Statistics & Decisions 19, 1–37 (2001)

Minimax estimators of the coverage probability of the impermissible error for a location family

Miguel A. Arcones

Received: Month-1 99, 2003; Accepted: Month-2 99, 2004

Summary: We consider estimation for a multivariate location family. Between all confidence regions with volume less than a fixed value L, we find the equivariant confidence region with the biggest coverage probability. This region maximizes the infimum of the coverage probability over all confidence regions with volume less than L. As an application, we find an estimator of a location parameter with the property that minimizes the supremum of the probability that the error of the estimation exceeds a fixed constant. We also find a confidence region and an estimator having the previous properties, but based on the maximum likelihood estimator. In the one dimensional case, we find the Bahadur slope of the two obtained estimators. We show that except for certain families of distributions, the estimator based on the whole sample is superior to the estimator based upon the m.l.e. Hence, we get that m.l.e.'s are not asymptotically sufficient.

February 6, 2008

1 Introduction

We consider the estimation of the parameter θ indexing a family $\{f(\cdot,\theta) : \theta \in \Theta\}$ of p.d.f.'s, where Θ is a Borel subset of $\mathbb{R}^d$. Let $\{X_j\}_{j=1}^{\infty}$ be a sequence of i.i.d.r.v.'s with a p.d.f. belonging to $\{f(\cdot,\theta) : \theta \in \Theta\}$. Let $T_n := T_n(X_1,\dots,X_n)$ be an estimator of θ. The error of the estimator T_n is $|T_n-\theta|$, where $|\cdot|$ is the Euclidean norm. Suppose that we select a number b, $b>0$, such that any error of estimation bigger than b is impermissible. The maximum coverage probability of the impermissible error of the estimator T_n is

$$\sup_{\theta\in\Theta} \mathbb{P}_\theta\{|T_n-\theta|>b\}, \tag{1.1}$$

where $\mathbb{P}_\theta$ is the probability measure when the r.v.'s $X_1,\dots,X_n$ have p.d.f. $f(\cdot,\theta)$. In this paper, we study an estimator $T_{HL,b}$ minimizing (1.1) for a location family. The estimator $T_{HL,b}$ may depend on b.

AMS 2000 subject classification: Primary 62F12; secondary, 62F25, 62C20.
Key words and phrases: error of estimation, impermissible error, confidence interval, location family, minimax, translation equivariant.

Coverage probability of an impermissible error appears in the literature as a way to assess estimators. In the one dimensional case, Bahadur [4, 5, 6] proved that, if T_n is a consistent estimator of θ, for each $\theta \in \Theta$, then, for each $\theta \in \Theta$,

$$\lim_{b \to 0} \liminf_{n \to \infty} b^{-2} n^{-1} \log \left(\mathbb{P}_\theta \{ |T_n - \theta| \geq b \} \right) \geq -2^{-1} v(\theta), \tag{1.2}$$

where $v(\theta)$ is the Fisher information at θ, i.e.

$$v(\theta) := E_\theta \left[\left(\frac{\partial \log f(X, \theta)}{\partial \theta} \right)^2 \right] = -E_\theta \left[\frac{\partial^2 \log f(X, \theta)}{\partial \theta^2} \right]. \tag{1.3}$$

Bahadur also proved, under regularity conditions, that for each θ

$$\lim_{\epsilon \to 0} \liminf_{n \to \infty} \epsilon^{-2} n^{-1} \log \left(\mathbb{P}_\theta \{ |\hat{\theta}_n - \theta| \geq \epsilon \} \right) = -2^{-1} v(\theta), \tag{1.4}$$

where $\hat{\theta}_n$ is a (m.l.e.) maximum likelihood estimator of θ.

Bahadur [6] (see also Bahadur, Zabell and Gupta [7]) showed that if T_n is a consistent estimator of θ, for each $\theta \in \Theta$, then, for each $\theta \in \Theta$,

$$-\liminf_{n \to \infty} n^{-1} \log \left(\mathbb{P}_\theta \{ |T_n - \theta| > b \} \right) \leq B(b) \tag{1.5}$$

where

$$B(b) := \inf\{ K(f(\cdot, \theta_1), f(\cdot, \theta)) : \theta_1 \text{ satisfying } |\theta_1 - \theta| > b \} \tag{1.6}$$

and K is the Kullback–Leibler information of the densities $f(\cdot, \theta_1)$ and $f(\cdot, \theta)$. Given densities f and g with respect to a measure μ,

$$K(f, g) = \int \log(f(t)/g(t)) f(t)\, d\mu(t).$$

The limit

$$-\liminf_{n \to \infty} n^{-1} \log \left(\mathbb{P}_\theta \{ |T_n - \theta| > b \} \right) \tag{1.7}$$

is called the inaccuracy rate of the estimator T_n. Several authors have studied the inaccuracy rates of estimators. Fu [12, 14] gave conditions in order that sequences of estimators satisfy (1.4). Fu [13] showed that the inaccuracy rate of an estimator is related with the asymptotic behavior of its density (if it has one).

We consider location families, i.e. families of the form $\{ f(\cdot - \theta) : \theta \in \mathbb{R}^d \}$. For one dimensional location families, we find a translation equivariant estimator $T_{n,HL,b}$ such that for any other estimator T_n,

$$\sup_{\theta \in \mathbb{R}} \mathbb{P}_\theta \{ |T_n - \theta| > b \} \geq \sup_{\theta \in \mathbb{R}} \mathbb{P}_\theta \{ |T_{n,HL,b} - \theta| > b \}, \tag{1.8}$$

(see Section 2). Notice that $\mathbb{P}_\theta \{ |T_{n,HL,b} - \theta| > b \}$ does not depend on θ. $T_{n,HL,b}$ is the middle point of the highest likelihood region of length $2b$ (assuming that this region is an interval). $T_{n,HL,b}$ is also the best equivariant estimator for the considered problem.

It is known (see Girshick and Savage [15]) that in certain situations the best equivariant estimator is minimax. We show that the inaccuracy rate of $T_{n,HL,b}$ is

$$S(2b) = -\lim_{n\to\infty} n^{-1}\log\left(\mathbb{P}_\theta\{|T_{n,HL,b}-\theta|>b\}\right), \tag{1.9}$$

where

$$S(t) := -\inf_{0<\lambda<1}\log\int_{\mathbb{R}}(f(x-t))^{\lambda}(f(x))^{1-\lambda}\,dx, t\in\mathbb{R}.$$

The function $S(\cdot)$ appeared in Chernoff [8], as the large deviations of the sum of the two errors in a simple hypothesis testing problem. $S(\cdot)$ is a measure of the information in a sample. The function $S(\cdot)$ also appeared in Sievers [23] as the limit of

$$\lim_{n\to\infty} n^{-1}\log\left(\max\left(\mathbb{P}_\theta\{T_n^{(S)}-\theta>b\},(\mathbb{P}_\theta\{T_n^{(S)}-\theta<-b\}\right)\right), \tag{1.10}$$

where $T_n^{(S)}$ is the translation equivariant estimator minimizing

$$\max\left(\mathbb{P}_\theta\{T_n-\theta>b\},\mathbb{P}_\theta\{T_n-\theta<-b\}\right),$$

over all translation equivariant estimators.

$T_{HL,b}$ determines a confidence interval of length $2b$. Finding $T_{HL,b}$ is equivalent to find a confidence interval of length $2b$ with coverage probability as large as possible. A confidence region $C(X_1,\dots,X_n)$ is determined by a map from $(\mathbb{R}^d)^n$ to $\mathcal{B}(\mathbb{R}^d)$, where $\mathcal{B}(\mathbb{R}^d)$ is the Borel σ–field of $\mathbb{R}^d$. Suppose that we consider confidence regions $C(X_1,\dots,X_n)$ of Lebesgue measure less or equal than L. Let $C_{HL,L}(X_1,\dots,X_n)$ be the highest likelihood region having Lebesgue measure L. For a location family, $C_{HL,L}(X_1,\dots,X_n)$ satisfies that for any confidence region $C(X_1,\dots,X_n)$ of Lebesgue measure less or equal than L,

$$\sup_{\theta\in\mathbb{R}^d}\mathbb{P}_\theta\{\theta\notin C(X_1,\dots,X_n)\}\ge\sup_{\theta\in\mathbb{R}^d}\mathbb{P}_\theta\{\theta\notin C_{HL,L}(X_1,\dots,X_n)\}, \tag{1.11}$$

(see Theorem 2.2). $C_{HL,L}(X_1,\dots,X_n)$ is the best translation equivariant region. If $d=1$ and $C_{HL,2b}(X_1,\dots,X_n)$ is an interval, then $T_{n,HL,b}(X_1,\dots,X_n)$ is the middle point of $C_{HL,2b}(X_1,\dots,X_n)$.

As to regular confidence regions, the results in Kudō [19] (see also Valand [25]; and Joshi [17]) show that for a location family the level $1-\alpha$ confidence interval obtained by taking the highest density region of the distribution of parameter given the sample, when the parameter has a "uniform" distribution, minimizes the maximum expected length of a confidence interval over all the level $1-\alpha$ confidence intervals. These confidence intervals are different from the ones obtained. In general, these confidence intervals have a random length.

Since m.l.e.'s are asymptotically optimal in some situations, it is of interest knowing whether they are optimal for the considered problem. We obtain the confidence regions $C_{piv,L}(\hat\theta_n)$ based on the m.l.e. $\hat\theta_n$ which maximize the coverage probability of the parameter between all confidence regions of volume less or equal than a constant. These confidence regions $C_{piv,L}(\hat\theta_n)$ are obtained taking the highest density region of the pivot

$\theta - \hat{\theta}_n$ (see Theorem 2.3). In Section 2, we show that for any confidence region $C(\hat{\theta}_n)$ based on the m.l.e. of expected Lebesgue measure less or equal than L,

$$\sup_{\theta \in \mathbb{R}^d} \mathbb{P}_\theta\{\theta \notin C(\hat{\theta}_n)\} \geq \sup_{\theta \in \mathbb{R}^d} \mathbb{P}_\theta\{\theta \notin C_{piv,L}(\hat{\theta}_n)\}. \tag{1.12}$$

$C_{piv,L}(\hat{\theta}_n)$ is a translation equivariant confidence region. We find

$$A(L) := -\lim_{n \to \infty} n^{-1} \log \mathbb{P}_\theta\{\theta \notin C_{piv,L}(\hat{\theta}_n)\}. \tag{1.13}$$

$C_{piv,L}(\hat{\theta}_n)$ can be written as $\hat{\theta}_n + C_{piv,2b}$, where $C_{piv,2b}$ is a nonrandom set. The estimator $T_{n,piv,b}$ based on the m.l.e. which minimizes (1.1) between all estimators based on the m.l.e. is the middle point of $\hat{\theta}_n + C_{piv,2b}$.

In general, the confidence regions $C_{HL,L}(X_1, \ldots, X_n)$ and $C_{piv,L}(\hat{\theta}_n)$ are different. We have that $C_{HL,L}(X_1, \ldots, X_n)$ and $C_{piv,L}(\hat{\theta}_n)$ agree for the location families which are members of an exponential family. But, when $d = 1$, they also do for a location family which is a two dimensional curved exponential family (see Example 3.14). We show that for symmetric one dimensional distributions, not of the form above, then for some $b > 0$, $A(2b) < S(2b)$. This means the m.l.e. is not asymptotically sufficient for the considered problem.

$T_{n,HL}$ is not necessarily a consistent estimator of θ. In Theorem 3.5 we obtain the limit in probability of $T_{n,HL,b}$. $T_{n,HL,b}$ is consistent for each $b > 0$ if and only if the Kullback–Leibler information is symmetric.

We present examples, where $S(2b) < B(b)$, $S(2b) = B(b)$ and $S(2b) > B(b)$. This seems to indicate that the bound $B(\cdot)$ in (1.5) is not optimal. In Example 3.15, we have that $T_{n,HL}$ is consistent, but $S(2b) < B(b)$.

The organization of the paper is as follows. Section 2 contains the results about minimax confidence regions. Section 3 deals with the large deviations of the considered confidence regions and estimators. Section 4 contains the proofs.

2 Maximin confidence regions of a fixed Lebesgue measure.

In this section, we find a confidence region of Lebesgue measure less or equal than a constant L which maximizes the infimum of the coverage probability of the parameter. We will use the following lemma:

Lemma 2.1 *Let $f, g : \mathbb{R}^d \to [0, \infty)$ be two measurable functions. Let $L > 0$.*

(i) Suppose that $C \in \mathcal{B}(\mathbb{R}^d)$ satisfies that $\int_C f(x)\,dx = L$ and that there exists $\infty \geq \lambda \geq 0$ such that $\{x : \lambda f(x) < g(x)\} \subset C \subset \{x : \lambda f(x) \leq g(x)\}$. Then, for each $B \in \mathcal{B}(\mathbb{R}^d)$ such that $\int_B f(x)\,dx \leq L$,

$$\int_B g(x)\,dx \leq \int_C g(x)\,dx.$$

(ii) Suppose that $L < \int f(x)\,dx$ *and* $\int g(x)\,dx < \infty$. *Let*

$$\lambda := \inf\{t \geq 0 : \int_{x \in \mathbb{R}^d : tf(x) \leq g(x)} f(x)\,dx \leq L\}.$$

Then, $0 \leq \lambda < \infty$ *and*

$$\int_{x \in \mathbb{R}^d : \lambda f(x) < g(x)} f(x)\,dx \leq L \leq \int_{x \in \mathbb{R}^d : \lambda f(x) \leq g(x)} f(x)\,dx.$$

Hence, there exists a set $C \in \mathcal{B}(\mathbb{R}^d)$ *such that* $\int_C f(x)\,dx = L$ *and* $\{x \in \mathbb{R}^d : \lambda f(x) < g(x)\} \subset C \subset \{x \in \mathbb{R}^d : \lambda f(x) \leq g(x)\}$.

The previous lemma is a variation on the Neyman–Pearson lemma (see also Guenther [16] and Juola [18]). Guenther [16] and Juola [18] use a variation of the previous lemma to construct the shortest $1-\alpha$ level confidence interval for a one dimensional parameter based on a pivotal quantity.

First, we find between all translation equivariant confidence regions for θ with Lebesgue measure less or equal than L, the one maximizing

$$\inf_{\theta \in \mathbb{R}^d} \mathbb{P}_\theta\{\theta \in C(X_1, \ldots, X_n)\}. \tag{2.1}$$

We use the Pitman [20] transformation (see also Shao [22, Section 4.2]). Let $\vec{Z} := (Z_1, \ldots, Z_{n-1}) = (X_1 - X_n, \ldots, X_{n-1} - X_n)$. Since $(X_1, \ldots, X_n)$ has p.d.f. $\prod_{j=1}^n f(x_j - \theta)$, $(X_n, \vec{Z})$ has p.d.f. $\prod_{j=1}^{n-1} f(z_j + x_n - \theta) \times f(x_n - \theta)$. We denote $f_{X_n,\vec{Z}}(x_n, \vec{z}) := \prod_{j=1}^{n-1} f(z_j + x_n) \times f(x_n)$. Then, $(X_n, \vec{Z})$ has p.d.f. $f_{X_n,\vec{Z}}(x_n - \theta, \vec{z})$. The marginal p.d.f. of $\vec{Z}$ is

$$f_{\vec{Z}}(\vec{z}) = \int_{\mathbb{R}^d} \prod_{j=1}^{n-1} f(z_j + y) \times f(y)\,dy. \tag{2.2}$$

Notice that $\vec{Z}$ is an ancillary statistic. The conditional pdf of X_n given $\vec{Z}$ is

$$f_{X_n|\vec{Z}}(x_n - \theta|\vec{z}) = (f_{\vec{Z}}(\vec{z}))^{-1} f_{X_n,\vec{Z}}(x_n - \theta, \vec{z}). \tag{2.3}$$

If $C(X_1, \ldots, X_n)$ is a translation equivariant confidence region, then $C(X_1, \ldots, X_n) = X_n + C(Z_1, \ldots, Z_{n-1}, 0)$. Let $C^*(\vec{Z}) = C(Z_1, \ldots, Z_{n-1}, 0)$. We would like to find $C^*(\vec{Z})$ maximizing

$$\begin{aligned} &\inf_{\theta \in \mathbb{R}^d} \mathbb{P}_\theta\{\theta \in X_n + C^*(\vec{Z})\} = \mathbb{P}_0\{-X_n \in C^*(\vec{Z})\} \qquad (2.4)\\ &= \int\int I(-x_n \in C^*(\vec{z})) \textstyle\prod_{j=1}^{n-1} f(z_j + x_n) \times f(x_n)\,dx_n\,d\vec{z} \\ &= \int\int I(y \in C^*(\vec{z})) \textstyle\prod_{j=1}^{n-1} f(z_j - y) \times f(-y)\,dy\,d\vec{z}. \end{aligned}$$

subject to

$$\int_{\mathbb{R}^m} I(y \in x_n + C^*(\vec{z}))\,dy = \int_{\mathbb{R}^m} I(y \in C^*(\vec{z}))\,dy \leq L, \tag{2.5}$$

for each $\vec{z}$. We apply Lemma 2.1 to $f \equiv 1$ and $g \equiv \prod_{j=1}^{n-1} f(z_j - y) \times f(-y)$ conditionally on $\vec{z}$. Take $C^*_{HL}(\vec{z})$ such that for each $\vec{z}$, $\int I(y \in C^*_{HL}(\vec{z}))\, dy = L$, and

$$\begin{aligned} & \{y \in \mathbb{R}^d : \lambda_{HL}(\vec{z}) < \textstyle\prod_{j=1}^{n-1} f(z_j - y) \times f(-y)\} \subset C^*_{HL}(\vec{z}) \\ \subset \quad & \{y \in \mathbb{R}^d : \lambda_{HL}(\vec{z}) \le \textstyle\prod_{j=1}^{n-1} f(z_j - y) \times f(-y))\}. \end{aligned} \tag{2.6}$$

where

$$\lambda_{HL,L}(\vec{z}) = \inf\{t \ge 0 : \int_{t \le \prod_{j=1}^{n-1} f(z_j - y) \times f(-y)} dy \le L\}. \tag{2.7}$$

Then, $C_{HL,L}(X_1, \dots, X_n) := X_n + C^*_{HL,L}(X_1 - X_n, \dots, X_1 - X_n)$ maximizes (2.1) between all possible translation equivariant confidence regions of Lebesgue measure less or equal than L.

It follows from (2.6) that

$$\begin{aligned} & \{\theta \in \mathbb{R}^d : \lambda_{HL}(\vec{z}) < \textstyle\prod_{j=1}^{n} f(x_j - \theta)\} \subset C_{HL,L}(X_1, \dots, X_n) \\ \subset \quad & \{\theta \in \mathbb{R}^d : \lambda_{HL}(\vec{z}) \le \textstyle\prod_{j=1}^{n} f(x_j - \theta)\}. \end{aligned} \tag{2.8}$$

Next theorem gives sufficient conditions so that $C_{HL,L}(X_1, \dots, X_n)$ maximizes (2.1) between all possible confidence regions of Lebesgue measure less or equal than L.

Theorem 2.2 *Suppose that:*
(i) For each $0 < M < \infty$, $\inf\{f(x) : |x| \le M\} > 0$.
(ii) $\lim_{|x| \to \infty} f(x) = 0$.
(iii) $\sup_{x \in \mathbb{R}^d} f(x) < \infty$.
Let $C(X_1, \dots, X_n)$ *be a confidence region for* θ *such that for each* $x_1, \dots, x_n$, $\int_{\mathbb{R}^d} I(y \in C(x_1, \dots, x_n))\, dy \le L$. *Then,*

$$\inf_{\theta \in \mathbb{R}^d} \mathbb{P}_\theta\{\theta \in C(X_1, \dots, X_n)\} \le \inf_{\theta \in \mathbb{R}^d} \mathbb{P}_\theta\{\theta \in C_{HL,L}(X_1, \dots, X_n)\}. \tag{2.9}$$

Observe that the confidence region $C_{HL,L}(X_1, \dots, X_n)$ is not necessarily convex. In the one dimensional situation, $C_{HL,L}(X_1, \dots, X_n)$ is not necessarily an interval.

Next, we consider the maximim confidence region based upon an mle. An mle $\hat{\theta}_n$ of θ, over a parametric family $\{f(x, \theta) : \theta \in \Theta\}$ is a r.v. such that

$$n^{-1} \sum_{j=1}^{n} \log f(X_j, \hat{\theta}_n) = \sup_{\theta \in \Theta} n^{-1} \sum_{j=1}^{n} \log f(X_j, \theta). \tag{2.10}$$

Suppose that there exists a r.v. $\hat{\theta}_n = \hat{\theta}_n(X_1, \dots, X_n)$ satisfying (2.10) when $f(x, \theta) = f(x - \theta)$. We may assume that $\hat{\theta}_n$ is equivariant. This implies that $\hat{\theta}_n - \theta$ is a pivotal quantity, i.e. the distribution of $\hat{\theta}_n - \theta$ does not depend on θ. Suppose that $\theta - \hat{\theta}_n$ has a distribution absolutely continuous with respect to the Lebesgue measure, when θ obtains. Let $h_n(\cdot)$ be the pdf of $\theta - \hat{\theta}_n$ when θ obtains. Let $L > 0$. Given $C \in \mathcal{B}(\mathbb{R}^d)$, a confidence region for θ is determined by $\{\theta \in \mathbb{R}^d : \theta - \hat{\theta}_n \in C\} = \{\theta \in \mathbb{R}^d : \theta \in \hat{\theta}_n + C\}$. Between all sets C such that

$$\int I(\theta \in \hat{\theta}_n + C)\, d\theta = \int_C 1\, d\theta \leq L, \tag{2.11}$$

a set C which maximizes

$$\mathbb{P}_\theta\{\theta - \hat{\theta}_n \in C\} = \int_C h_n(\theta)\, d\theta \tag{2.12}$$

is obtained by applying Lemma 2.1 to $f \equiv 1$ and $g \equiv h_n$. Let

$$\lambda_{piv,L} := \inf\{t \geq 0 : \int_{x \in \mathbb{R}^d : t \leq h_n(x)} 1\, dx \leq L\}. \tag{2.13}$$

Take $C_{piv,L} \subset \mathbb{R}^d$ such that

$$\int_{C_{piv,L}} 1\, dx = L \tag{2.14}$$

and

$$\{x \in \mathbb{R}^d : \lambda_{piv,L} < h_n(x)\} \subset C_{piv,L} \subset \{x \in \mathbb{R}^d : \lambda_{piv,L} \leq h_n(x)\}. \tag{2.15}$$

Hence, using the mle, a confidence region satisfying (2.11) and maximizing (2.12) is

$$\hat{\theta}_n + C_{piv,L}, \tag{2.16}$$

where $C_{piv,L}$ satisfies (2.14) and (2.15). The obtained confidence region $\hat{\theta}_n + C_{piv,L}$ is translation equivariant. From (2.15), we get that

$$\{\theta \in \mathbb{R}^d : \lambda_{piv,L} < f_n(\hat{\theta}_n, \theta)\} \subset \hat{\theta}_n + C_{piv,L} \subset \{\theta \in \mathbb{R}^d : \lambda_{piv,L} \leq f_n(\hat{\theta}_n, \theta)\}, \tag{2.17}$$

were $f_n(x, \theta) = h_n(\theta - x)$, $x \in \mathbb{R}$, is the p.d.f. of $\hat{\theta}_n$ when θ obtains.

In this situation, the obtained confidence regions maximize the coverage probability over all the regions with expected Lebesgue measure less or equal than L:

Theorem 2.3 *Let $C(\hat{\theta}_n)$ be a confidence region for θ such that*

$$\sup_{\theta \in \mathbb{R}^d} E_\theta \left[\int_{\mathbb{R}^d} I(y \in C(\hat{\theta}_n))\, dy \right] \leq L. \tag{2.18}$$

Then,

$$\inf_{\theta \in \mathbb{R}^d} \mathbb{P}_\theta\{\theta \in C(\hat{\theta}_n)\} \leq \inf_{\theta \in \mathbb{R}^d} \mathbb{P}_\theta\{\theta \in \hat{\theta}_n + C_{piv,L}\}. \tag{2.19}$$

The previous theorem implies that if $C(\hat{\theta}_n)$ is a confidence region for θ such that for each $x \in \mathbb{R}^d$,

$$\int_{\mathbb{R}^d} I(y \in C(x))\, dy \leq L, \tag{2.20}$$

then, (2.19) holds.

Theorem 2.2 does not hold using the expected value of the Lebesgue measure of the confidence region. In general, the confidence regions which are maximin when the expected value of the Lebesgue measure is used do not have constant Lebesgue measure.

3 Inaccuracy rates of minimax estimators.

Before presenting the results in this section, we recall some notation on the large deviation principle (LPD). General references on the LDP are Deuschel and Stroock [10] and Dembo and Zeitouni [9]. Let M be a metric space. Let $\mathcal{B}(M)$ be the Borel σ–field of M. We say that a sequence of r.v.'s. $\{U_n\}_{n=1}^{\infty}$ with values in M is said to follow the LDP with speed ϵ_n^{-1}, where $\{\epsilon_n\}$ is a sequence of positive numbers converging to zero, and with good rate function I if:

(i) For each $0 \le c < \infty$, $\{z \in M : I(z) \le c\}$ is a compact set of S.

(ii) For each set $A \in \mathcal{B}(M)$,

$$\begin{aligned} -\inf\{I(z) : z \in A^o\} &\le \liminf_{n\to\infty} \epsilon_n \log(\Pr\{U_n \in A\}) \\ \le \limsup_{n\to\infty} \epsilon_n \log(\Pr\{U_n \in A\}) &\le -\inf\{I(z) : z \in \bar{A}\}, \end{aligned}$$

where A^o (resp. $\bar{A}$) denotes the interior (resp. closure) of A.

A function $\Phi : \mathbb{R} \to \bar{\mathbb{R}}$ is said to be a Young function if it is convex, $\Phi(0) = 0$; $\Phi(x) = \Phi(-x)$ for each $x > 0$; and $\lim_{x\to\infty} \Phi(x) = \infty$. Let X be a r.v. with values in a measurable space $(S, \mathcal{S})$. The Orlicz space $\mathcal{L}^{\Phi}(S, \mathcal{S})$ (abbreviated to $\mathcal{L}^{\Phi}$) associated with the Young function Φ is the class of measurable functions $f : (S, \mathcal{S}) \to \mathbb{R}$ such that $E[\Phi(\lambda f(X))] < \infty$ for some $\lambda > 0$. The Minkowski (or gauge) norm of the Orlicz space $\mathcal{L}^{\Phi}(S, \mathcal{S})$ by

$$N_\Phi(f) := \inf\{t > 0 : E[\Phi(f(X)/t)] \le 1\}.$$

It is well known that the vector space $\mathcal{L}^{\Phi}$ with the norm N_Φ is a Banach space. Define

$$\mathcal{L}^{\Phi_1} := \{f : S \to \mathbb{R} : E[\Phi_1(\lambda|f(X)|)] < \infty \text{ for some } \lambda > 0\},$$

where $\Phi_1(x) = e^{|x|} - |x| - 1$. Let $(\mathcal{L}^{\Phi_1})^*$ be the dual of $(\mathcal{L}^{\Phi_1}, N_{\Phi_1})$. The function $f \in \mathcal{L}^{\Phi_1} \mapsto \log\left(E[e^{f(X)}]\right) \in \mathbb{R}$ is a convex lower semicontinuous function. The Fenchel–Legendre conjugate of the previous function is:

$$J(l) := \sup_{f\in\mathcal{L}^{\Phi_1}} \left(l(f) - \log\left(E[e^{f(X)}]\right)\right),\ l \in (\mathcal{L}^{\Phi_1})^*. \tag{3.1}$$

J is a function with values in $[0, \infty]$. Since J is a Fenchel–Legendre conjugate, it is a nonnegative convex lower semicontinuous function. It is easy to see that if $J(l) < \infty$, then:

(i) $l(\mathbf{1}) = 1$, where $\mathbf{1}$ denotes the function constantly 1.

(ii) l is a nonnegative definite functional: if $f(X) \ge 0$ a.s., then $l(f) \ge 0$.

Given a nonnegative function γ on S such that $E[\gamma(X)] = 1$ and $E[\Psi_2(\gamma(X))] < \infty$, where $\Psi_2(x) = x \log x$, then $l_\gamma(f) = E[f(X)\gamma(X)]$ defines a continuous linear functional in $\mathcal{L}^{\Phi_1}$. Besides, it is easy to see that

$$J(l_\gamma) = \sup_{f\in\mathcal{L}^{\Phi_1}} E[f(X)\gamma(X) - \Phi_2(f(X))] = E[\Psi_2(\gamma(X))]. \tag{3.2}$$

The previous function J can be used to determine the rate function in the large deviation of statistics. Let $\{X_j\}_{j=1}^{\infty}$ be a sequence of i.i.d.r.v.'s with the distribution of X. If

$f \in \mathcal{L}^{\Phi_1}$, then $\{n^{-1}\sum_{j=1}^n f(X_j)\}$ satisfies the LDP with rate function

$$I_f(t) := \sup_{\lambda\in\mathbb{R}} \left(\lambda t - \log\left(E[\exp(\lambda f(X))]\right)\right), t \in \mathbb{R} \tag{3.3}$$

(see for example Dembo and Zeitouni [9, Theorem 2.2.3]). By Arcones [3, Lemma 2.2],

$$I_f(t) = \inf\left\{J(l) : l \in (\mathcal{L}^{\Phi_1})^*, l(f) = t\right\}. \tag{3.4}$$

It is well known that $I_f(\mu_f) = 0$, where $\mu_f = E[f(X)]$, I_f is convex, I_f is nondecreasing in $[\mu_f, \infty)$ and I is nonincreasing in $(-\infty, \mu_f]$ (see e.g. Dembo and Zeitouni [9, Lemma 2.2.5]). In particular, if $t \geq \mu_f$,

$$\inf\{J(l) : l \in (\mathcal{L}^{\Phi_1})^*, l(f) \geq t\} = I_f(t) \tag{3.5}$$

and for each $t \leq \mu_f$,

$$\inf\{J(l) : l \in (\mathcal{L}^{\Phi_1})^*, l(f) \leq t\} = I_f(t)$$

(see for example Dembo and Zeitouni [9, Corollary 2.2.19]). Chernoff [8, page 496] shows that if $\mathbb{P}\{f(X) = t) < 1$ and $t < E[f(X)]$, then

$$I_f(t) = \sup_{\lambda<0}(\lambda t - \log E[\exp(\lambda f(X))]) \tag{3.6}$$

Chernoff [8, page 495] shows that: (i) if $\mathbb{P}\{f(X) > t) > 0$ and $\mathbb{P}\{f(X) < t) > 0$, then there exists a unique finite λ_t such that

$$I_f(t) = \lambda_t t - \log E[\exp(\lambda_t f(X))]; \tag{3.7}$$

(ii) if $\mathbb{P}\{f(X) \geq t) > 0$ and $\mathbb{P}\{f(X) \leq t) > 0$, then

$$I_f(t) < \infty. \tag{3.8}$$

We will also use that for each $k \geq 0$ and each function $f \in \mathcal{L}^{\Phi_1}$ and each

$$|l(f)| \leq (J(l) + 1 + 2^{1/2})N_{\Phi_1}(f), \tag{3.9}$$

(see Arcones [3, Lemma 5.1]).

When θ obtains, we denote $\mathcal{L}^{\Phi_1}$, $(\mathcal{L}^{\Phi_1})^*$ and J by $\mathcal{L}_\theta^{\Phi_1}$, $(\mathcal{L}_\theta^{\Phi_1})^*$ and J_θ, respectively.

In Arcones [2] the large deviations of the confidence regions of Lebesgue measure less or equal than a constant were studied for a general family of p.d.f.'s. Let $\{X_j\}_{j=1}^\infty$ be a sequence of i.i.d.r.v.'s from the parametric family $\{f(\cdot, \theta) : \theta \in \Theta\}$, where Θ is a Borel set of $\mathbb{R}^d$. Let

$$\lambda_{HL,L}(X_1, \ldots, X_n) := \inf\{t \geq 0 : \int_{t\leq G_n(\theta)} d\,\theta \leq L\},$$

where $G_n(t) = n^{-1}\sum_{j=1}^n \log f(X_j, t)$. Take a set $C_{HL,L}(X_1, \ldots, X_n)$ such that

$$\int_{\theta\in C_{HL,L}(X_1,\ldots,X_n)} d\theta = L$$

and

$$\{\theta \in \Theta : \lambda_{HL,L}(X_1,\dots,X_n) < G_n(\theta)\} \subset C_{HL,L}(X_1,\dots,X_n)$$
$$\subset \{\theta \in \Theta : \lambda_{HL,L}(X_1,\dots,X_n) \le G_n(\theta)\}.$$

We will need the following variation of Arcones [2, Theorem 3.1].

Theorem 3.1 *With the notation above. Let $\{K_M\}_{M\ge 1}$ be a sequence of compact convex sets of $\mathbb{R}^d$ contained in Θ and containing θ. Let $L > 0$. Suppose that the following conditions are satisfied:*

(i) Θ is an open convex set of $\mathbb{R}^d$.

(ii) For each $t \in \Theta$, $\log f(X,t) \in \mathcal{L}_\theta^{\Phi_1}$.

(iii) For each x, $\log f(x,\cdot)$ is a concave function.

(iv)

$$\lim_{M\to\infty} \sup_{t\in\partial K_M} \inf_{\lambda\in\mathbb{R}} E_\theta[\exp(\lambda(\log f(X,t) - \log f(X,\theta)))] = 0.$$

Then,

$$\begin{aligned} & -\inf\{J_\theta(l) : l \in (\mathcal{L}_\theta^{\Phi_1})^*, \\ & \textstyle\int_{\mathbb{R}^d} I(t\in\Theta : l(\log f(\cdot,t) - \log f(\cdot,\theta)) > 0)\,dt > L\} \\ \le\ & \liminf_{n\to\infty} n^{-1}\log\left(\mathbb{P}_\theta\{\theta \notin C_{HL,L}(X_1,\dots,X_n)\}\right) \\ \le\ & \limsup_{n\to\infty} n^{-1}\log\left(\mathbb{P}_\theta\{\theta \notin C_{HL,L}(X_1,\dots,X_n)\}\right) \\ \le\ & -\inf\{J_\theta(l) : l \in (\mathcal{L}_\theta^{\Phi_1})^*, \textstyle\int_{\mathbb{R}^d} I(t\in\Theta : l(\log f(\cdot,t) - \log f(\cdot,\theta)) \ge 0)\,dt \ge L\} \end{aligned} \tag{3.10}$$

The differences between the previous theorem and Theorem 3.1 in Arcones [2] are that we assume that Θ is an open and convex set, but we do not assume that $\log f(x,\cdot)$ is a strictly concave function. It is easy to see that minor variations in the proof of Theorem 3.1 in Arcones [2] gives the previous theorem.

The previous theorem gives the following when applied to a location family.

Corollary 3.2 *Let $\{X_j\}_{j=1}^\infty$ be a sequence of $\mathbb{R}^d$–valued i.i.d.r.v.'s from the family $\{f(\cdot - \theta) : \theta \in \mathbb{R}^d\}$. Suppose that the following conditions are satisfied:*

(i) $\log f(\cdot)$ is a concave function.

(ii) For each $t \in \Theta$, there exists $\lambda_t > 0$ such that

$$E_0[\exp(\lambda_t |\log f(X - t)|)] < \infty.$$

(iii)

$$\lim_{M\to\infty} \sup_{|t|=M} \inf_{\lambda\in\mathbb{R}} E_\theta[\exp(\lambda(\log f(X-t) - \log f(X-\theta)))] = 0.$$

Then,

$$\begin{aligned} & -\inf\{J_\theta(l) : l \in (\mathcal{L}_\theta^{\Phi_1})^*, \textstyle\int_{\mathbb{R}^d} I(t\in\Theta : \\ & l(\log(f(\cdot - t)/f(\cdot-\theta)) > 0)\,dt > L\} \\ \le\ & \liminf_{n\to\infty} n^{-1}\log\left(\mathbb{P}_\theta\{\theta \notin C_{HL,L}(X_1,\dots,X_n)\}\right) \\ \le\ & \limsup_{n\to\infty} n^{-1}\log\left(\mathbb{P}_\theta\{\theta \notin C_{HL,L}(X_1,\dots,X_n)\}\right) \\ \le\ & -\inf\{J_\theta(l) : l \in (\mathcal{L}_\theta^{\Phi_1})^*, \textstyle\int_{\mathbb{R}^d} I(t\in\Theta : l(\log f(\cdot - t)/f(\cdot-\theta)) \ge 0)\,dt \ge L\}. \end{aligned} \tag{3.11}$$

We are able to find an expression for the bounds in (3.11) for one dimensional location families. To do that, we need the following lemma:

Lemma 3.3 *Let X be a r.v. with p.d.f. f. Let*

$$S(t) := -\inf_{0<\lambda<1} \log \int_{\mathbb{R}} (f(x-t))^{\lambda}(f(x))^{1-\lambda}\, dx, t \in \mathbb{R}.$$

Suppose that:
(i) For each $x \in \mathbb{R}$, $f(x) > 0$.
(ii) $\log f(\cdot)$ is a concave function.
(iii) For each t $\neq 0$, there exists $\lambda_t > 0$ such that

$$E[\exp(\lambda_t |\log(f(X-t)/f(X))|)] < \infty.$$

Then,
(i)

$$S(t) = -\inf_{\lambda\in\mathbb{R}} \log \textstyle\int_{\mathbb{R}} (f(x-t))^{\lambda}(f(x))^{1-\lambda}\, dx.$$

(ii) For each t $\neq 0$, there exists a unique $\lambda_t \in (0,1)$ such that

$$S(t) = -\log \int_{\mathbb{R}} (f(x-t))^{\lambda_t}(f(x))^{1-\lambda_t}\, dx,\ t \in \mathbb{R}.$$

Besides, for each t $\neq 0$, $S(t) > 0$.
(iii) S is increasing on $[0,\infty)$ and S is decreasing in $(-\infty, 0]$.
(iv) S is a continuous function.
(v) $\lim_{|t|\to\infty} S(t) = \infty$.

Theorem 3.4 *Let $\{X_j\}_{j=1}^{\infty}$ be a sequence of i.i.d.r.v.'s from the family $\{f(\cdot-\theta) : \theta \in \mathbb{R}\}$. Suppose that:*
(i) For each $x \in \mathbb{R}$, $f(x) > 0$.
(ii) $\log f(\cdot)$ is a strictly concave function
(iii) For each t $\in \mathbb{R}$, there exists $\lambda_t > 0$ such that

$$E_0[\exp(\lambda_t |\log f(X-t)|)] < \infty.$$

Then,

$$\lim_{n\to\infty} n^{-1} \log \mathbb{P}_\theta\{\theta \notin C_{HL,L}(X_1,\ldots,X_n)\} = -S(L).$$

Under the conditions in the previous theorem, the smaller set in (2.6) is an open interval, and the bigger set in (2.6) is a closed interval. Hence, we may take $C^*_{HL,L}(\vec{Z})$ to be a closed interval. Let $T_{n,HL,b} := T_{n,HL,2b}(X_1,\ldots,X_n)$ be the middle point of $X_n + C^*_{HL,2b}(\vec{Z})$. $T_{n,HL}$ satisfies (1.8) and (1.9).

Recall that for a convex function, the right and left derivatives exists everywhere, and they are equal everywhere except for countably many points (see Proposition 5.16 in Royden [21]). By an abuse of notation, we will denote by f' to the right derivative of f. The following theorem deals with the consistency of $T_{n,HL,b}$.

Theorem 3.5 *Consider a one–dimensional location family. Suppose that:*
(i) For each $x \in \mathbb{R}$, $f(x) > 0$.
(ii) $\log f(\cdot)$ is a concave function
(iii) $\int_{\mathbb{R}} f'(x)\,dx = 0$.
(iv) For each $a \in \mathbb{R}$, $\int_{\mathbb{R}} |\log f(x-a)| f(x) < \infty$.
Then,
(i) $E_\theta[\log f(X-t)] = Q(t-\theta)$, where $Q(a) = \int_{\mathbb{R}} \log f(x-a) f(x)\,dx$.
(ii) Q is decreasing in $[0,\infty)$ and increasing in $(-\infty,0]$.
(iii) Q is continuous.
(iv) There exists a unique $a \in [-b,b]$ such that $Q(a-b) = Q(a+b)$.
(v) $T_{n,HL,b} \overset{\mathbb{P}_\theta}{\to} \theta + a$.

It follows from the previous theorem that $T_{n,HL,b}$ is a consistent estimator of θ if and only if $Q(-b) = Q(b)$.

Next theorem gives an expression for the bound in (1.5) for a location family.

Theorem 3.6 *Consider a one–dimensional location family. Suppose that:*
(i) For each $x \in \mathbb{R}$, $f(x) > 0$.
(ii) $\log f(\cdot)$ is a concave function
(iii) $\int_{\mathbb{R}} f'(x)\,dx = 0$.
(iv) For each $a \in \mathbb{R}$, $\int_{\mathbb{R}} |\log f(x-a)| f(x) < \infty$.
Then,

$$\begin{aligned} &B(b) := \inf\{K(f(\cdot-\theta_1), f(\cdot-\theta)) : \theta_1 \text{ satisfying } |\theta_1-\theta| > b\} \\ &= \min(K(-b), K(b)), \end{aligned}$$

where $K(a) = \int_{\mathbb{R}} \log(f(x)/f(x+a)) f(x)\,dx$.

Next, we consider the asymptotics of the confidence region using the pivot $\theta - \hat{\theta}_n$. we see that the conditions in the previous theorem hold for one dimensional location classes such that $\log f$ is a strictly concave function and other regularity conditions hold.

Theorem 3.7 *Assume that $d = 1$. Suppose that the following conditions are satisfied:*
(i) For each $x \in \mathbb{R}$, $f(x) > 0$.
(ii) $\log f$ is a strictly concave function.
(iii) $\int_{\mathbb{R}} f'(x)\,dx = 0$.
(iv) $E[(f(X))^{-2}(f'(X))^2] < \infty$.
Then, $\hat{\theta}_n$ is well defined and $\theta - \hat{\theta}_n$, when θ obtains, has pdf

$$h_n(t) := E_0\left[I\left(\sum_{j=1}^n (f(X_j+t))^{-1} f'(X_j+t) > 0\right)\sum_{j=1}^n (f(X_j))^{-1} f'(X_j)\right], t \in \mathbb{R}.$$

Besides, h_n is nonincreasing in $[0,\infty)$ and nondecreasing in $(-\infty,0]$.

Under the conditions in the previous theorem, we may choose $C_{piv,L}$ to be a closed interval $[u_{piv,L}, v_{piv,L}]$. Hence, $\hat{\theta}_n + C_{piv,L} = [\hat{\theta}_n + u_{piv,L}, \hat{\theta}_n + v_{piv,L}]$ and $T_{n,piv,b} = \hat{\theta}_n + 2^{-1}(u_{piv,2b} + v_{piv,2b})$.

If there exists a one dimensional sufficient statistic $S(X_1, \ldots, X_n)$, then both $C_{HL,L}(X_1, \ldots, X_n)$ and $C_{piv,L}(\hat{\theta}_n)$ depend on this statistic and they agree. However, next theorem show that $C_{HL,L}(X_1, \ldots, X_n)$ and $C_{piv,L}(\hat{\theta}_n)$ can agree under a more general condition:

Theorem 3.8 *Suppose that:*

(i) For each $x_1, \ldots, x_n \in \mathbb{R}^d$, $C_{HL,L}(x_1, \ldots, x_n) = \{\theta \in \mathbb{R}^d : \prod_{j=1}^n f(x_j - \theta) \geq k(x_1, \ldots, x_n)\}$.

(ii) There are functions $T : (\mathbb{R}^d)^n \to \mathbb{R}^d$, $g : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^m$, $h : (\mathbb{R}^d)^n \to \mathbb{R}^k$, *and* $\tau : \mathbb{R}^m \times \mathbb{R}^k \to \mathbb{R}$ *such that for each* $x_1, \ldots, x_n, \theta \in \mathbb{R}^d$,

$$\prod_{j=1}^n f(x_j - \theta) = \tau(g(T(X_1, \ldots, X_n), \theta), h(X_1, \ldots, X_n)).$$

(iii) $\hat{\theta}_n = \eta(T(X_1, \ldots, X_n))$ *where* $\eta : B \to \mathbb{R}^d$ *is a one–to–one function and* B *is the range of* T.

(iv) When θ *obtains,* $\theta - \hat{\theta}_n$ *has p.d.f.* h_n.

Then, $C_{HL,L}(X_1, \ldots, X_n)$ *is based on the m.l.e.* $\hat{\theta}_n$.

In the situation of the previous theorem, for convenient choices $C_{piv,L}(\hat{\theta}_n)$ and $C_{HL,L}(X_1, \ldots, X_n)$ agree. The previous theorem applies to Examples 3.12–3.14. We have that if there exists a one dimensional sufficient statistic, the previous theorem applies. But, Theorem 3.8 applies to families which do not have a natural sufficient statistic of dimension one (see Example 3.14). The family in Example 3.14 is a curved exponential family of dimension two.

Next, we consider the large deviations of the complementary of the coverage probability.

Theorem 3.9 *Suppose that:*

(i) When θ *obtains,* $\hat{\theta}_n - \theta$ *satisfies the LDP with speed* n *and continuous rate function*

$$R(t) := -\inf_{\lambda \in \mathbb{R}^d} \log\left(E_\theta[\exp\left(\lambda' \nabla_t \log f(X - t)\right)]\right), \tag{3.12}$$

where ∇_t *denotes the (vector of partial derivatives) gradient of* $\log f(x - t)$.

(ii) For each $0 < M < \infty$,

$$\lim_{n\to\infty} \sup_{|x| \leq M} |n^{-1} \log h_n(x) + R(-x)| \to 0.$$

(iii)

$$\lim_{M\to\infty} \limsup_{n\to\infty} \sup_{|x| \geq M} n^{-1} \log h_n(x) = -\infty.$$

(iv) For each $u \geq 0$, $\int I(x \in \mathbb{R}^d : R(x) = u)\, dx = 0$.

Then,

$$\lim_{n\to\infty} n^{-1}\log \mathbb{P}_\theta\{\theta-\hat\theta_n \notin C_{piv,L}(\hat\theta_n)\} = -A(L) \tag{3.13}$$

where

$$A(L) := \inf\{u \ge 0 : \int_{y\in\mathbb{R}^d : R(y)\le u} 1\, dy \ge L\}. \tag{3.14}$$

In the one dimensional situation, we are able to give an expression for the rate in the previous theorem.

Theorem 3.10 *Assume that* $d = 1$*. Suppose that the following conditions are satisfied:*
(i) For each $x \in \mathbb{R}$*,* $f(x) > 0$*.*
(ii) $\log f$ *is a concave function.*
(iii) $E_0[(f(X))^{-1}f'(X)] = 0$*, where* f' *is the right derivative of* f*.*
(iv) $E_0[(f(X))^{-2}(f'(X))^2] < \infty$*.*
(v) $\hat\theta_n - \theta$ *satisfies the LDP with speed* n *and continuous rate function*

$$R(t) := -\inf_{\lambda\in\mathbb{R}} \log E_0[\exp(\lambda(f(X-t))^{-1}f'(X-t))], t \in \mathbb{R}. \tag{3.15}$$

Then,

$$\lim_{n\to\infty} n^{-1}\log \mathbb{P}_\theta\{\theta-\hat\theta_n \notin C_{piv,L}(\hat\theta_n)\} = -R(t_0), \tag{3.16}$$

where $t_0 \in (0, L)$ *satisfies that* $R(t_0) = R(t_0 - L)$*. Besides,*

$$R(t_0) = \inf\{u \ge 0 : \int_{x:R(x)\le u} 1\, dx \ge L\}.$$

Notice that if R is even, then $R(2^{-1}L) = \inf\{u \ge 0 : \int_{x:R(x)\le u} 1\, dx \ge L\}$. Arcones [3, Theorem 3.4] gives sufficient conditions to condition (v) in the previous theorem to hold.

It follows from the theorem that the inaccuracy rate of the optimal estimator $\hat\theta_n + 2^{-1}(u_{piv,L} + v_{piv,L})$ is $R(t_0)$ when $L = 2b$.

Theorem 3.11 *Suppose that:*
(i) For each $x \in \mathbb{R}$*,* $f(x) > 0$*.*
(ii) $\log f$ *is a concave function.*
(iii) f *is even.*
Then,
(i) For each $b > 0$*,* $R(b) \le S(2b)$*.*
(ii) Assume that f *has a third derivative and that* f *is not a p.d.f. from the families of p.d.f.'s in examples 3.12 and 3.14, then for some* $b \ne 0$*,* $R(b) < S(2b)$*.*

Next, we present how the previous theorems apply to several examples.

Example 3.12 If

$$f(x) = (2\pi)^{-1/2}\sigma^{-1}\exp(-2^{-1}\sigma^{-2}x^2), x \in \mathbb{R},$$

where $\sigma > 0$, then, it is easy to see that:
(i) $T_{HL,b} = T_{piv,b} = \hat{\theta}_n = \bar{X} := n^{-1} \sum_{j=1}^{n} X_j$.
(ii) $C_{HL,2b}(X_1, \ldots, X_n) = C_{piv,2b}(\hat{\theta}_n) = [\bar{X} - b, \bar{X} + b]$.
(iii) $A(2b) = S(2b) = R(b) = B(b) = 2^{-1}\sigma^{-2}b^2$.
(iv) For each $\theta \in \mathbb{R}$, $T_{HL,b} \xrightarrow{\mathbb{P}_\theta} \theta$.

Example 3.13 If

$$f(x) = (\Gamma(\alpha))^{-1}|\gamma|\alpha^\alpha \exp\left(\alpha\gamma x - \alpha e^{\gamma x}\right) \tag{3.17}$$

where $\alpha > 0$ and $\gamma \neq 0$, then, it is possible to see that:
(i) $\hat{\theta}_n = \gamma^{-1} \log \left(n^{-1} \sum_{j=1}^{n} e^{\gamma X_j} \right)$.
(ii)

$$\begin{aligned} & C_{HL,2b}(X_1, \ldots, X_n) = C_{piv,2b}(\hat{\theta}_n) \\ = \ & \hat{\theta}_n \\ & +[\gamma^{-1} \log\left(2^{-1}\gamma^{-1}b^{-1}(e^{2\gamma b} - 1)\right) - 2b, \gamma^{-1} \log\left(2^{-1}\gamma^{-1}b^{-1}(e^{2\gamma b} - 1)\right)]. \end{aligned}$$

(iii) $T_{HL,b} = T_{piv,b} = \hat{\theta}_n + \gamma^{-1} \log\left(2^{-1}\gamma^{-1}b^{-1}(e^{\gamma b} - e^{-\gamma b})\right)$.
(iv) For each $b > 0$,

$$\begin{aligned} & S(2b) = A(2b) = R(b) \\ = \ & \alpha \log(2^{-1}|\gamma|^{-1}b^{-1}(e^{2|\gamma|b} - 1)) - \alpha(e^{2|\gamma|b} - 1)^{-1}(e^{2|\gamma|b} - 1 - 2|\gamma|b). \end{aligned}$$

(v) For each $b > 0$, $B(b) = \alpha\left(e^{-|\gamma|b} - 1 + |\gamma|b\right)$.
(vi) For each $b > 0$, $S(2b) > B(b)$.
(vii) For each $b > 0$ and each $\theta \in \mathbb{R}$,

$$T_{HL,b} \xrightarrow{\mathbb{P}_\theta} \theta + \gamma^{-1}\log(2^{-1}\gamma^{-1}b^{-1}(e^{\gamma b} - e^{-\gamma b})) \neq \theta.$$

By Theorem 2 in Ferguson [11] the location family in examples 3.12 and 3.13 are the only one dimensional location families, which are exponential families.

Example 3.14 If

$$f(x) = c\exp(-a_1\exp(\tau_1 x) - a_2\exp(-\tau_2 x)), \tag{3.18}$$

where $a_1, a_2, \tau_1, \tau_2 > 0$ and $c = \left(\int_{\mathbb{R}} \exp(-a_1\exp(\tau_1 x) - a_2\exp(-\tau_2 x))\,dx\right)^{-1}$. Then,
(i) $(\sum_{j=1}^{n} e^{\tau_1 X_j}, \sum_{j=1}^{n} e^{-\tau_2 X_j})$ is a minimal sufficient statistic for θ.
(ii) The family $\{f(\cdot - \theta) : \theta \in \mathbb{R}\}$ is a curved exponential family.
(iii)

$$\hat{\theta}_n = (\tau_1 + \tau_2)^{-1} \log \left(a_2^{-1}\tau_2^{-1}a_1\tau_1 \left(\sum_{j=1}^{n} e^{-\tau_2 X_j} \right)^{-1} \sum_{j=1}^{n} e^{\tau_1 X_j} \right).$$

(iv)

$$C_{HL,2b}(X_1, \ldots, X_n) = C_{piv,2b}(\hat{\theta}_n) = \hat{\theta}_n + [t_0 - 2b, t_0],$$

where

$$t_0 = (\tau_1+\tau_2)^{-1}\log\left(\tau_1^{-1}\tau_2(e^{2b\tau_1}-1)(1-e^{-2b\tau_2})^{-1}\right).$$

(v)

$$T_{HL,b} = T_{piv,b} = \hat{\theta}_n + (\tau_1+\tau_2)^{-1}\log\left(\tau_1^{-1}\tau_2(e^{b\tau_1}-e^{-b\tau_1})(e^{b\tau_2}-e^{-b\tau_2})^{-1}\right).$$

Example 3.15 If

$$f(x) = 2^{-1}\exp(-|x|), x\in\mathbb{R}, \tag{3.19}$$

then, $\hat{\theta}_n$ is not uniquely defined. The function $\log f(\cdot)$ is concave, but not strictly concave. Let $X_{(1)},\dots,X_{(n)}$ be the order statistics. $X_1,\dots,X_n$ are all different with probability one. Assume that $X_1,\dots,X_n$ are all different. If n is odd, $\hat{\theta}_n = X_{(2^{-1}(n+1))}$. If n is even, then $\hat{\theta}_n = X_{(2^{-1}n)}$ and $\hat{\theta}_n = X_{(2^{-1}n+1)}$ are both m.l.e.'s. It is easy to see that the previous choice for the m.l.e. theorems 3.2–3.8 apply giving that:

(i) $C_{HL,2b}(X_1,\dots,X_n) \neq C_{piv,2b}(\hat{\theta}_n)$.
(ii) For each $b>0$, $S(2b) = b-\log(1+b)$.
(iii) For each $b>0$, $A(2b) = R(b) = b - 2^{-1}\log\left(2e^b-1\right)$.
(iii) For each $b>0$, $B(b) = e^{-b}-1+b$.
(iv) For each $b>0$, $A(2b) < S(2b) < B(b)$.
(v) For each $\theta\in\mathbb{R}$, $T_{HL,b} \overset{\mathbb{P}_\theta}{\to} \theta$.

Notice that in the first three examples, $C_{HL,2b}(X_1,\dots,X_n) = C_{piv,2b}(\hat{\theta}_n)$. Theorem 3.8 applies to the first three examples. In Example 3.13, $S(2b) > B(b)$, but in Example 3.15, $S(2b) < B(b)$.

4 Proofs.

The proof of Lemma 2.1 is omitted.

Proof of Theorem 2.2: We may assume that the confidence region is based upon $\vec{Z}$ and X_n. Consider the loss function $L(\theta,C) = I(\theta\notin C)$, where $\theta\in\mathbb{R}^d$ and $C\subset\mathcal{B}(R^d)$. The risk of the confidence region $C(\vec{Z},X_n)$ is

$$\begin{aligned} R(\theta, C(\vec{Z},X_n)) &= E_\theta[I(\theta\notin C(\vec{Z},X_n))] \\ &= \int_{(\mathbb{R}^d)^{n-1}}\int_{\mathbb{R}^d} I(\theta\notin C(\vec{z},x_n)) f_{X_n|\vec{Z}}(x_n-\theta|\vec{z}) f_{\vec{Z}}(\vec{z})\, dx_n\, d\vec{z}. \end{aligned} \tag{4.1}$$

We need to show that if $C(\vec{Z},X_n)$ is a confidence region such that for each $\vec{z},x_n$, $\int I(y\in C(\vec{z},x_n))\,dy \le L$, then

$$\sup_{\theta\in\mathbb{R}^d} R(\theta, C(\vec{Z},X_n)) \ge \sup_{\theta\in\mathbb{R}^d} R(\theta, X_n + C^*_{HL,L}(\vec{Z})). \tag{4.2}$$

Given $\delta>0$, take $M>0$ such that $(2M)^d > L$ and

$$P_0\{(\vec{Z},X_n)\in([-M,M]^d)^n\} \ge 1-\delta.$$

Let $c_f := \sup_{x\in\mathbb{R}^d} f(x)$ and let $a := \inf_{x\in[-2M,2M]^d}(f(x))^n > 0$. Take $\tau > M$ such that $\sup_{x\notin[-\tau,\tau]^d} f(x) \le 2^{-1}c_f^{-(n-1)}a$. Consider the a priori p.d.f. $\pi_m(\theta) = (2m)^{-d}I(\theta\in[-m,m]^d)$, where $m>\tau$. Then, the Bayes risk of $C(\vec{Z},X_n)$ is

$$B(\pi_m, C(\vec{Z},X_n)) = (2m)^{-d}\int_{[-m,m]^d} R(\theta, C(\vec{Z},X_n))\,d\theta.$$

By the changes of variable $x_n-\theta=-y$, Fubini's theorem and the change of variables $y-\theta=u$, we have that

$$\begin{aligned}
&\sup_{\theta\in\mathbb{R}^d} R(\theta, C(\vec{Z},X_n)) \ge B(\pi_m, C(\vec{Z},X_n)) \qquad (4.3)\\
=\ & (2m)^{-d}\int_{[-m,m]^d}\int_{(\mathbb{R}^d)^{n-1}}\int_{\mathbb{R}^d} I(\theta\notin C(\vec{z},x_n)) f_{X_n|\vec{Z}}(x_n-\theta|\vec{z}) f_{\vec{Z}}(\vec{z})\,dx_n\,d\vec{z}\,d\theta\\
=\ & (2m)^{-d}\int_{[-m,m]^d}\int_{(\mathbb{R}^d)^{n-1}}\int_{\mathbb{R}^d} I(\theta\notin C(\vec{z},\theta-y)) f_{X_n|\vec{Z}}(-y|\vec{z}) f_{\vec{Z}}(\vec{z})\,dy\,d\vec{z}\,d\theta\\
=\ & (2m)^{-d}\int_{(\mathbb{R}^d)^{n-1}}\int_{\mathbb{R}^d}\int_{[-m,m]^d} I(\theta\notin C(\vec{z},\theta-y)) f_{X_n|\vec{Z}}(-y|\vec{z}) f_{\vec{Z}}(\vec{z})\,d\theta\,dy\,d\vec{z}\\
=\ & (2m)^{-d}\int_{(\mathbb{R}^d)^{n-1}}\int_{\mathbb{R}^d}\int_{y_j-m\le u_j\le y_j+m,\text{ for each }1\le j\le d}\\
& \times I(y-u\notin C(\vec{z},-u)) f_{X_n|\vec{Z}}(-y|\vec{z}) f_{\vec{Z}}(\vec{z})\,du\,dy\,d\vec{z}.\\
=\ & (2m)^{-d}\int_{(\mathbb{R}^d)^{n-1}}\int_{\mathbb{R}^d}\int_{u_j-m\le y_j\le u_j+m,\text{ for each }1\le j\le d}\\
& \times I(y-u\notin C(\vec{z},-u)) f_{X_n|\vec{Z}}(-y|\vec{z}) f_{\vec{Z}}(\vec{z})\,dy\,du\,d\vec{z}.
\end{aligned}$$

We also have that

$$\begin{aligned}
& (2m)^{-d}\int_{(\mathbb{R}^d)^{n-1}}\int_{\mathbb{R}^d}\int_{u_j-m\le y_j\le u_j+m,\text{ for each }1\le j\le d} \qquad (4.4)\\
& \times I(y-u\notin C(\vec{z},-u)) \times f_{X_n|\vec{Z}}(-y|\vec{z}) f_{\vec{Z}}(\vec{z})\,dy\,du\,d\vec{z}\\
\ge\ & (2m)^{-d}\int_{(\mathbb{R}^d)^{n-1}}\int_{-[(m-\tau),m-\tau]^d}\int_{u_j-m\le y_j\le u_j+m,\text{ for each }1\le j\le d} I(y-u\notin C(\vec{z},-u))\\
& \times f_{X_n|\vec{Z}}(-y|\vec{z}) f_{\vec{Z}}(\vec{z})\,dy\,du\,d\vec{z}\\
\ge\ & (2m)^{-d}\int_{(\mathbb{R}^d)^{n-1}}\int_{[-(m-\tau),m-\tau]^d}\int_{[\tau,\tau]^d} I(y\notin u+C(\vec{z},-u))\\
& \times f_{X_n|\vec{Z}}(-y|\vec{z}) f_{\vec{Z}}(\vec{z})\,dy\,du\,d\vec{z}\\
\ge\ & (2m)^{-d}\int_{([-M,M]^d)^{n-1}}\int_{[-(m-\tau),m-\tau]^d}\int_{[\tau,\tau]^d} I(y\notin u+C(\vec{z},-u))\\
& \times f_{X_n|\vec{Z}}(-y|\vec{z}) f_{\vec{Z}}(\vec{z})\,dy\,du\,d\vec{z}.
\end{aligned}$$

Notice that if $y\in[\tau,\tau]^d$ and $u\in[-(m-\tau),m-\tau]^d$, then $u_j-m\le y_j\le u_j+m$, for each $1\le j\le d$.

Now, $u+C(\vec{z},-u)$ is a confidence region such that for each $\vec{z}$ and u,

$$\int_{\mathbb{R}^d} I(y\in u+C(\vec{z},-u))\,dy \le L.$$

Between all confidence regions $C(\vec{z},u)$ with

$$\int_{\mathbb{R}^d} I(y\in C(\vec{z},u))\,dy \le L, \qquad (4.5)$$

for each $u, \vec{z}$, the one minimizing

$$\int_{[\tau,\tau]^d} I(y \notin C(\vec{z},u)) f_{X_n|\vec{Z}}(-y|\vec{z}) f_{\vec{Z}}(\vec{z})\, dy \tag{4.6}$$

is obtained applying Lemma 2.1 to $f \equiv 1$ and

$$g \equiv I(y \in [-\tau,\tau]^d) f_{X_n|\vec{Z}}(-y|\vec{z}) f_{\vec{Z}}(\vec{z}) = I(y \in [-\tau,\tau]^d) \prod_{j=1}^{n-1} f(z_j - y) \times f(y).$$

Notice that this region does not have to depend on u. Let

$$\lambda_{HL,L,\tau}(\vec{z}) = \inf\{t \geq 0 : \int_{t \leq I(y \in [-\tau,\tau]^d) \prod_{j=1}^{n-1} f(z_j - y) \times f(y)} dy \leq L\}. \tag{4.7}$$

A confidence region $C_{HL,L,\tau}(\vec{z})$ minimizes (4.6) subject to (4.5) if

$$\begin{aligned} &\{y \in \mathbb{R}^d : t < I(y \in [-\tau,\tau]^d) \textstyle\prod_{j=1}^{n-1} f(z_j - y) \times f(y), \} \subset C_{HL,L,\tau}(\vec{z}) \\ \subset \quad &\{y \in \mathbb{R}^d : t \leq I(y \in [-\tau,\tau]^d) \textstyle\prod_{j=1}^{n-1} f(z_j - y) \times f(y), \}. \end{aligned} \tag{4.8}$$

Since $\vec{z} \in ([-M,M]^d)^{n-1}$, $[-M,M]^d \subset \{y \in \mathbb{R}^d : a \leq \prod_{j=1}^{n-1} f(z_j - y) \times f(y)\}$. This implies that $a \leq \lambda_{HL,L,\tau}(\vec{z})$. We also have that if $\vec{z} \in ([-M,M]^d)^{n-1}$, then $\{y \in \mathbb{R}^d : a \leq \prod_{j=1}^{n-1} f(z_j - y) \times f(y)\} \subset [-\tau,\tau]^d$. Hence, the factor $I(y \in [\tau,\tau]^d)$ is superfluous in (4.7) and (4.8). This means we can take $C^*_{HL,L}(\vec{z})$ as the confidence region minimizing (4.6). Hence,

$$\begin{aligned} &\sup_{\theta \in \mathbb{R}^d} R(\theta, C(\vec{Z}, X_n)) \\ \geq\ &(2m)^{-d} \textstyle\int_{([-M,M]^d)^{n-1}} \int_{-[(m-\tau),m-\tau]^d} \int_{[-\tau,\tau]^d} I(y \notin C^*_{HL}(\vec{z})) \\ &f_{X_n|\vec{Z}}(-y|\vec{z}) f_{\vec{Z}}(\vec{z})\, dy\, du\, d\vec{z}. \\ =\ &(m-\tau)^d m^{-d} \textstyle\int_{([-M,M]^d)^{n-1}} \int_{[-\tau,\tau]^d} I(y \notin C^*_{HL}(\vec{z})) \\ &f_{X_n|\vec{Z}}(-y|\vec{z}) f_{\vec{Z}}(\vec{z})\, dy\, d\vec{z}. \end{aligned} \tag{4.9}$$

Letting $m \to \infty$, we get that

$$\begin{aligned} &\sup_{\theta \in \mathbb{R}^d} R(\theta, C(\vec{Z}, X_n)) \\ \geq\ &\textstyle\int_{([-M,M]^d)^{n-1}} \int_{[-\tau,\tau]^d} I(y \notin C^*_{HL}(\vec{z})) f_{X_n|\vec{Z}}(-y|\vec{z}) f_{\vec{Z}}(\vec{z})\, dy\, d\vec{z}. \\ \geq\ &P_0\{0 \notin X_n + C^*_{HL}(\vec{Z})\} - \delta. \end{aligned} \tag{4.10}$$

Since $\delta > 0$ is arbitrary, the claim follows. □

In the proof of the next theorem, we need the following lemma:

Lemma 4.1 *Let $h : \mathbb{R}^d \to \mathbb{R}$ be a measurable function. Let $\lambda \in \mathbb{R}$. Let $C_\lambda \in \mathcal{B}(\mathbb{R}^d)$ be such that*

$$\{x \in \mathbb{R}^d : h(x) > \lambda\} \subset C_\lambda \subset \{x \in \mathbb{R}^d : h(x) \geq \lambda\}.$$

Then, for any $C \in \mathcal{B}(\mathbb{R}^d)$,

$$\int_C (\lambda - h(x))\,dx \geq \int_{C_\lambda} (\lambda - h(x))\,dx.$$

The proof of the previous lemma is omitted.

Proof of Theorem 2.3: We abbreviate $\lambda_{piv} = \lambda_{piv,L}$ and $C_{piv} = C_{piv,L}$. If $\lambda_{piv} = 0$, then $\int I(x : 0 < h_n(x))\,dx \leq L$ and $\{x : 0 < h_n(x)\} \subset C_{piv}$. Hence,

$$P_0\{0 \in \hat{\theta}_n + C_{piv}\} = \int_{C_{piv}} h_n(x)\,dx = 1$$

and the claim follows.

Assume that $\lambda_{piv} > 0$. Consider the loss function

$$L(\theta, C) = \lambda_{piv} \int_{\mathbb{R}^d} I(t \in C)\,dt + I(\theta \notin C).$$

We restrict ourselves to decision rules depending on $\hat{\theta}_n$, i.e. confidence regions of the form $C(\hat{\theta}_n)$. The risk of the decision rule $C(\hat{\theta}_n)$ is

$$\begin{aligned} &R(\theta, C(\hat{\theta}_n)) = \textstyle\int_{\mathbb{R}^d} \left(\lambda_{piv} \int_{\mathbb{R}^d} I(t \in C(x))\,dt + I(\theta \notin C(x))\right) h_n(\theta - x)\,dx \\ &= \textstyle\int_{\mathbb{R}^d} \left(\lambda_{piv} \int_{\mathbb{R}^d} I(t \in C(\theta - y))\,dt + I(\theta \notin C(\theta - y))\right) h_n(y)\,dy. \end{aligned}$$

Let $m > \tau > M > 0$. Consider the Bayes a priori p.d.f. $\pi_m(\theta) = (2m)^{-d} I(\theta \in [-m, m]^d)$. Then, by the change of variables $\theta - y = -u$

$$\begin{aligned} &\sup_{\theta \in \mathbb{R}^d} R(\theta, C(\hat{\theta}_n)) = B(\pi_m, C(\hat{\theta}_n)) \\ &= (2m)^{-d} \textstyle\int_{[-m,m]^d} \int_{\mathbb{R}^d} \left(\lambda_{piv} \int_{\mathbb{R}^d} I(t \in C(\theta - y))\,dt\right. \\ &\quad \left.+ I(\theta \notin C(\theta - y))\right) h_n(y)\,dy\,d\theta \\ &= (2m)^{-d} \textstyle\int_{\mathbb{R}^d} \int_{[-m,m]^d} \left(\lambda_{piv} \int_{\mathbb{R}^d} I(t \in C(\theta - y))\,dt\right. \\ &\quad \left.+ I(\theta \notin C(\theta - y))\right) h_n(y)\,d\theta\,dy \\ &= (2m)^{-d} \textstyle\int_{\mathbb{R}^d} \int_{y_j - m \leq u_j \leq y_j + m} \left(\lambda_{piv} \int_{\mathbb{R}^d} I(t \in C(-u))\,dt\right. \\ &\quad \left.+ I(y \notin u + C(-u))\right) h_n(y)\,du\,dy \\ &= (2m)^{-d} \textstyle\int_{\mathbb{R}^d} \int_{u_j - m \leq y_j \leq u_j + m} \left(\lambda_{piv} \int I(t \in C(-u))\,dt\right. \\ &\quad \left.+ I(y \notin u + C(-u))\right) h_n(y)\,dy\,du \\ &\geq (2m)^{-d} \textstyle\int_{[-(m-\tau), m-\tau]^d} \int_{[-\tau,\tau]^d} \left(\lambda_{piv} \int_{\mathbb{R}^d} I(t \in C(-u))\,dt\right. \\ &\quad \left.+ I(y \notin u + C(-u))\right) h_n(y)\,dy\,du. \end{aligned}$$

We have that

$$\begin{aligned} &(2m)^{-d} \textstyle\int_{[-(m-\tau), m-\tau]^d} \int_{[-\tau,\tau]^d} \lambda_{piv} \int I(t \in C(-u))\,dt h_n(y)\,dy\,du \\ &= \mathbb{P}\{\theta - \hat{\theta}_n \in [-\tau, \tau]^d\} (2m)^{-d} \textstyle\int_{[-(m-\tau), m-\tau]^d} \lambda_{piv} \int_{\mathbb{R}^d} I(t \in C(-u))\,dt\,du \\ &= \mathbb{P}\{\theta - \hat{\theta}_n \in [-\tau, \tau]^d\} (2m)^{-d} \textstyle\int_{[-(m-\tau), m-\tau]^d} \int_{\mathbb{R}^d} \lambda_{piv} I(y \in C(-u))\,dy\,du \\ &= \mathbb{P}\{\theta - \hat{\theta}_n \in [-\tau, \tau]^d\} (2m)^{-d} \textstyle\int_{[-(m-\tau), m-\tau]^d} \int_{\mathbb{R}^d} \lambda_{piv} I(y \in u + C(-u))\,dy\,du \\ &\geq \mathbb{P}\{\theta - \hat{\theta}_n \in [-\tau, \tau]^d\} (2m)^{-d} \\ &\quad \times \textstyle\int_{[-(m-\tau), m-\tau]^d} \int_{[-\tau,\tau]^d} \lambda_{piv} I(y \in u + C(-u))\,dy\,du \end{aligned}$$

and

$$\begin{aligned}
&(2m)^{-d}\int_{[-(m-\tau),m-\tau]^d}\int_{[-\tau,\tau]^d} I(y\notin u+C(-u))h_n(y)\,dy\,du\\
\ge\ &\mathbb{P}\{\theta-\hat\theta_n\in[-\tau,\tau]^d\}(2m)^{-d}\int_{[-(m-\tau),m-\tau]^d}\int_{[-\tau,\tau]^d}\\
&\times I(y\notin u+C(-u))h_n(y)\,dy\,du\\
=\ &(m-\tau)^d m^{-d}\left(\mathbb{P}\{\theta-\hat\theta_n\in[-\tau,\tau]^d\}\right)^2\\
&-\mathbb{P}\{\theta-\hat\theta_n\in[-\tau,\tau]^d\}(2m)^{-d}\int_{[-(m-\tau),m-\tau]^d}\int_{[-\tau,\tau]^d}\\
&\times I(y\in u+C(-u))h_n(y)\,dy\,du.
\end{aligned}$$

Thus,

$$\begin{aligned}
&R(\theta,C(\hat\theta_n))\\
\ge\ &(m-\tau)^d m^{-d}\left(\mathbb{P}\{\theta-\hat\theta_n\in[-\tau,\tau]^d\}\right)^2\\
&+\mathbb{P}\{\theta-\hat\theta_n\in[-\tau,\tau]^d\}(2m)^{-d}\\
&\times\int_{[-(m-\tau),m-\tau]^d}\int_{[-\tau,\tau]^d}(\lambda_{piv}-h_n(y))\,I(y\in u+C(-u))\,dy\,du.
\end{aligned}$$

By Lemma 4.1,

$$\begin{aligned}
&(2m)^{-d}\int_{[-(m-\tau),m-\tau]^d}\int_{[-\tau,\tau]^d}(\lambda_{piv}-h_n(y))\,I(y\in u+C(-u))\,dy\,du\\
\ge\ &(2m)^{-d}\int_{[-(m-\tau),m-\tau]^d}\int_{[-\tau,\tau]^d}(\lambda_{piv}-h_n(y))\,I(y\in C_{piv})\,dy\,du\\
=\ &(m-\tau)^d(m)^{-d}\int_{[-\tau,\tau]^d}(\lambda_{piv}-h_n(y))\,I(y\in C_{piv})\,dy.
\end{aligned}$$

Hence,

$$\begin{aligned}
&\sup_{\theta\in\mathbb{R}^d} R(\theta,C(\hat\theta_n))\\
\ge\ &(m-\tau)^d m^{-d}\left(\mathbb{P}\{\theta-\hat\theta_n\in[-\tau,\tau]^d\}\right)^2\\
&+\mathbb{P}\{\theta-\hat\theta_n\in[-\tau,\tau]^d\}(m-\tau)^d(m)^{-d}\int_{[-\tau,\tau]^d}(\lambda_{piv}-h_n(y)I(y\in C_{piv})\,dy.
\end{aligned}$$

Letting $m\to\infty$, and then $\tau\to\infty$, we get that

$$\begin{aligned}
&\sup_{\theta\in\mathbb{R}^d} R(\theta,C(\hat\theta_n)) \qquad (4.11)\\
\ge\ &1+\int_{\mathbb{R}^d}(\lambda_{piv}-h_n(y)I(y\in C_{piv})\,dy\\
=\ &\int_{\mathbb{R}^d}\left(\lambda_{piv}I(y\in C_{piv})+h_n(y)I(y\notin C_{piv})\right)\,dy=R(\theta,\hat\theta_n+C_{piv}).
\end{aligned}$$

If $C(\hat\theta_n)$ satisfies (2.18), (4.11) implies that $C(\hat\theta_n)$ satisfies (2.19). $\square$

Proof of Lemma 3.3: (i) If $t=0$, (i) is trivial. Assume that $t\neq 0$. If $f(x-t)$ were equal to $f(x)$ a.e. with respect to the Lebesgue measure, then for each a

$$\mathbb{P}\{X\le a\}=\int_{-\infty}^a f(s)\,ds=\int_{-\infty}^a f(s-t)\,ds=\mathbb{P}\{X\le a-t\},$$

in contradiction. Hence, by the strict concavity of log, $E_0[\log(f(X-t)/f(X))]<0$ and $E_t[\log(f(X)/f(X-t))]<0$. Thus, by (3.6)

$$\begin{aligned}
&-\inf_{\lambda\in\mathbb{R}}\log\int_{\mathbb{R}}(f(x-t))^\lambda(f(x))^{1-\lambda}\,dx\\
=\ &-\inf_{\lambda\in\mathbb{R}}\log E_0[\exp(\lambda\log(f(X-t)/f(X)))],\\
=\ &-\inf_{\lambda>0}\log E_0[\exp(\lambda\log(f(X-t)/f(X)))],
\end{aligned}$$

and

$$\begin{aligned} & -\inf_{\lambda\in\mathbb{R}} \log \textstyle\int_{\mathbb{R}} (f(x-t))^{\lambda} (f(x))^{1-\lambda}\, dx \\ = & -\inf_{\lambda\in\mathbb{R}} \log E_t[\exp((1-\lambda)\log(f(X)/f(X-t)))], \\ = & -\inf_{\lambda<1} \log E_t[\exp((1-\lambda)\log(f(X)/f(X-t)))], \end{aligned}$$

By the previous two inequalities,

$$\begin{aligned} & -\inf_{\lambda\in\mathbb{R}} \log \textstyle\int_{\mathbb{R}} (f(x-t))^{\lambda} (f(x))^{1-\lambda}\, dx \\ = & -\inf_{0<\lambda<1} \log \textstyle\int_{\mathbb{R}} (f(x-t))^{\lambda} (f(x))^{1-\lambda}\, dx \end{aligned}$$

(ii) Let $t \neq 0$. Since it is not true that $f(x-t) = f(x)$ a.e. x,

$$m(\lambda) = \log \int_{\mathbb{R}} (f(x-t))^{\lambda} (f(x))^{1-\lambda}\, dx, 0 \le \lambda \le 1,$$

is a strictly convex function with $m(0) = m(1) = 0$. Hence, there exists a unique $\lambda_t \in (0,1)$ such that

$$S(t) = -\log \int_{\mathbb{R}} (f(x-t))^{\lambda_t} (f(x))^{1-\lambda_t}\, dx.$$

Besides, $S(t) > 0$.

(iii) Given $0 \le s < t$, we show that $S(s) < S(t)$. By concavity,

$$\log(f(x-s)) \ge (t^{-1}s)\log(f(x-t)) + (1-t^{-1}s)\log(f(x)).$$

Let us prove that

$$\int I(\log(f(x-s)) > (t^{-1}s)\log(f(x-t)) + (1-t^{-1}s)\log(f(x)))\, dx > 0, \quad (4.12)$$

by contradiction. Assume that

$$\log(f(x-s)) = (t^{-1}s)\log(f(x-t)) + (1-t^{-1}s)\log(f(x)) \text{ a.e.}$$

Then, by the Hölder inequality,

$$\begin{aligned} 1 = & \textstyle\int f(x-s)\, dx = \int (f(x-t))^{t^{-1}s} (f(x))^{1-t^{-1}s}\, dx \\ \le & \left(\textstyle\int f(x-t)\, dx\right)^{t^{-1}s} \left(\textstyle\int f(x)\, dx\right)^{1-t^{-1}s} = 1. \end{aligned}$$

Hence, by the reverse of the Hölder inequality (see e.g. Theorem 5.32.2, in Royden [21]), there are $a, b \ge 0$ with $a + b > 0$ such that $af(x-t) = bf(x)$ a.e. in contradiction.

By (ii), there exists $0 < \lambda_s < 1$ such that

$$S(s) = -\log \int_{\mathbb{R}} (f(x-s))^{\lambda_s} (f(x))^{1-\lambda_s}\, dx$$

Hence, by (4.12),

$$\begin{aligned} S(s) = & -\log \textstyle\int_{\mathbb{R}} \exp\left(\lambda_s \log(f(x-s)/f(x))\right) f(x)\, dx \\ < & -\log \textstyle\int_{\mathbb{R}} \exp\left(\lambda_s t^{-1}s \log(f(x-t)/f(x))\right) f(x)\, dx \le S(t). \end{aligned}$$

The proof of that S is decreasing in $(-\infty, 0]$ is similar and it is omitted.

(iv) Since $\log f(\cdot)$ is a concave function and $\lim_{|x|\to\infty} \log f(x) = -\infty$, there exists $x_0 \in \mathbb{R}$, such that f is nondecreasing in $(-\infty, x_0]$ and nonincreasing in $[x_0, \infty)$. By a change of variables, we may assume that $x_0 = 0$. We have that:

$$\begin{array}{ll} \text{if } x \le -1, & f(x-h) \le f(x+1); \\ \text{if } -1 \le x \le 1, & f(x-h) \le f(0); \\ \text{if } 1 \le x, & f(x-h) \le f(x-1). \end{array}$$

Therefore,

$$\int_{\mathbb{R}} \sup_{|h|\le 1} f(x-h)\, dx < \infty. \tag{4.13}$$

Suppose that $t_n \to t_0 \in \mathbb{R}$, we need to prove that $S(t_n) \to S(t_0)$. There are $\lambda_{t_n}, \lambda_0 \in [0,1]$ such that

$$S(t_n) = -\log \int_{\mathbb{R}} (f(x-t_n))^{\lambda_{t_n}} (f(x))^{1-\lambda_{t_n}}\, dx$$

and

$$S(t_0) = -\log \int_{\mathbb{R}} (f(x-t_0))^{\lambda_{t_0}} (f(x))^{1-\lambda_{t_0}}\, dx.$$

We have that for n large enough, and $0 \le \lambda \le 1$,

$$(f(x-t_n))^{\lambda}(f(x))^{1-\lambda} \le \lambda f(x-t_n) + (1-\lambda) f(x) \le \sup_{|h|\le 1} f(x-t_0+h) + f(x), \tag{4.14}$$

whose integral in $\mathbb{R}$ is finite. Hence, by the dominated convergence theorem,

$$\begin{array}{l} \liminf_{n\to\infty} S(t_n) \ge \liminf_{n\to\infty} -\log \int_{\mathbb{R}} (f(x-t_n))^{\lambda_{t_0}} (f(x))^{1-\lambda_{t_0}}\, dx \\ = -\log \int_{\mathbb{R}} (f(x-t_0))^{\lambda_{t_0}} (f(x))^{1-\lambda_{t_0}}\, dx = S(t_0). \end{array}$$

Let $\{n_k\}$ be a subsequence such that

$$\lim_{k\to\infty} S(t_{n_k}) = \limsup_{n\to\infty} S(t_n)$$

and $\lambda_{n_k} \to \bar{\lambda}$, for some $\bar{\lambda}$. By (4.14),

$$\begin{array}{l} \limsup_{n\to\infty} S(t_n) = \lim_{k\to\infty} -\log \int_{\mathbb{R}} (f(x-t_{n_k}))^{\lambda_{t_{n_k}}} (f(x))^{1-\lambda_{t_{n_k}}}\, dx \\ = -\log \int_{\mathbb{R}} (f(x-t_0))^{\bar{\lambda}} (f(x))^{1-\bar{\lambda}}\, dx \le S(t_0). \end{array}$$

Hence, $S(t_n) \to S(t_0)$.

(v) Take $t_0 > 0$ such that $f(t_0) < f(0)$ and $f(-t_0) < f(0)$. By concavity, for each $x > t_0$,

$$\log(f(t_0)/f(0)) \ge t_0 x^{-1} \log(f(x)/f(0)).$$

Hence,

$$f(x) \le f(0) \exp(-x t_0^{-1} \log(f(0)/f(t_0))).$$

Similarly, we get that for $x < -t_0$,

$$f(x) \le f(0)\exp(-|x|t_0^{-1}\log(f(0)/f(-t_0)).$$

Hence, for each $\lambda > 0$,

$$\int_{\mathbb{R}} (f(x))^{\lambda}\,dx < \infty. \tag{4.15}$$

We have that

$$\sup_{t\in\mathbb{R}} (f(x-t))^{1/2}(f(x))^{1/2} \le (f(0))^{1/2}(f(x))^{1/2}$$

and

$$\int_{\mathbb{R}} (f(0))^{1/2}(f(x))^{1/2}\,dx < \infty.$$

Hence, by the dominated convergence theorem,

$$\begin{array}{l} \lim_{|t|\to\infty} S(t) \ge \lim_{|t|\to\infty} -\log\int_{\mathbb{R}}(f(x-t))^{1/2}(f(x))^{1/2}\,dx \\ = -\log\int_{\mathbb{R}} \lim_{|t|\to\infty}(f(x-t))^{1/2}(f(x))^{1/2}\,dx = \infty. \quad \square \end{array}$$

We will need the following lemmas.

Lemma 4.2 *Let X be a r.v. Let A be a Borel set of $\mathbb{R}$. Let Φ_1^{-1} be the inverse of Φ_1 restricted to $[0,\infty)$. Then,*

$$N_{\Phi_1}(I(X\in A)) = (\Phi_1^{-1}(1/\mathbb{P}\{X\in A\}))^{-1}.$$

Proof: We have that

$$E[\Phi_1(t^{-1}I(X\in A))] = \mathbb{P}\{X\in A\}(e^{t^{-1}} - 1 - t^{-1}).$$

So, $1 \ge E[\Phi_1(t^{-1}I(X\in A))]$ is equivalent to $t \ge (\Phi_1^{-1}(1/\mathbb{P}\{X\in A\}))^{-1}$.

$\square$

Lemma 4.3 *Let $g, h : \mathbb{R}\to\mathbb{R}$ be two measurable functions. Let $l \in (\mathcal{L}^{\Phi_1})^*$ with $J(l) < \infty$. If $\mathbb{P}\{g(X) < h(X)\} = 1$, then $l(g) < l(h)$.*

Proof: Using Lemma 4.2 and $\mathbb{P}\{h(X) - g(X) \ge m^{-1}\} \to 1$, as $m\to\infty$, we get that $I(h(X) - g(X) \ge m^{-1}) \xrightarrow{N_{\Phi_1}} 1$, as $m\to\infty$. By (3.9), $l(I(h(X) - g(X) \ge m^{-1}) \to l(1) = 1$, as $m\to\infty$. Take m such that $l(I(h(X) - g(X) \ge m^{-1}) > 0$. Then,

$$\begin{array}{l} l(h-g) \\ = l((h-g)I(h-g\ge m^{-1})) + l((h-g)I(m^{-1} > h-g \ge m^{-1} \ge 0)) \\ \quad + l((h-g)I(0 > h-g)) \\ \ge l((h-g)I(h-g\ge m^{-1}) > 0. \end{array}$$

$\square$

Proof of Theorem 3.4: We apply Corollary 3.2. Hypothesis (i) and (ii) in Corollary 3.2 are assumed. By Lemma 3.3,

$$\begin{aligned} &\sup_{|t|=M} \inf_{\lambda\in\mathbb{R}} E_\theta[\exp(\lambda(\log f(X-t) - \log f(X-\theta)))] \\ = &\sup_{|t|=M} \inf_{\lambda\in\mathbb{R}} E_0[\exp(\lambda(\log f(X-t) - \log f(X)))] \\ = &\sup_{|t|=M} \exp(-S(t)) \to 0, \end{aligned}$$

as $M \to \infty$. Hence, Corollary 3.2 gives that

$$\begin{aligned} &-S_1(\theta, L) \le \liminf_{n\to\infty} n^{-1} \log\left(\mathbb{P}_\theta\{\theta \notin C_{HL,L}(X_1,\ldots,X_n)\}\right) \\ \le &\limsup_{n\to\infty} n^{-1} \log\left(\mathbb{P}_\theta\{\theta \notin C_{HL,L}(X_1,\ldots,X_n)\}\right) \le -S_2(\theta, L) \end{aligned}$$

where

$$\begin{aligned} &S_1(\theta, L) = \inf\{J_\theta(l) : l \in (\mathcal{L}_\theta^{\Phi_1})^*, \\ &\textstyle\int_{\mathbb{R}} I(t \in \mathbb{R} : l(\log(f(\cdot - t)/f(\cdot - \theta))) > 0)\, dt > L\} \end{aligned}$$

and

$$\begin{aligned} &S_2(\theta, L) = \inf\{J_\theta(l) : l \in (\mathcal{L}_\theta^{\Phi_1})^*, \\ &\textstyle\int_{\mathbb{R}} I(t \in \mathbb{R} : l(\log(f(\cdot - t)/f(\cdot - \theta))) \ge 0)\, dt \ge L\}. \end{aligned}$$

So, we need to prove that

$$S_1(\theta, L) = S_2(\theta, L) = S(L).$$

First, we prove that $S_1(\theta, L)$ and $S_2(\theta, L)$ do not depended on θ. Given $\theta \in \mathbb{R}$ and $l \in (\mathcal{L}_\theta^{\Phi_1})^*$, define $l_0 \in (\mathcal{L}_0^{\Phi_1})^*$ as

$$l_0(g) = l(g - \log(f(\cdot - \theta)/f(\cdot))), g \in \mathcal{L}_0^{\Phi_1}.$$

By (3.1),

$$\begin{aligned} J_\theta(l) &= \sup_{g\in\mathcal{L}_\theta^{\Phi_1}} \left(l(g) - \log\left(E_\theta[\exp(g(X))]\right)\right) \\ &= \sup_{g\in\mathcal{L}_\theta^{\Phi_1}} \left(l_0(g + \log(f(\cdot - \theta)/f(\cdot))\right) \\ &\quad - \log\left(E_0[\exp(g(X) + \log(f(X-\theta)/f(X)))]\right)) \\ &= \sup_{g\in\mathcal{L}_0^{\Phi_1}} \left(l_0(g) - \log\left(E_0[\exp(g(X))]\right)\right) = J_0(l_0) \end{aligned}$$

Hence,

$$\begin{aligned} &\inf\{J_\theta(l) : l \in (\mathcal{L}_\theta^{\Phi_1})^*, \textstyle\int_{\mathbb{R}} I(t \in \mathbb{R} : l(\log(f(\cdot - t)/f(\cdot - \theta))) > 0)\, dt > L\} \\ = &\inf\{J_0(l_0) : l_0 \in (\mathcal{L}_0^{\Phi_1})^*, \\ &\textstyle\int_{\mathbb{R}} I(t \in \mathbb{R} : l_0(\log(f(\cdot - t)/f(\cdot - \theta)) + \log(f(\cdot - \theta)/f(\cdot))) > 0)\, dt > L\} \\ = &\inf\{J_0(l) : l \in (\mathcal{L}_0^{\Phi_1})^*, \textstyle\int_{\mathbb{R}} I(t \in \mathbb{R} : l(\log(f(\cdot - t)/f(\cdot))) > 0)\, dt > L\}. \end{aligned}$$

Thus, $S_1(\theta, L)$ does not depend on θ. The proof that $S_2(\theta, L)$ does not depend on θ is similar.

Suppose that $l \in (\mathcal{L}_0^{\Phi_1})^*$ satisfies that $J_0(l) < \infty$, then by Lemma 4.3, $l(\log f(\cdot - t))$, $t \in \mathbb{R}$ is a strictly concave function. Besides we have that $\lim_{|t|\to\infty} l(\log f(\cdot - t)) = -\infty$. Therefore, if $J_0(l) < \infty$ and

$$\int_{\mathbb{R}} I(t \in \mathbb{R} : l(\log(f(\cdot - t)/f(\cdot))) > 0)\, dt > L,$$

the set

$$\{t \in \mathbb{R} : l(\log(f(\cdot - t)/f(\cdot))) > 0\}$$

is an interval of the type either $(s, 0)$ with $s < -L$, or $(0, s)$ with $s > L$. Hence,

$$\begin{aligned} S_1(0, L) &= \inf\{J_0(l) : l \in (\mathcal{L}_0^{\Phi_1})^*, l(\log(f(\cdot - t)) = 0, \\ &\quad \text{for some } s \in (-\infty, -L) \cup (L, \infty)\} \\ &= \min(\inf_{s<-L} S(t), \inf_{t>L} S(t)) = \min(S(-L), S(L)) = S(L). \end{aligned}$$

The proof that $S_2(0, 2L) = S(L)$ is similar and it is omitted. $\square$

Proof of Theorem 3.5: (i) We have that

$$\begin{aligned} &E_\theta[\log f(X - t)] = \textstyle\int_{\mathbb{R}} \log f(x - t) f(x - \theta)\, dx \\ &= \textstyle\int_{\mathbb{R}} \log f(x - t + \theta) f(x)\, dx = Q(t - \theta). \end{aligned}$$

(ii) Since $\log f(\cdot)$ is a concave function, $(f(x))^{-1} f'(x)$ is nonincreasing. So, for $a > 0$, $-(f(x))^{-1} f'(x) \ge -(f(x - a))^{-1} f'(x - a)$. Hence,

$$\begin{aligned} &0 = Q'(0) = \textstyle\int_{\mathbb{R}} (-(f(x))^{-1} f'(x)) f(x)\, dx \\ &\ge \textstyle\int_{\mathbb{R}} (-(f(x - a))^{-1} f'(x - a)) f(x)\, dx = Q'(a). \end{aligned}$$

But, if $Q'(a) = 0$, then $-(f(x))^{-1} f'(x) = -(f(x - a))^{-1} f'(x - a)$ almost everywhere. Hence, for each $x > 0$,

$$\begin{aligned} &\log(f(x)/f(0)) = \textstyle\int_0^x (f(t))^{-1} f'(t))\, dt \\ &= \textstyle\int_0^x (f(t - a))^{-1} f'(t - a))\, dt = \log(f(x - a)/f(-a)). \end{aligned}$$

Similarly, for $x < 0$, $\log(f(x)/f(0)) = \log(f(x - a)/f(-a))$. This implies that for each $x \in \mathbb{R}$, $f(x)/f(0) = f(x - a)/f(-a)$. Since f is a p.d.f. $f(0) = f(-a)$ and for each $x \in \mathbb{R}$, $f(x) = f(x - a)$. Hence, for each $t \in \mathbb{R}$, $\mathbb{P}_0\{X \le t\} = \mathbb{P}_0\{X \le t - a\}$, in contradiction. Hence, $Q'(a) < 0$. We got that for each $a > 0$, $Q'(a) < 0$. Hence, Q is decreasing in $[0, \infty)$. The proof that Q is increasing in $(-\infty, 0]$ is omitted.

(iii) Given $a \in \mathbb{R}$, we would like to show that by the dominated convergence theorem, $\lim_{t \to a} Q(t) = Q(a)$. Given $\tau > 0$, we show that

$$\int_{\mathbb{R}} \sup_{a-\tau<t<a+\tau} |\log f(x - t)|\, f(x)\, dx < \infty. \tag{4.16}$$

Since f is continuous and $\lim_{|x|\to\infty} f(x) = 0$, $c_f := \sup_{x\in\mathbb{R}} f(x) < \infty$. Let t be such that $a - \tau < t < a + \tau$. Since f is continuous, $\{x \in \mathbb{R} : f(x) \ge 1\} = [\alpha, \beta]$, for some $\alpha < \beta$. So,

$$\log f(x - t) \le (\log c_f) I(x \in [\alpha + a - \tau, \beta + a + \tau]).$$

By concavity,

$$\begin{aligned} &\log f(x - t) \ge \min(\log f(x - a - \tau), \log(x - a + \tau)) \\ &\ge -|\log f(x - a - \tau)| - |\log(x - a + \tau)|. \end{aligned}$$

Hence,

$$\begin{aligned}&|\log f(x-t)|\\ \le\ &(\log c_f) I(x \in [\alpha+a-\tau, \beta+a+\tau]) + |\log f(x-a-\tau)| + |\log(x-a+\tau)|.\end{aligned}$$

(iv) Consider $h(t) = Q(t+b) - Q(t-b)$, $-b \le t \le b$. We have that h is decreasing, $h(-b) = Q(0) - Q(-2b) > 0$, and $h(b) = Q(2b) - Q(0) < 0$. Therefore, there exists a unique $a \in (-b, b)$ such that $h(a) = 0$.

(v) Since $T_{n,HL,L}$ is translation equivariant, we may assume that $\theta = 0$. Let $G_n(s) = n^{-1}\sum_{j=1}^n \log f(X_j - s)$. By the law of the numbers, for each s, $G_n(s) \xrightarrow{\mathbb{P}_0} Q(s)$ a.s. Let $0 < \tau < 2^{-2}b$. Hence, in a set of probability one with respect to $\mathbb{P}_0$, for n large enough,

$$G_n(a-b-2\tau) < G_n(a-b-\tau) < G_n(a+b) < G_n(a-b+\tau) < G_n(a-b+2\tau)$$

and

$$G_n(a+b+2\tau) < G_n(a+b+\tau) < G_n(a-b) < G_n(a+b-\tau) < G_n(a+b-2\tau).$$

This implies that G_n attains its maximum at some point d_n in $[a-b+2\tau, a+b-2\tau]$, it is nonincreasing in $(-\infty, d_n)$ and nondecreasing on (d_n, ∞). Let

$$\lambda_{HL,2b} = \inf\{t \ge 0 : \int_{t \le G_n(s)} ds \le 2b\}.$$

Let $C_{HL,2b}(X_1, \ldots, X_n)$ be a convex set such that

$$\int_{\theta \in C_{HL,2b}(X_1,\ldots,X_n)} d\theta = L$$

and

$$\begin{aligned}&\{\theta \in \Theta : \lambda_{HL,L}(X_1,\ldots,X_n) < G_n(\theta)\} \subset C_{HL,2b}(X_1,\ldots,X_n)\\ \subset\ &\{\theta \in \Theta : \lambda_{HL,L}(X_1,\ldots,X_n) \le G_n(\theta)\}.\end{aligned}$$

We have that if $\int_{t \le G_n(s)} ds \le 2b$, then $\lambda_{HL,2b} \le t$; and if $2b < \int_{t \le G_n(s)} ds$, then $t \le \lambda_{HL,2b}$. Since $G_n(a+b) > G_n(a-b-\tau)$,

$$\int_{G_n(a-b-\tau) \le G_n(s)} ds \ge 2b + \tau.$$

So, $\lambda_{HL,2b} \ge G_n(a-b-\tau) > G_n(a-b-2\tau)$. Similarly, we get that $\lambda_{HL,2b} > G_n(a+b+2\tau)$. Hence, $a-b-2\tau, a+b+2\tau \notin C_{HL,2b}$. Since $G_n(a+b) < G_n(a-b+\tau)$,

$$\int_{G_n(a-b+\tau) \le G_n(s)} ds \le 2b - \tau.$$

So, $\lambda_{HL,2b} \le G_n(a-b+\tau) < G_n(a-b+2\tau)$. Similarly, we get that $\lambda_{HL,2b} < G_n(a+b-2\tau)$. Hence, $[a-b+2\tau, a+b-2\tau] \subset C_{HL,2b}$. Since

$$[a-b+2\tau, a+b-2\tau] \subset C_{HL,2b} \subset (a-b-2\tau, a+b+2\tau),$$

$-2\tau \le T_{HL,2b} - a \le 2\tau$. Since $\tau > 0$ is arbitrary, $T_{HL,2b} \overset{\mathbb{P}_0}{\to} a$ a.s. $\square$

Proof of Theorem 3.6: We have that

$$\begin{aligned} &K(f(\cdot - t), f(\cdot - \theta)) = \int_{\mathbb{R}} \log(f(x-t)/f(x-\theta)) f(x-t)\,dx \\ &= \int_{\mathbb{R}} \log(f(x)/(f(x+t-\theta)) f(x)\,dx = K(t-\theta). \end{aligned}$$

Since $K(a) = Q(0) - Q(-a)$, we have that K is continuous, decreasing in $(-\infty, 0)$ and increasing in $[0, \infty)$. So, $B(a) = \min(K(a), K(-a))$. $\square$

Proof of Theorem 3.7: By conditions (i)–(ii), $\hat{\theta}_n$ is well defined. We have that for each $t \in \mathbb{R}$,

$$\begin{aligned} &\mathbb{P}_\theta\{\theta - \hat{\theta}_n \le t\} = \mathbb{P}_0\{-\hat{\theta}_n \le t\} = \mathbb{P}_0\{\tfrac{d}{d\theta} \textstyle\sum_{j=1}^n \log f(X_j - \theta)|_{-t} \ge 0\} \\ &= \mathbb{P}_0\{\textstyle\sum_{j=1}^n (f(X_j + t))^{-1} f'(X_j + t) \le 0\} \\ &= \int_{\mathbb{R}^n} I(\textstyle\sum_{j=1}^n (f(x_j + t))^{-1} f'(x_j + t) \le 0) \prod_{j=1}^n f(x_j)\, dx_1 \cdots dx_n \\ &= \int_{\mathbb{R}^n} I(\textstyle\sum_{j=1}^n (f(x_j))^{-1} f'(x_j) \le 0) \prod_{j=1}^n f(x_j - t)\, dx_1 \cdots dx_n \\ &= \int_{\mathbb{R}^n} I(\textstyle\sum_{j=1}^n (f(x_j))^{-1} f'(x_j) \le 0) \exp\left(\sum_{j=1}^n \log f(x_j - t)\right) dx_1 \cdots dx_n. \end{aligned}$$

Hence,

$$\begin{aligned} &h_n(t) = -\int_{\mathbb{R}^n} I(\textstyle\sum_{j=1}^n (f(x_j))^{-1} f'(x_j) \le 0) \exp\left(\sum_{j=1}^n \log f(x_j - t)\right) \\ &\times \left(\textstyle\sum_{j=1}^n (f(x_j - t))^{-1} f'(x_j - t)\right) dx_1 \cdots dx_n \\ &= -\int_{\mathbb{R}^n} I(\textstyle\sum_{j=1}^n (f(x_j))^{-1} f'(x_j) \le 0) \\ &\times \left(\textstyle\sum_{j=1}^n (f(x_j - t))^{-1} f'(x_j - t)\right) \prod_{j=1}^n f(x_j - t)\, dx_1 \cdots dx_n \\ &= -\int_{\mathbb{R}^n} I(\textstyle\sum_{j=1}^n (f(x_j + t))^{-1} f'(x_j + t) \le 0) \left(\sum_{j=1}^n (f(x_j))^{-1} f'(x_j)\right) \\ &\textstyle\prod_{j=1}^n f(x_j)\, dx_1 \cdots dx_n \\ &= -E_0[I(\textstyle\sum_{j=1}^n (f(X_j + t))^{-1} f'(X_j + t) \le 0) \sum_{j=1}^n (f(X_j))^{-1} f'(X_j)] \\ &= E_0[I(\textstyle\sum_{j=1}^n (f(X_j + t))^{-1} f'(X_j + t) > 0) \sum_{j=1}^n (f(X_j))^{-1} f'(X_j)], \end{aligned}$$

where we have used that $E_0[\sum_{j=1}^n (f(X_j))^{-1} f'(X_j)] = 0$.

Since $\log f$ is a strictly concave function, so is $\sum_{j=1}^n \log f(X_j + t)$, $t \in \mathbb{R}$. Hence, $\sum_{j=1}^n (f(X_j + t))^{-1} f'(X_j + t)$, $t \in \mathbb{R}$, is a decreasing function. Hence, for $t > s \ge 0$,

$$\sum_{j=1}^n (f(X_j + t))^{-1} f'(X_j + t) < \sum_{j=1}^n (f(X_j + s))^{-1} f'(X_j + s) \le \sum_{j=1}^n (f(X_j))^{-1} f'(X_j),$$

$$\left\{\sum_{j=1}^n (f(X_j + t))^{-1} f'(X_j + t) > 0\right\} \subset \left\{\sum_{j=1}^n (f(X_j))^{-1} f'(X_j) > 0\right\}$$

and

$$\begin{aligned} &E_0[I(\textstyle\sum_{j=1}^n (f(X_j + t))^{-1} f'(X_j + t) > 0) \sum_{j=1}^n (f(X_j))^{-1} f'(X_j)] \\ &< E_0[I(\textstyle\sum_{j=1}^n (f(X_j + s))^{-1} f'(X_j + s) > 0) \sum_{j=1}^n (f(X_j))^{-1} f'(X_j)]. \end{aligned}$$

Note that in the set $I(\sum_{j=1}^n (f(X_j+t))^{-1} f'(X_j+t) > 0)$, $\sum_{j=1}^n (f(X_j))^{-1} f'(X_j) > 0$. This implies that h_n is nonincreasing in $[0,\infty)$. A similar argument gives that h_n nondecreasing in $(-\infty, 0]$. $\square$

Proof of Theorem 3.8: We have that

$$\begin{aligned}
&C_{HL,L}(x_1,\ldots,x_n)\\
&= \{\theta \in \mathbb{R}^d : \textstyle\prod_{j=1}^n f(x_j-\theta) \ge k(x_1,\ldots,x_n)\}\\
&= \{\theta \in \mathbb{R}^d : \tau(g(T(x_1,\ldots,x_n),\theta), h(x_1,\ldots,x_n)) \ge k(x_1,\ldots,x_n)\}\\
&= \{\theta \in \mathbb{R}^d : g(T(x_1,\ldots,x_n),\theta) \in A(x_1,\ldots,x_n)\}\\
&= \{\theta \in \mathbb{R}^d : g(\eta^{-1}(\hat\theta_n),\theta) \in A(x_1,\ldots,x_n)\},
\end{aligned}$$

where

$$A(x_1,\ldots,x_n) := \{t \in \mathbb{R} : \tau(t, h(x_1,\ldots,x_n)) \ge k(x_1,\ldots,x_n)\}.$$

This implies that $C_{HL,L}(X_1,\ldots,X_n)$ is based on the m.l.e. $\square$

Proof of Theorem 3.9: We have that if $\theta - \hat\theta_n \in C_{piv,L}$, then $\lambda_{piv,L} \le h_n(\theta-\hat\theta_n)$. We also have that $\int_{x:\lambda_{piv,L}<h_n(x)} 1\,dx \le L$. Thus,

$$\int_{x:h_n(\theta-\hat\theta_n)<h_n(x)} 1\,dx \le \int_{x:\lambda_{piv,L}<h_n(x)} 1\,dx \le L.$$

Hence,

$$\begin{aligned}
&\mathbb{P}_\theta\{\theta-\hat\theta_n \in C_{piv,L}\} \le \mathbb{P}_\theta\{\textstyle\int_{x:h_n(\theta-\hat\theta_n)<h_n(x)} 1\,dx \le L\}\\
&\le \mathbb{P}_\theta\{\textstyle\int_{x:-n^{-1}\log h_n(x)<-n^{-1}\log h_n(\theta-\hat\theta_n)} 1\,dx \le L\}.
\end{aligned}$$

Let $k_n(x) = \int_{y\in\mathbb{R}:-n^{-1}\log h_n(y)<-n^{-1}\log h_n(-x)} 1\,dy$ and let
$k(x) = \int_{y\in\mathbb{R}:R(y)<R(-x)} 1\,dy$. By hypotheses (ii)–(iv), if $x_n \to x$, then, $k_n(x_n) \to k(x)$. Hence, by Arcones [1, Theorem 2.1], $k_n(\theta-\hat\theta_n)$ satisfies the LDP with speed n and rate function

$$\inf\{R(x) : k(-x) = t\}, t \in \mathbb{R}.$$

Therefore,

$$\begin{aligned}
&\liminf_{n\to\infty} n^{-1}\log(\mathbb{P}_\theta\{\theta-\hat\theta_n \notin C_{piv,L}(\hat\theta_n)\}\\
&\ge \liminf_{n\to\infty} n^{-1}\log(\mathbb{P}_\theta\{k_n(\hat\theta_n-\theta) > L\})\\
&\ge -\inf\{R(x) : k(-x) > L\} = -\inf\{R(x) : \textstyle\int_{y\in\mathbb{R}^d:R(y)<R(x)} 1\,dy > L\}\\
&= -\inf\{u \ge 0 : \textstyle\int_{y\in\mathbb{R}^d:R(y)\le u} 1\,dy \ge L\}.
\end{aligned}$$

We also have that if $h_n(\theta-\hat\theta_n) > 0$ and $\int_{y:n^{-1}(n-1)h_n(\theta-\hat\theta_n)\le h_n(y)} 1\,dy \le L$, then $\lambda_{piv,L} < h_n(\hat\theta_n-\theta)$, and $\theta-\hat\theta_n \in C_{piv,L}(\hat\theta_n)$. Therefore,

$$\begin{aligned}
&\mathbb{P}_\theta\{\textstyle\int_{y:n^{-1}(n-1)h_n(\theta-\hat\theta_n)\le h_n(y)} 1\,dy \le L\}\\
&\le \mathbb{P}_\theta\{\theta-\hat\theta_n \in C_{piv,L}(\hat\theta_n)\} + \mathbb{P}_\theta\{h_n(\theta-\hat\theta_n) = 0\}
\end{aligned}$$

So,

$$\begin{aligned}&\mathbb{P}_\theta\{\theta-\hat\theta_n \notin C_{piv,L}(\hat\theta_n)\} \\ \le\ &\mathbb{P}_\theta\{\textstyle\int_{y:n^{-1}(n-1)h_n(\theta-\hat\theta_n)\le h_n(y)} 1\,dy > L\} + \mathbb{P}_\theta\{h_n(\theta-\hat\theta_n)=0\}\end{aligned} \tag{4.17}$$

As before,

$$\begin{aligned}&\limsup_{n\to\infty} n^{-1}\log(\mathbb{P}_\theta\{\textstyle\int_{x:n^{-1}(n-1)h_n(\theta-\hat\theta_n)\le h_n(x)} 1\,dx > L\}) \\ \le\ &-\inf\{R(x) : \textstyle\int_{y\in\mathbb{R}^d:R(y)\le R(x)} 1\,dy > L\}.\end{aligned} \tag{4.18}$$

Given $\epsilon, M>0$, there exists n_0 such that for $n\ge n_0$,

$$\sup_{|x|\le M} |n^{-1}\log h_n(x) + R(-x)| \le \epsilon,$$

Hence, for each $n\ge n_0$ and each $|x|\le M$,

$$-\epsilon \le n^{-1}\log h_n(x) + R(-x) \le n^{-1}\log h_n(x) + a,$$

where $a := \sup_{|x|\le M} R(-x) < \infty$. Hence, for each $n\ge n_0$ and each $|x|\le M$,

$$0 < \exp(-n(\epsilon+a)) \le h_n(x).$$

Thus,

$$\begin{aligned}&\limsup_{n\to\infty} n^{-1}\log\Big(\mathbb{P}_\theta\{h_n(\theta-\hat\theta_n)=0\}\Big) \\ \le\ &\limsup_{n\to\infty} n^{-1}\log\Big(\mathbb{P}_\theta\{|\theta-\hat\theta_n|\ge M\}\Big) \\ \le\ &-\inf\{R(x) : |x|\ge M\}\end{aligned}$$

Letting $M\to\infty$, we get that

$$\limsup_{n\to\infty} n^{-1}\log\Big(\mathbb{P}_\theta\{h_n(\theta-\hat\theta_n)=0\}\Big) = -\infty. \tag{4.19}$$

By (4.17)–(4.19),

$$\begin{aligned}&\limsup_{n\to\infty} n^{-1}\log(\mathbb{P}_\theta\{\theta-\hat\theta_n\notin C_{piv,L}(\hat\theta_n)\}) \\ \le\ &-\inf\{R(x) : \textstyle\int_{y\in\mathbb{R}^d:R(y)\le R(x)} 1\,dy > L\} \\ =\ &-\inf\{u\ge 0 : \textstyle\int_{y\in\mathbb{R}^d:R(y)\le u} 1\,dy \ge L\}. \quad \square\end{aligned}$$

Lemma 4.4 *Under the conditions of Theorem 3.10,*

(i) For each $t\in\mathbb{R}$, $R(t)<\infty$.

(ii) $R(0)=0$

(iii) R *is increasing in* $[0,\infty)$ *and decreasing in* $(-\infty,0]$.

(iv) R *is continuous in* $\mathbb{R}$.

(v) For each $0<M<\infty$,

$$\sup_{|x|\le M} |n^{-1}\log h_n(x) + R(-x)| \to 0,$$

as $n\to\infty$.

(vi)

$$\lim_{M\to\infty}\limsup_{n\to\infty}\sup_{|x|\ge M} n^{-1}\log h_n(x) = -\infty.$$

Proof: (i) Let

$$\tau := \min(\lim_{x\to-\infty}(f(x))^{-1}f'(x), -\lim_{x\to\infty}(f(x))^{-1}f'(x)). \tag{4.20}$$

Let x_0 be such that $\sup_{x\in\mathbb{R}} f(x) = f(x_0)$. Since f is a p.d.f. there are $x_1 < x_0 < x_2$ such that $f(x_1), f(x_2) < f(x_0)$. Since $\log f(\cdot)$ is concave,

$$\lim_{x\to-\infty}(f(x))^{-1}f'(x) \geq (x_0 - x_1)^{-1}(f(x_0) - f(x_1)) > 0$$

and

$$\lim_{x\to\infty}(f(x))^{-1}f'(x) \leq (x_2 - x_0)^{-1}(f(x_2) - f(x_0)) < 0.$$

Hence, $\tau > 0$. So, by (3.8), $R(t) < \infty$.

(ii) holds because $E[(f(X))^{-1}f'(X)] = 0$.

(iii) Let $0 \leq s < t$, then $R(s) < R(t)$. Since $(f(x-u))^{-1}f'(x-u)$, $u \in \mathbb{R}$, is increasing $E_0[(f(X-s))^{-1}f'(X-s)] > 0$. By (3.7), there exists $\lambda_s < 0$ such that

$$R(s) = -\log E_0\left[\exp\left(\lambda_s(f(X-s))^{-1}f'(X-s)\right)\right].$$

Since $\int I(x \in \mathbb{R} : (f(X-s))^{-1}f'(X-s) < (f(X-t))^{-1}f'(X-t))\,dx > 0$,

$$\begin{aligned}&R(s) = -\log E_0\left[\exp\left(\lambda_s(f(X-s))^{-1}f'(X-s)\right)\right]\\ &< -\log E_0\left[\exp\left(\lambda_s(f(X-t))^{-1}f'(X-t)\right)\right]\\ &\leq -\inf_{\lambda\in\mathbb{R}}\log E_0\left[\exp\left(\lambda(f(X-t))^{-1}f'(X-t)\right)\right]\\ &= R(t).\end{aligned}$$

Hence, R is increasing in $[0,\infty)$. The proof that R is decreasing in $(-\infty, 0]$ is similar and it is omitted.

(iv) Suppose that $t_n \to t_0$. We show that $R(t_n) \to R(t_0)$. Then, there are $\lambda_n, \lambda_0 \in \mathbb{R}$ such that

$$R(t_n) = -\log E_0[\exp(\lambda_n(f(X-t_n))^{-1}f(X-t_n))]$$

and

$$R(t_0) = -\log E_0[\exp(\lambda_0(f(X-t_0))^{-1}f(X-t_0))].$$

Let $\eta := 2\sup_{n\geq 1}|t_n|$.

By (4.20), there exists $0 < M < \infty$ such that $M > 2\tau$; for $x < -M$, $(f(x))^{-1}f'(x) \geq 2^{-1}\tau$; and for $x \geq M$, $(f(x))^{-1}f'(x) \leq -2^{-1}\tau$. For $\lambda > 0$,

$$\begin{aligned}&E_0[\exp(\lambda(f(X-t))^{-1}f'(X-t))]\\ &\geq \int_{x\leq t-M}\exp(\lambda(f(x-t))^{-1}f'(x-t))f(x)\,dx\\ &\geq \exp(2^{-1}\lambda\tau)\mathbb{P}\{X \leq t-M\}.\end{aligned} \tag{4.21}$$

For $\lambda < 0$,

$$E_0[\exp(\lambda f(X-t))^{-1}f'(X-t))] \geq \exp(-2^{-1}\lambda\tau)\mathbb{P}\{X \geq t+M\}. \tag{4.22}$$

By (4.21), for $\lambda > 0$,

$$E_0[\exp(\lambda(f(X-t_n))^{-1}f'(X-t_n))] \geq \exp(2^{-1}\lambda\tau)\mathbb{P}\{X \leq -\tau - M\}.$$

By (4.22), for $\lambda < 0$,

$$E_0[\exp(\lambda(f(X-t_n))^{-1}f'(X-t_n))] \geq \exp(-2^{-1}\lambda\tau)\mathbb{P}\{X \geq -\tau + M\}.$$

Take $\lambda^* > 0$ such that

$$\min(\exp(2^{-1}\lambda^*\tau)\mathbb{P}\{X \leq -\tau - M\}, \exp(-2^{-1}\lambda^*\tau)\mathbb{P}\{X \geq -\tau + M\}) \geq 2.$$

Then, $\sup_{n\geq 1}|\lambda_n| \leq \lambda^*$. Take a subsequence n_k such that

$$\limsup_{n\to\infty} R(t_n) = \lim_{k\to\infty} R(t_{n_k})$$

and $\lim_{k\to\infty}\lambda_{n_k} = \bar{\lambda}$ exists. Since $(f(\cdot))^{-1}f'(\cdot)$ is nonincreasing,

$$\begin{aligned}&\exp(\lambda_{t_{n_k}}(f(X-t_{n_k})^{-1}f'(X-t_{n_k}))\\ \leq\ &\exp(\lambda^*|(f(X-2\eta))^{-1}f'(X-2\eta)|) + \exp(\lambda^*|(f(X+2\eta))^{-1}f'(X+2\eta)|).\end{aligned}$$

Hence, by the dominated convergence theorem

$$\lim_{k\to\infty} R(t_{n_k}) = -\log E_0[\exp(\bar{\lambda}(f(X-t_0))^{-1}f'(X-t_0))] \leq R(t_0).$$

So,

$$\limsup_{n\to\infty} R(t_n) \leq R(t_0).$$

We also have that

$$\begin{aligned}&\exp(\lambda_{t_0}(f(X-t_n))^{-1}f'(X-t_n))|\\ \leq\ &\exp(\lambda^*|(f(X-2\eta))^{-1}f'(X-2\eta)|) + \exp(\lambda^*|(f(X+2\eta))^{-1}f'(X+2\eta)|).\end{aligned}$$

Hence, by the dominated convergence theorem,

$$\begin{aligned}R(t_n) &\geq -\log E_0[\exp(\lambda_{t_0}(f(X-t_n))^{-1}f'(X-t_n))]\\ &\to -\log E[\exp(\lambda_{t_0}(f(X-t_0))^{-1}f'(X-t_0))] = R(t_0).\end{aligned}$$

Thus,

$$\liminf_{n\to\infty} R(t_n) \geq R(t_0).$$

(v) Since h_n is nonincreasing in $(0,\infty)$, for each $t > 0$ and each $\epsilon > 0$,

$$\mathbb{P}_\theta\{t+\epsilon \geq \theta - \hat{\theta}_n \geq t\} \leq \epsilon h_n(t).$$

Hence,

$$\begin{aligned}&-\inf\{R(-x) : x \in (t, t+\epsilon)\} \leq \liminf_{n\to\infty} n^{-1}\log(\mathbb{P}_\theta\{t+\epsilon \geq \theta - \hat{\theta}_n \geq t\})\\ \leq\ &\liminf_{n\to\infty} n^{-1}\log(h_n(t)).\end{aligned}$$

Letting $\epsilon \to 0$, we get that

$$-R(-t) \le \liminf_{n\to\infty} n^{-1} \log(h_n(t)).$$

Since a similar argument holds for the $\limsup$ (and for each $t < 0$), we have that for each $t \neq 0$,

$$-R(-t) = \lim_{n\to\infty} n^{-1} \log(h_n(t)).$$

By the CLT, $n^{-1/2} V^{-1/2} \sum_{j=1}^n (f(X_j))^{-1} f'(X_j) \xrightarrow{\mathrm{d}} Z_1$, where V is the Fisher information for the location family, i.e.

$$V := E\left[\left(\frac{\partial \log f(X-\theta)}{\partial\theta}\right)^2\right] = E[(f(X))^{-2}(f'(X))^2].$$

By uniform integrability,

$$\begin{aligned} &n^{-1/2}V^{-1/2}h_n(0) = n^{-1/2}V^{-1/2}E_0[I(\textstyle\sum_{j=1}^n (f(X_j))^{-1} f'(X_j) > 0) \\ &\times \textstyle\sum_{j=1}^n (f(X_j))^{-1} f'(X_j)] \\ \to\ &-E[I(Z_1 > 0)Z_1] = \phi(0). \end{aligned}$$

Note that

$$\begin{aligned} &E_0\left[\left(n^{-1/2} I(\textstyle\sum_{j=1}^n (f(X_j))^{-1} f'(X_j) > 0) \sum_{j=1}^n (f(X_j))^{-1} f'(X_j)\right)^2\right] \\ \le\ & n^{-1} E_0\left[\left(\textstyle\sum_{j=1}^n (f(X_j))^{-1} f'(X_j)\right)^2\right] = E_0\left[\left((f(X))^{-1} f'(X)\right)^2\right]. \end{aligned}$$

We have that

$$n^{-1} \log(h_n(0)) \to 0.$$

Hence, for each $t \in \mathbb{R}$,

$$-R(-t) = \lim_{n\to\infty} n^{-1} \log(h_n(t)).$$

Since h_n is decreasing in $[0,\infty)$ and increasing in $(-\infty, 0]$, for each $0 < M < \infty$,

$$\lim_{n\to\infty} \sup_{|t|\le M} |n^{-1}\log(h_n(t)) + R(-t)| = 0,$$

(vi) We have that

$$\begin{aligned} &\lim_{M\to\infty} \limsup_{n\to\infty} \sup_{t\le -M} n^{-1} \log(h_n(t)) \\ =\ &\lim_{M\to\infty} \limsup_{n\to\infty} n^{-1} \log(h_n(-M)) \\ =\ &\lim_{M\to\infty} -R(-M) = -\infty \end{aligned}$$

and

$$\lim_{M\to\infty} \limsup_{n\to\infty} \sup_{t\ge M} n^{-1} \log(h_n(t)) = -\infty$$

□

Proof of Theorem 3.10: It follows from Theorem 3.9 and Lemma 4.4. $\square$

Proof of Theorem 3.11: (i) We have that

$$R(b) = \inf\{J_0(l) : l \in (\mathcal{L}_0^{\Phi_1})^*, l((f'(\cdot - b))^{-1} f(\cdot - b)) = 0\}.$$

Let

$$\gamma_0(x) = (f(x-2b)/f(x))^{1/2}(E_0[(f(X-2b)/f(X))^{1/2}])^{-1}.$$

By (3.2),

$$\begin{aligned}
&J_0(l_{\gamma_0}) = E_0[\gamma_0(X)\log\gamma_0(X)] \\
=\ & E_0[(f(X-2b)/f(X))^{1/2}(E_0[(f(X-2b)/f(X))^{1/2}])^{-1} \\
&\times(2^{-1}\log(f(X-2b)/f(X)) - \log(E_0[(f(X-2b)/f(X))^{1/2}]))] \\
=\ & 2^{-1}(E_0[(f(X-2b)/f(X))])^{-1}E_0[(f(X-2b)/f(X))^{1/2}\log(f(X-2b)/f(X))] \\
&-\log(E_0[(f(X-2b)/f(X))^{1/2}]).
\end{aligned}$$

Using that f is even and the change of variables $-x+2b=y$,

$$\begin{aligned}
&E_0[(f(X-2b)/f(X))^{1/2}\log(f(X-2b)/f(X))] \\
=\ & \int_{\mathbb{R}} \log(f(x-2b)/f(x))(f(x-2b)f(x))^{1/2}\,dx \\
=\ & \int_{\mathbb{R}} \log(f(-x+2b)/f(-x))(f(-x+2b)f(-x))^{1/2}\,dx \\
=\ & \int_{\mathbb{R}} \log(f(y)/f(y-2b))(f(y)f(y-2b))^{1/2}\,dy \\
=\ & -\int_{\mathbb{R}} \log(f(x-2b)/f(x))(f(x-2b)f(x))^{1/2}\,dx = 0.
\end{aligned}$$

Hence,

$$J_0(l_{\gamma_0}) = -\log(E_0[(f(X-2b)/f(X))^{1/2}]).$$

By the Hölder inequality,

$$H_{2b}(\lambda) := \log\int_{\mathbb{R}}(f(x-2b))^{\lambda}(f(x))^{1-\lambda}\,dx, 0 \le \lambda \le 1,$$

is a convex function. Since f is an even function, we have that $H_{2b}(\lambda) = H_{2b}(1-\lambda)$, for each $0 \le \lambda \le 1$. Thus,

$$S(2b) = -\log H_{2b}(1/2) = -\log(E_0[(f(X-2b)/f(X))^{1/2}]) = J_0(l_{\gamma_0}).$$

We have that

$$\begin{aligned}
&l_{\gamma_0}(f'(\cdot - b)/f(\cdot - b)) \\
=\ & (E_0[(f(X-b)/f(X))^{1/2}])^{-1}\int_{\mathbb{R}}(f'(x-b)/f(x-b))(f(x-2b)f(x))^{1/2}\,dx.
\end{aligned}$$

Now,

$$\begin{aligned}
&\int_{\mathbb{R}}(f'(x-b)/f(x-b))(f(x-2b)f(x))^{1/2}\,dx \\
=\ & \int_{\mathbb{R}}(f'(x)/f(x))(f(x-b)f(x+b))^{1/2}\,dx \\
=\ & \int_{\mathbb{R}}(f'(-x)/f(-x))(f(-x-b)f(-x+b))^{1/2}\,dx \\
=\ & -\int_{\mathbb{R}}(f'(x)/f(x))(f(x+b)f(x-b))^{1/2}\,dx = 0.
\end{aligned}$$

Hence, $R(b) \le J_0(l_{\gamma_0}) = S(2b)$.

(ii) We need to prove that if $R(b) = S(2b)$, for each $b \in \mathbb{R}$, then f is either normal or one of the distributions in Example 3.14. By (3.7), there exists $\lambda(b) \in \mathbb{R}$ such that

$$R(b) = -\log E_0[\exp(\lambda(b)(f'(X-b)/f(X-b)))].$$

This implies that

$$0 = E_0[(f'(X-b)/f(X-b))\exp(\lambda(b)(f'(X-b)/f(X-b)))].$$

Let

$$\gamma_1(x) = (E_0[\exp(\lambda(b)(f'(X-b)/f(X-b)))])^{-1}\exp(\lambda(b)(f'(x-b)/f(x-b))).$$

Then,

$$l_{\gamma_1}(f'(\cdot-b)/f(\cdot-b)) = 0$$

and

$$R(b) = J_0(l_{\gamma_1}) = E_0[\gamma_1(X)\log\gamma_1(X)].$$

If $\gamma_0(x)$ were not equal to $\gamma_1(x)$ a.e. with respect to the Lebesgue measure, then for each $0 < \lambda < 1$

$$l_{\lambda\gamma_0+(1-\lambda)l_{\gamma_1}}(f'(\cdot-b)/f(\cdot-b)) = 0.$$

and

$$\begin{aligned} &J_0(l_{\lambda\gamma_0+(1-\lambda)\gamma_1}) = E_0[\Psi_2(\lambda\gamma_0(X)+(1-\lambda)\gamma_1(X))] \\ <\ &\lambda E_0[\Psi_2(\gamma_0(X))] + (1-\lambda)E_0[\Psi_2(\gamma_1(X))] = \lambda J_0(l_{\gamma_0}) + (1-\lambda)J_0(l_{\gamma_1}), \end{aligned}$$

in contradiction. Note that $\Psi_2(x) = x\log x$ is a strict convex function. Thus, $\gamma_0(x) = \gamma_1(x)$ a.e. with respect to the Lebesgue measure. Notice that $\gamma_0(x)$ and $\gamma_1(x)$ depend on b. So, we assume that for each $b \in \mathbb{R}$,

$$\begin{aligned} &(f(x-2b)/f(x))^{1/2}(E_0[(f(X-2b)/f(X)])^{1/2}])^{-1} \\ =\ &\exp(\lambda(b)(f'(x-b)/f(x-b)))(E_0[\exp(\lambda(b)(f'(X-b)/f(X-b)))])^{-1} \text{ a.e.} \end{aligned}$$

Using that f is continuous, we get that for each $x \in \mathbb{R}$,

$$\begin{aligned} &2^{-1}\log(f(x-2b)/f(x)) = \lambda(b)(f'(x-b)/f(x-b)) - S(2b) + R(b) \\ =\ &\lambda(b)(f'(x-b)/f(x-b)). \end{aligned}$$

Let $g(x) = \log f(x)$. We have that for each $x, b \in \mathbb{R}$,

$$g(x-2b) - g(x) = 2\lambda(b)g'(x-b).$$

Changing x into $x+b$, we have that for each $x, b \in \mathbb{R}$,

$$g(x-b) - g(x+b) = 2\lambda(b)g'(x).$$

Interchanging x and b, we get that

$$g(b-x) - g(x+b) = 2\lambda(x)g'(b).$$

Since g is even, we have that

$$2\lambda(b)g'(x) = 2\lambda(x)g'(b).$$

Since f is a p.d.f. there exists t such that $g'(t) \neq 0$. Then,

$$\lambda(x) = \frac{\lambda(t)g'(x)}{g'(t)} = 2^{-1}cg'(x),$$

where $c = \frac{2\lambda(t)}{g'(t)}$. Hence,

$$g(b-x) - g(x+b) = cg'(x)g'(b).$$

Taking two derivatives with respect to x and with respect to b, we get that

$$g''(b-x) - g''(x+b) = cg^{(3)}(x)g'(b)$$

and

$$g''(b-x) - g''(x+b) = cg'(x)g^{(3)}(b).$$

Hence,

$$g^{(3)}(x)g'(b) = g'(x)g^{(3)}(b).$$

If $g^{(3)}(x) = 0$, for each $x \in \mathbb{R}$, then $g(x) = ax^2 + b$, for some $a, b \in \mathbb{R}$. Hence, f is the p.d.f. of a normal distribution with mean zero. If $g^{(3)}(x_0) \neq 0$, for some $x_0 \in \mathbb{R}$, then

$$g^{(3)}(x) = \tau g'(x), \tag{4.23}$$

where $\tau = (g'(x_0))^{-1}g^{(3)}(x_0)$. The solutions of this differential equation (see e.g. Simmons [24, Section 17]) are

$$g'(x) = a_1 \sin(bx) + a_2 \cos(bx), \text{ if } \tau < 0,$$

and

$$g'(x) = a_1 \sinh(bx) + a_2 \cosh(bx), \text{ if } \tau > 0,$$

where $a_1, a_2 \in \mathbb{R}$ and $b^2 = |\tau|$. Since g' is an odd function and $e^{g(x)}$ is a p.d.f. The only solution is

$$f(x) = c\exp(-a\cosh(bx))$$

where $a, b > 0$ and $c = \left(\int_{\mathbb{R}} \exp(-a\cosh(bx))\,dx\right)^{-1}$. $\square$

Acknowledgement: I would like to referee for pointing out several typos.

References

[1] M. A. Arcones. The large deviation principle for stochastic processes I. *Theor. Probab. Applic.*, 47:567–583, 2003.

[2] M. A. Arcones. Confidence regions of fixed volume based on the likelihood ratio test. *Far East J. Theor. Statist.* 15:121–141, 2005.

[3] M. A. Arcones. Large deviations for M–estimators. *Ann. Instit. Statist. Mathem.* 58:21–52, 2006.

[4] R. R. Bahadur. Asymptotic efficiency of tests and estimates. *Shankhya.* 22:229–252, 1960.

[5] R. R. Bahadur. Rates of convergence of estimates and test statistics. *Ann. Mathem. Statist.* 38:303–324, 1967.

[6] R. R. Bahadur. *Some Limit Theorems in Statistics*. SIAM, Philadelphia, PA, 1971.

[7] R. R. Bahadur, S. L. Zabell, and J. C. Gupta. Large deviations, tests, and estimates. In *Asymptotic Theory of Statistical Tests and Estimation*, 33–64. Academic Press, New York, 1979.

[8] H. Chernoff. A measure of asymptotic efficiency for tests of hypothesis based on the sum of observations. *Ann. Math. Statist.* 23:493–507, 1952.

[9] A. Dembo, and O. Zeitouni. *Large Deviations Techniques and Applications. Second Edition.*. Springer, New York, 1998.

[10] J. D. Deuschel, and D. W. Stroock. *Large Deviations*. Academic Press, Inc., Boston, MA, 1989.

[11] T. S. Ferguson. Location and scale parameters in exponential families of distributions. *Ann. Mathemat. Statist.* 33:986–1001, 1962.

[12] J. C. Fu. On a theorem of Bahadur on the rate of convergence of point estimators. *Ann. Statist.* 1:745–749, 1973.

[13] J. C. Fu. The rate of convergence of consistent point estimators. *Ann. Statist.* 3 234–240, 1975.

[14] J. C. Fu. Large sample point estimation: a large deviation theory approach. *Ann. Statist.* 10:762–771, 1982.

[15] M. A. Girshick, and L. J. Savage. Bayes and minimax estimates for quadratic loss functions. *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability* 53–73. University of California Press, Berkeley and Los Angeles, 1951.

[16] W. C. Guenther. Shortest confidence intervals. *Amer. Statist.* 33:22–25, 1969.

[17] V. M. Joshi. Admissibility of invariant confidence procedures for estimating a location parameter. *Ann. Math. Statist.* 41 1568–1581, 1970.

[18] R. C. Juola. More on shortest confidence intervals. *Amer. Statist.* 47:117–119, 1993.

[19] H. Kudō. On minimax invariant estimates of the transformation parameter. *Nat. Sci. Rep. Ochanomizu Univ.* 6:31–73, 1955.

[20] E. J. G. Pitman. The estimation of the location and scale parameters of a continuous population of any given form. *Biomet.* 30:391–421, 1939.

[21] H. L. Royden. *Real Analysis*. MacMillan Publishing Comp. New York, 1968.

[22] J. Shao. *Mathematical Statistics. 2nd Edition.* Springer–Verlag, New York, 2003.

[23] G. L. Sievers. Estimates of location: a large deviation comparison. *Ann. Statist.* 6:610–618, 1978.

[24] G. F. Simmons. *Differential Equations with Applications and Historical Notes*. McGraw–Hill Book Co., New York, 1972.

[25] R. S. Valand. Invariant interval estimation of a location parameter. *Ann. Math. Statist.* 39:193–199, 1968.

Miguel A. Arcones
Department of Mathematical Sciences
Binghamton University
Binghamton, NY 13902, USA
arcones@math.binghamton.edu