute of Statistical Mathematics*. Latest version: http://arxiv.org/abs/math/0212077. doi:10.1007/s10463-008-0182-4.

Hayashi, M. (2008). Two non-regular extensions of large deviation bound. *Special Issue—Recent Advances in Statistical Inference. Communications in Statistics: Theory and Methods* (accepted). Latest version: http://arxiv.org/abs/math/0604197.

Ibragimov, I. A., Has'minskii, R. Z. (1981). *Statistical estimation*. New York: Springer.

Nagaoka, H. (1992). On the relation between Kullback divergence and Fisher information—from classical systems to quantum systems. In *Proceedings of Joint Mini-workshop for data compression theory and fundamental open problems in information theory* (pp. 63–72). Originally written in Japanese. Also appeared as Chap. 27 of *Asymptotic theory of quantum statistical inference*, M. Hayashi (ed.), 2005.

Nagaoka, H. (1994). Two quantum analogues of the large deviation Cramér-Rao inequality. In *Proceedings of 1994 IEEE International Symposium on Information Theory* (p. 118).