IMS Collections
High Dimensional Probability V: The Luminy Volume
Vol. 5 (2009) 293–307

DOI: 10.1214/09-IMSCOLL519

Asymptotic distribution of the most powerful invariant test for invariant families

Miguel A. Arcones[1]

Binghamton University

Abstract: We obtain the limit distribution of the test statistic of the most powerful invariant test for location families of densities. As an application, we obtain the consistency of this test. From these results similar results are obtained for the test statistic of the most powerful invariant test for scale families.

Contents

1. Introduction

A statistical problem which has received some attention in the literature is that of testing two separate composite hypothesis (for a bibliography on this topic see Pereira [16] and Pereira [17]). The meaning of separate is that no member of the null hypothesis can be obtained as a limit of a sequence of distributions in the alternative hypothesis. A natural test for this problem is the likelihood ratio test. Cox [5, 6] proposes a variation of likelihood ratio test for this problem. Lindley [13] using a Bayesian approach proposes to use the ratio of the posterior likelihoods. Atkinson [3] considers the case of discrimating between several separate families of distributions. White [19] obtains the asymptotic distribution of the tests in Cox [5, 6]. Loh [14] presents a modification of Cox's which has asymptotically the correct level. Several other methods have being proposed by different authors. Pace and Salvan [15] find the best estimator based on a sufficient statistic of the null hypothesis.

For invariant separate families of distributions, it is possible to find the most powerful invariant test. This test is based on the maximal invariant. In this paper, we obtain the asymptotic distribution of the most powerful invariant test for separate families of distributions, when the possible types of distributions are location or scale families.

[1]Department of Mathematical Sciences, Binghamton University, Binghamton, NY 13902, USA, e-mail: arcones@math.binghamton.edu

Keywords and phrases: invariant tests, separate families, most powerful test.

AMS 2000 subject classifications: Primary 62F05, 60F03; secondary 62E20.

First, we look at testing two location families of distributions. Let f_0 and f_1 be two densities in $\Re^d$. We consider the testing problem

$$H_0 : \text{the data comes from } \{f_0(\cdot - t) : t \in \Re^d\}$$

versus

$$H_a : \text{the data comes from } \{f_1(\cdot - t) : t \in \Re^d\}.$$

In order that this problem makes sense, we assume that $f_1 \notin \{f_0(\cdot - t) : t \in \Re^d\}$, i.e. there exists no $t \in \Re^d$ such that $f_1(x) = f_0(x - t)$ almost everywhere with respect to Lebesgue measure. It is well known that the most powerful invariant test for this problem rejects H_0 for large values of the statistic

$$T_{n,\text{loc}} := \frac{\int_{\Re^d} \prod_{j=1}^n f_1(X_j - t)\, dt}{\int_{\Re^d} \prod_{j=1}^n f_0(X_j - t)\, dt} \tag{1.1}$$

(see Lehmann and Romano [12, Example 6.3.2]). In Theorem 2.2, we obtain the asymptotic distribution of $T_{n,\text{loc}}$.

A variation of the previous problem consists in dealing with scale families. Let f_0 and f_1 be two densities in $(0, \infty)$. We analyze the testing problem

$$H_0 : \text{the data comes from } \{\lambda^{-1} f_0(\lambda^{-1}\cdot) : \lambda > 0\}$$

versus

$$H_a : \text{the data comes from } \{\lambda^{-1} f_1(\lambda^{-1}\cdot) : \lambda > 0\}.$$

We assume that $f_1 \notin \{\lambda^{-1} f_0(\lambda^{-1}\cdot) : \lambda > 0\}$. The most powerful invariant test for this problem rejects H_0 for large values of the statistic

$$T_{n,scale} := \frac{\int_0^\infty \left(\prod_{j=1}^n \lambda^{-1} f_1(\lambda^{-1} X_j)\right) \lambda^{-1}\, d\lambda}{\int_0^\infty \left(\prod_{j=1}^n \lambda^{-1} f_0(\lambda^{-1} X_j)\right) \lambda^{-1}\, d\lambda} \tag{1.2}$$

(see Lehmann and Romano [12, Problem 6.12]).

The goal of this manuscript is to obtain the asymptotic distribution of the statistics in (1.1) and (1.2). The tests in (1.1) and (1.2) are a particular case of the Bayesian tests in Lindley [13]. For these tests, Lindley [13, page 456] obtains the asymptotic distribution without making explicit the required regularity conditions. The goal of this article is to show that concavity and mild smooth conditions suffice to give the expansion obtained in Lindley [13, page 456] for the test in (1.1) and (1.2). Ducharme and Frichot [8, Theorem 1] claim that the expansion in Lindley [13] holds under no conditions by using Theorem 6.2 in Barndorff–Nielsen and Cox [4]. However, the asymptotic result in Theorem 6.2 in Barndorff–Nielsen and Cox [4] does not apply neither to Theorem 1 Ducharme and Frichot [8] nor to (1.1). Theorem 6.2 in Barndorff–Nielsen and Cox [4] gives an expansion for $\int_D g(x)(f(x))^n\, dx$, where f and and g are nonstochastic fixed functions. This is not the situation in (1.1) and (1.2).

In Section 2, we present the main results. We show that by using concavity, only mild additional smoothness conditions are needed to obtain the aymptotic expansions of certain integrals. This approach applies to location families as well as scale families. This approach does not apply to location and scales families. Concavity (or convexity) is an assumption used by many authors (see for example Daniels [7]; Andersen and Gill [1], Haberman [9]; Hjort and Pollard, [11]; Arcones [2]). Section 3 contains the proofs of the results in Section 2.

In the proof of Theorem 2.1, we use concavity to bound certain stochastic processes defined on $\Re^d$. One of the properties of concave functions which we will use is that a sequence of pointwise converging convex functions converges uniformly on compact sets (see for example Rockafellar [18, Theorem 10.8]). We will use the stochastic version of this theorem given by Andersen and Gill [1, Theorem II.1]. Andersen and Gill [1] show that if the finite dimensional distributions of sequence of convex stochastic processes converge in probability, then the convergence holds in probability uniformly over compact sets. We also use concavity in the proof of Theorem 2.1 (see (3.6)) to bound a concave stochastic process outside a compact set.

We will use the usual multivariate notation. For example, given $u = (u_1, \dots, u_d)' \in \Re^d$, $|u| = (\sum_{j=1}^d u_j^2)^{1/2}$. Given a $d \times d$ matrix A, we denote

$$|A| := \sup_{|u|,|v|\le 1} u'Av.$$

I_d denotes the $d \times d$ identity matrix. c_d denotes the Lebesgue measure of the unit ball of $\Re^d$. c will denote a constant which may vary from occurrence to occurrence.

2. Main results

We will derive our results from the following theorem, which is a kind of delta method for integrals of empirical processes.

Theorem 2.1. *Let $\{X_j\}_{j=1}^\infty$ be a sequence of $\Re^d$–valued i.i.d. r.v.'s with values in a measurable space $(S, \mathcal{S})$. Let X be a copy of X_1. Let Θ be a Borel subset of $\Re^d$. Let $g : S \times \Theta \to \Re$ be a function such that $g(\cdot, \theta) : S \to \Re$ is measurable for each $\theta \in \Theta$. Let $\phi : S \to \Re^d$ be a Borel measurable function. Let θ_0 be a point in the interior of Θ. Let $\{a_n\}_{n=1}^\infty$ be a sequence of real numbers which converges to infinity. Let $\{b_n\}_{n=1}^\infty$ be a sequence of real numbers. Suppose that:*

(i) For each $x \in S$, $g(x, \cdot) : \Theta \to \Re$ is a concave function.

(ii) There exists a $d \times d$ positive definite symmetric matrix W such that

$$E[g(X,\theta) - g(X,\theta_0)] = -2^{-1}(\theta - \theta_0)'W(\theta - \theta_0) + o(|\theta - \theta_0|^2),$$

as $\theta \to \theta_0$.

(iii) For each $\theta \in \Re^d$,

$$a_n^2 E\left[\min\left(|r(X, a_n^{-1}\theta) - b_n|, |r(X, a_n^{-1}\theta) - b_n|^2\right)\right] \to 0,$$

where

$$r(x,\theta) := g(x, \theta_0 + \theta) - g(x,\theta_0) - \theta'\phi(x).$$

(iv) $a_n n^{-1} \sum_{i=1}^n (\phi(X_i) - E[\phi(X_i)]) = O_{\Pr}(1)$.

Then,

$$a_n^d \int_\Theta \exp\left(a_n^2 n^{-1} \sum_{j=1}^n (g(X_j,\theta) - g(X_j,\theta_0))\right) d\theta - e^{2^{-1}Z_n' W^{-1} Z_n} \det(W^{-1/2})(2\pi)^{d/2} \xrightarrow{\Pr} 0, \tag{2.1}$$

where $Z_n := a_n n^{-1} \sum_{i=1}^n (\phi(X_i) - E[\phi(X_i)])$.

If Θ is a convex set, conditions (i) and (ii) in Theorem 2.1 imply that $E[g(X,\theta)]$, $\theta \in \Theta$, has a maximum at θ_0.

If for almost all x $\theta \mapsto g(x;\theta)$ is differentiable with respect to θ at θ_0, then $\phi(x)$ is the partial derivative $\dot{g}(x,\theta_0)$ of $g(x;\theta)$ with respect to θ at θ_0.

Next two propositions show that differentiability conditions imply condition (iii) in Theorem 2.1.

Proposition 2.1. *Suppose that:*

(i) $a_n = n^{1/2}$.

(ii) For each $x \in S$, $g(x,\cdot)$ *is differentiable with respect to* θ *in a neighborhood of* θ_0.

(iii)

$$\lim_{\delta \to 0+} E\left[\sup_{|\theta-\theta_0|\le\delta} |\dot{g}(X,\theta) - \dot{g}(X,\theta_0)|^2\right] = 0$$

where $\dot{g}(x,\theta)$ *is the vector of partial derivatives of* $g(x,\cdot)$ *at* θ_0.

Then, (iii) in Theorem 2.1 holds with $\phi(\cdot) = \dot{g}(\cdot,\theta_0)$.

Proposition 2.2. *Suppose that:*

(i) $a_n = n^{1/2}$.

(ii) For each $x \in S$, $g(x,\cdot)$ *is twice differentiable with respect to* θ *in a neighborhood of* θ_0.

(iii)

$$\lim_{\delta \to 0+} E\left[\sup_{|\theta-\theta_0|\le\delta} |\ddot{g}(X,\theta) - \ddot{g}(X,\theta_0)|\right] = 0$$

where $\ddot{g}(x,\theta)$ *is the* $d \times d$ *matrix of second partial derivatives of* $g(x,\cdot)$ *at* θ_0.

(iv) $E\left[|\ddot{g}(X,\theta_0)|\right] < \infty$.

(v) $E\left[\dot{g}(X,\theta_0)\right] = 0$.

Then, (ii) and (iii) in Theorem 2.1 hold with $W = -E\left[\ddot{g}(X,\theta)\right]$ *and* $\phi(\cdot) = \dot{g}(\cdot,\theta_0)$.

Out next theorem gives the asymptotic distribution of $T_{n,\mathrm{loc}}$ for an arbitrary sampling distribution.

Theorem 2.2. *Let* f_0 *and* f_1 *be two densities in* $\Re^d$ *such that* $f_0(x), f_1(x) > 0$, *for each* $x \in \Re^d$. *Let* $\{X_n\}_{n=1}^\infty$ *be a sequence of* $\Re^d$*–valued i.i.d. r.v.'s. Let* X *be a copy of* X_1. *Suppose that:*

(i) There exists $t_0 \in \Re^d$ *such that*

$$E[\log f_0(X - t_0)] = \sup_{t\in\Re^d} E[\log f_0(X - t)].$$

(ii) There are a measurable function $\phi_0 : \Re^d \to \Re$ *and a* $d \times d$ *matrix* W_0 *such that* $g_0(x,\theta) := \log f_0(x-\theta)$, $x,\theta \in \Re^d$, *such that the conditions in Theorem 2.1 are satisfied for* $g \equiv g_0$, $\theta_0 \equiv t_0$, $\phi \equiv \phi_0$, $W \equiv W_0$ *and* $a_n \equiv n^{1/2}$.

(iii) There exists $t_1 \in \Re^d$ *such that*

$$E[\log f_1(X - t_1)] = \sup_{t\in\Re^d} E[\log f_1(X - t)].$$

(iv) There are a measurable function $\phi_1 : \Re^d \to \Re$ *and a* $d \times d$ *matrix* W_1 *such that* $g_1(x,\theta) := \log f_1(x-\theta)$, $x,\theta \in \Re^d$, *such that the conditions in Theorem 2.1 are satisfied for* $g \equiv g_1$, $\theta_0 \equiv t_1$, $\phi \equiv \phi_1$, $W \equiv W_1$ *and* $a_n \equiv n^{1/2}$.

Then,

$$\begin{aligned}
&\log(T_{n,\mathrm{loc}}) - \textstyle\sum_{j=1}^n \log f_1(X_j - t_1) - \log\left(\det(W_1^{-1/2})\right) - 2^{-1}|W_1^{-1/2} Z_{n,1}|^2 \\
&+ \textstyle\sum_{j=1}^n \log f_0(X_j - t_0) + \log\left(\det(W_0^{-1/2})\right) + 2^{-1}|W_0^{-1/2} Z_{n,0}|^2 \\
&\overset{\mathrm{Pr}}{\to} 0,
\end{aligned} \tag{2.2}$$

where

$$Z_{n,0} := n^{-1/2} \sum_{i=1}^n (\phi_0(X_i) - E[\phi_0(X_i)])$$

and

$$Z_{n,1} := n^{-1/2} \sum_{i=1}^n (\phi_1(X_i) - E[\phi_1(X_i)]).$$

Consequently,

$$\begin{aligned}
&n^{1/2}\left(n^{-1}\log(T_{n,\mathrm{loc}}) - E[\log(f_1(X_1 - t_1)/f_0(X_1 - t_0))]\right) \\
&\overset{\mathrm{d}}{\to} N(0, \mathrm{Var}(\log(f_1(X_1 - t_1)/f_0(X_1 - t_0)))).
\end{aligned} \tag{2.3}$$

If f_j, $j = 1, 2$, is differentiable with derivative $\dot{f}_j$, then $\phi_j(x) = (f_j(x - t_j))^{-1} \times \dot{f}_j(x - t_j)$, $x \in \Re^d$.

It is well known that by Jensen's inequality, if a $\Re^d$–valued r.v. X has density f, then for any other density g,

$$E[\log g(X)] \le E[\log f(X)], \tag{2.4}$$

with equality only when f and g are equal almost everywhere with respect to the Lebesgue measure. Notice that

$$K(P, Q) := E[\log f(X)] - E[\log g(X)] = E[\log(f(X)/g(X))]$$

is the Kullback–Leibler divergence between the measures P and Q in $\Re^d$ having respective densities f and g. Equation (2.4) implies that for each $t_j \in \Re^d$, $j = 1, 2$,

$$E_{f_j,t_j}[\log f_j(X - t_j)] = \sup_{t \in \Re^d} E_{f_j,t_j}[\log f_j(X - t)], \tag{2.5}$$

where E_{f_j,t_j} is expectation with respect to the probability measure for which the r.v. X have p.d.f. $f_j(\cdot - t_j)$.

Theorem 2.2 gives the limit in probability of $n^{-1}\log(T_{n,\mathrm{loc}})$ both under the null and alternative hypothesis. Suppose that the conditions in Theorem 2.2 hold when the sampling is distribution is either the null or alternative hypothesis. Then, for each $t_j \in \Re^d$, $j = 1, 2$,

$$n^{-1}\log(T_{n,\mathrm{loc}}) \overset{P_{f_j,t_j}}{\longrightarrow} \sup_{t \in \Re^d} E_{f_j,t_j}[\log f_1(X - t)] - \sup_{t \in \Re^d} E_{f_j,t_j}[\log f_0(X - t)],$$

where P_{f_j,t_j} is the probability measure when the r.v.'s $X_1, \ldots, X_n$ have p.d.f. $f_j(\cdot - t_j)$. By (2.5) and (2.4)

$$\begin{aligned}
&\sup_{t \in \Re^d} E_{f_0,t_0}[\log f_1(X - t)] - \sup_{t \in \Re^d} E_{f_0,t_0}[\log f_0(X - t)] \\
&= \sup_{t \in \Re^d} E_{f_0,t_0}[\log f_1(X - t)] - E_{f_0,t_0}[\log f_0(X - t_0)] \\
&= -\inf_{t \in \Re^d} K(P_{f_0,t_0}, P_{f_1,t}) < 0
\end{aligned}$$

and

$$\begin{aligned}&\sup_{t\in\Re^d} E_{f_1,t_1}[\log f_1(X-t)] - \sup_{t\in\Re^d} E_{f_1,t_1}[\log f_0(X-t)]\\ =&E_{f_1,t_1}[\log f_1(X-t_1)] - \sup_{t\in\Re^d} E_{f_1,t_1}[\log f_0(X-t)]\\ =&\inf_{t\in\Re^d} K(P_{f_1,t_1},P_{f_0,t}) > 0.\end{aligned}$$

A level α test is obtained as follows. Let

$$a_{n,\alpha} := \inf\{\lambda \ge 0 : P_{f_0,t_0}\{n^{-1}\log(T_{n,\mathrm{loc}}) < \lambda\} \ge 1-\alpha\}. \tag{2.6}$$

A level α test rejects the null hypothesis if $n^{-1}\log(T_{n,\mathrm{loc}}) \ge a_{n,\alpha}$. We have that $P_{f_0,t_0}\{n^{-1}\log(T_{n,\mathrm{loc}}) \ge a_{n,\alpha}\} \le \alpha$, i.e. the type I error of the test is less or equal than α.

If the conditions in Theorem 2.2 hold when the sampling is distribution is either f_0 or f_1, then for each $t_1 \in \Re^d$,

$$\lim_{n\to\infty} a_{n,\alpha} = -\inf_{t\in\Re^d} K(P_{f_0,t_0},P_{f_1,t}) < 0$$

and

$$n^{-1}\log(T_{n,\mathrm{loc}}) \overset{P_{f_1,t_1}}{\longrightarrow} \inf_{t\in\Re^d} K(P_{f_1,t_1},P_{f_0,t}) > 0.$$

Hence, the test is consistent:

Corollary 2.1. *Suppose that the conditions in Theorem 2.2 hold when the sampling is distribution is either f_0 or f_1. Then, for each $1 > \alpha > 0$ and each $t_1 \in \Re^d$,*

$$\lim_{n\to\infty} P_{f_1,t_1}\{n^{-1}\log(T_{n,\mathrm{loc}}) \ge a_{n,\alpha}\} = 1.$$

Another application of Theorem 2.2 is to obtain the asymptotic distribution of $T_{n,\mathrm{loc}}$ under the null hypothesis. This can used to approach $a_{n,\alpha}$ by

$$\begin{aligned}&E_{f_0,t_0}[\log(f_1(X_1-t_1)/f_0(X_1-t_0))]\\ &+n^{-1/2}z_\alpha\left(\mathrm{Var}_{f_0,t_0}(\log(f_1(X_1-t_1)/f_0(X_1-t_0)))\right)^{1/2},\end{aligned}$$

where $P\{N(0,1) \ge z_\alpha\} = \alpha$ and $t_1 \in \Re^d$ satisfies

$$E_{f_0,t_0}[\log f_1(X-t_1)] = \sup_{t\in\Re^d} E_{f_0,t_0}[\log f_1(X-t)].$$

By the lemma of Neyman–Pearson, the most powerful test for $H_0 : f_0(\cdot - t_0)$ versus $H_a : f_1(\cdot - t)$, rejects for large values of the statistic

$$T_{n,\mathrm{loc},t} := \frac{\prod_{j=1}^n f_1(X_j - t)}{\prod_{j=1}^n f_0(X_j - t_0)}.$$

Under regularity conditions, by the law of the large numbers,

$$n^{-1}\log(T_{n,\mathrm{loc},t}) \overset{P_{f_0,t_0}}{\longrightarrow} E_{f_0,t_0}[\log f_1(X_1-t) - \log f_0(X_1-t_0)] = -K(P_{f_0,t_0},P_{f_1,t}).$$

By (2.4) $-K(P_{f_0,t_0},P_{f_1,t}) < 0$. The bigger $-K(P_{f_0,t_0},P_{f_1,t})$ becomes, the more difficult is to differentiate between the two distributions P_{f_0,t_0} and $P_{f_1,t}$. We have that

$$\sup_{t\in\Re^d}\left(-K(P_{f_0,t_0},P_{f_1,t})\right) = -K(P_{f_0,t_0},P_{f_1,t_1}).$$

The value of t for which $-K(P_{f_0,t_0},P_{f_1,t})$ is biggest is t_1. Theorem 2.2 shows that under the null hypothesis $T_{n,\mathrm{loc}}$ behaves asymptotically as T_{n,loc,t_1}.

Next corollary shows an application of Theorem 2.2.

Corollary 2.2. *Let $\{X_n\}_{n=1}^\infty$ be a sequence of $\Re^d$–valued i.i.d. r.v.'s with density $f_0(\cdot - t_0)$, where $t_0 \in \Re^d$. Suppose that:*

(i) $f_0(x) = \gamma_{p_0} a_0^d e^{-a_0^{p_0}|x|^{p_0}}$, $x \in \Re^d$, where $a_0 > 0$, $p_0 \geq 1$ and $\gamma_{p_0} = (\int_{\Re^d} e^{-|x|^{p_0}}\, dx)^{-1}$.

(ii) $f_1(x) = \gamma_{p_1} a_1^d e^{-a_1^{p_1}|x|^{p_1}}$, $x \in \Re^d$, where $a_1 > 0$, $p_1 \geq 1$ and $\gamma_{p_1} = (\int_{\Re^d} e^{-|x|^{p_1}}\, dx)^{-1}$.

(iii) $(a_0, p_0) \neq (a_1, p_1)$.

Then, (2.2) and (2.3) hold with $\theta_0 = \theta_1 = t_0$.

Next, we consider the case of scale families. We have that if X has density $\lambda^{-1} f(\lambda^{-1} x)$, $x \geq 0$, then $Y = \log X$ has density $\lambda^{-1} f(\lambda^{-1} e^y) e^y = f(e^{y - \log\lambda}) \times e^{y-\log\lambda}$, $y \in \Re$. This transformation changes a scale family into a location family.

Theorem 2.3. *Let f_0 and f_1 be two densities in $(0,\infty)$ such that $f_0(x), f_1(x) > 0$, for each $x \in (0,\infty)$. Let $\{X_n\}_{n=1}^\infty$ be a sequence of $(0,\infty)$–valued i.i.d. r.v.'s with density $\lambda_0^{-1} f(\lambda_0^{-1}\cdot)$, where $\lambda_0^{-1} \in (0,\infty)$. Suppose that:*

(i) There exists $\lambda_0 > 0$ such that

$$E[\log \lambda_0^{-1} f_0(\lambda_0^{-1} X_1)] = \sup_{\lambda>0} E[\log \lambda^{-1} f_0(\lambda^{-1} X_1)].$$

(ii) There are a measurable function $\phi_0 : \Re \to \Re$ and $W_0 > 0$ such that $g_0(x,\theta) := \log(f_0(e^{x-\theta})e^{x-\theta})$ such that the conditions in Theorem 2.1 are satisfied for $\theta_0 \equiv \log\lambda_0$, $\phi \equiv \phi_0$, $W \equiv W_0$, $a_n \equiv n^{1/2}$ and $X_j \equiv \log X_j$.

(iii) There exists $\lambda_1 > 0$ such that

$$E[\log \lambda_1^{-1} f_1(\lambda_1^{-1} X_1)] = \sup_{\lambda>0} E[\log \lambda^{-1} f_1(\lambda^{-1} X_1)].$$

(iv) There are a measurable function $\phi_1 : \Re \to \Re$ and $W_1 > 0$ such that $g_1(x,\theta) := \log(f_1(e^{x-\theta})e^{x-\theta})$ such that the conditions in Theorem 2.1 are satisfied for $\theta_0 \equiv \log\lambda_1$, $\phi \equiv \phi_1$, $W \equiv W_1$, $a_n \equiv n^{1/2}$ and $X_j \equiv \log X_j$.

Then,

$$\begin{aligned} &\log(T_{n,\mathrm{scale}}) - \textstyle\sum_{j=1}^n \log(\lambda_1^{-1} f_1(\lambda_1^{-1} X_j)) - \log\left(W_1^{-1/2}\right) - 2^{-1}|W_g^{-1/2} Z_{n,1}|^2 \\ &\quad + \textstyle\sum_{j=1}^n \log(\lambda_0^{-1} f_0(\lambda_0^{-1} X_j)) + \log\left(W_0^{-1/2}\right) + 2^{-1}|W_0^{-1/2} Z_{n,0}|^2 \\ &\quad \xrightarrow{\Pr} 0, \end{aligned} \tag{2.7}$$

where

$$Z_{n,0} := n^{-1/2} \sum_{i=1}^n (\phi_0(\log X_i) - E[\phi_0(\log X_i)]),$$
$$\text{and } Z_{n,1} := n^{-1/2} \sum_{i=1}^n (\phi_1(\log X_i) - E[\phi_1(\log X_i)]).$$

Consequently,

$$\begin{aligned} &n^{1/2}\left(n^{-1}\log(T_{n,\mathrm{scale}}) - E[\log(\lambda_1^{-1} f_1(\lambda_1^{-1} X_1))/\lambda_0^{-1} f_0(\lambda_0^{-1} X_1))]\right) \\ \xrightarrow{\mathrm{d}} &\quad N(0, \mathrm{Var}(\log((\lambda_1^{-1} f_1(\lambda_1^{-1} X_1))/\lambda_0^{-1} f_0(\lambda_0^{-1} X_1))). \end{aligned} \tag{2.8}$$

Corollary 2.3. *Let $\{X_n\}_{n=1}^{\infty}$ be a sequence of $(0,\infty)$–valued i.i.d. r.v.'s with density $\lambda_0^{-1} f(\lambda_0^{-1}\cdot)$, where $\lambda_0 > 0$. Suppose that:*

(i) $f_0(x) = \frac{x^{\alpha_0-1}e^{-x}}{\Gamma(\alpha_0)}$, $x > 0$, where $\alpha_0 > 0$.

(ii) $f_1(x) = \frac{x^{\alpha_1-1}e^{-x}}{\Gamma(\alpha_1)}$, $x > 0$, where $\alpha_1 > 0$.

(iii) $\alpha_0 \neq \alpha_1$.

Then, (2.7) and (2.8) hold with $\lambda_1 = \lambda_0\alpha_0\alpha_1^{-1}$.

3. Proofs

We will need the following lemmas:

Lemma 3.1. *Let $Y_{n,1}, \dots, Y_{n,n}$ be i.i.d. r.v.'s. Let $\{b_n\}_{n=1}^{\infty}$ be a sequence of real numbers. If*

$$\lim_{n\to\infty} nE[\min(|Y_{n,1} - b_n|, (Y_{n,1} - b_n)^2)] = 0,$$

then

$$\sum_{j=1}^{n}(Y_{n,j} - E[Y_{n,j}]) \xrightarrow{\Pr} 0.$$

Proof. Without loss of generality, we may assume that $b_n = 0$. The claim follows noticing that

$$\left|E\left[\sum_{j=1}^{n}(Y_{n,j}I(|Y_{n,j}| \le 1) - E[Y_{n,j}I(|Y_{n,j}| \le 1)])\right]\right| \le 2nE[|Y_{n,1}|I(|Y_{n,1}| \le 1) \\ \le 2nE[\min(|Y_{n,1}|, Y_{n,1}^2)]$$

and

$$\mathrm{Var}\left(\sum_{j=1}^{n}(Y_{n,j}I(|Y_{n,j}| > 1) - E[Y_{n,j}I(|Y_{n,j}| > 1)])\right) = n\mathrm{Var}(Y_{n,1}I(|Y_{n,1}| > 1)) \\ \le nE[Y_{n,1}^2 I(|Y_{n,1}| > 1)] \le nE[\min(|Y_{n,1}|, Y_{n,1}^2)].$$ □

Proof of Theorem 2.1. By a change of variables, we have that

$$a_n^d \int_{\Theta} \exp\left(a_n^2 n^{-1}\sum_{j=1}^{n}(g(X_j,\theta) - g(X_j,\theta_0))\right) d\theta \\ = \int_{\theta_0 + a_n^{-1}t\in\Theta} \exp\left(a_n^2 n^{-1}\sum_{j=1}^{n}(g(X_j,\theta_0 + a_n^{-1}t) - g(X_j,\theta_0))\right) dt.$$

Hence, we would like to prove that

$$\int_{\theta_0 + a_n^{-1}t\in\Theta} e^{U_n(t)}\,dt - \int_{\Re^d} e^{V_n(t)}\,dt \xrightarrow{\Pr} 0, \tag{3.1}$$

where $U_n(t) := a_n^2 n^{-1}\sum_{j=1}^{n}(g(X_j,\theta_0 + a_n^{-1}t) - g(X_j,\theta_0))$, $t \in \Re^d$, and $V_n(t) := t'Z_n - 2^{-1}t'Wt$, $t \in \Re^d$.

By Lemma 3.1 and hypothesis (v), for each $t \in \Re^d$,

$$\sum_{j=1}^{n}(r(X_j, a_n^{-1}t) - E[r(X_j, a_n^{-1}t)]) \xrightarrow{\Pr} 0.$$

By hypothesis (ii),

$$E\left[a_n^2 n^{-1}\sum_{j=1}^{n}(g(X_j,\theta_0+a_n^{-1}t)-g(X_j,\theta_0))\right]=E[a_n^2(g(X_1,\theta_0+a_n^{-1}t)-g(X_1,\theta_0))]$$
$$\longrightarrow -2^{-1}t'Wt,$$

as $n\to\infty$. Hence, for each $t\in\Re^d$

$$\begin{aligned}(3.2)\qquad & U_n(t)-V_n(t)\\ &=\textstyle\sum_{j=1}^{n}(r(X_j,a_n^{-1}t)-E[r(X_j,a_n^{-1}t)])+E[a_n^2(g(X_1,\theta_0+a_n^{-1}t)-g(X_1,\theta_0))]\\ &\quad+2^{-1}t'Wt\xrightarrow{\Pr}0.\end{aligned}$$

Using the concavity of $U_n(\cdot)-V_n(\cdot)$, Theorem II.1 in Andersen and Gill [1], implies that for each $0<M<\infty$,

$$\sup_{t\in\Re^d,|t|\le M}|U_n(t)-V_n(t)|\xrightarrow{\Pr}0.$$

Hence,

$$\begin{aligned}(3.3)\qquad &\left|\int_{t\in\Re^d,|t|\le M}\left(e^{U_n(t)}-e^{V_n(t)}\right)dt\right|\le c_dM^d\sup_{t\in\Re^d,|t|\le M}\left|e^{U_n(t)}-e^{V_n(t)}\right|\\ &\le c_dM^d\sup_{t\in\Re^d,|t|\le M}e^{V_n(t)}\left|e^{U_n(t)-V_n(t)}-1\right|\\ &\le c_dM^d\sup_{t\in\Re^d,|t|\le M}e^{V_n(t)}|U_n(t)-V_n(t)|e^{|U_n(t)-V_n(t)|}\xrightarrow{\Pr}0.\end{aligned}$$

By (3.3) to show (3.1), we need to prove that for each $\epsilon>0$,

$$(3.4)\qquad \lim_{M\to\infty}\limsup_{n\to\infty}\Pr\left\{\int_{|t|>M,\theta_0+a_n^{-1}t\in\Theta}e^{U_n(t)}\,dt\ge\epsilon\right\}=0$$

and

$$(3.5)\qquad \lim_{M\to\infty}\limsup_{n\to\infty}\Pr\left\{\int_{|t|>M}e^{V_n(t)}\,dt\ge\epsilon\right\}=0.$$

Using that U_n is concave, for each $|t|\ge M$, we have that

$$\begin{aligned}U_n(|t|^{-1}Mt)&=U_n((1-|t|^{-1}M)0+|t|^{-1}Mt)\\ &\ge(1-|t|^{-1}M)U_n(0)+|t|^{-1}MU_n(t)=|t|^{-1}MU_n(t).\end{aligned}$$

Hence, for each $|t|\ge M$,

$$(3.6)\qquad U_n(t)\le M^{-1}|t|U_n(|t|^{-1}Mt)\le M^{-1}|t|\sup_{|s|=M}U_n(s).$$

Thus, for each $\epsilon > 0$ and each $0 < M < \infty$,

$$\begin{aligned}
&\limsup_{n\to\infty} \Pr\left\{\int_{|t|>M,\theta_0+a_n^{-1}t\in\Theta} e^{U_n(t)}\,dt \ge \epsilon\right\}\\
\le&\limsup_{n\to\infty} \Pr\left\{\int_{|t|>M,\theta_0+a_n^{-1}t\in\Theta} e^{M^{-1}|t|\sup_{|s|=M} U_n(s)}\,dt \ge \epsilon\right\}\\
=&\limsup_{n\to\infty} \Pr\left\{\int_{|t|>M} e^{M^{-1}|t|\sup_{|s|=M} V_n(s)}\,dt \ge \epsilon\right\}\\
=&\Pr\left\{\int_{|t|>M} e^{M^{-1}|t|\sup_{|s|=M} U(s)}\,dt \ge \epsilon\right\},
\end{aligned}$$

where $U(t) = t'Z - 2^{-1}t'Wt$, $t \in \Re^d$ and Z is normal $\Re^d$–valued r.v. with mean zero and variance $\mathrm{Var}(\dot{h}(X_1))$. Since V and W are two positive definite matrices,

$$\sup_{|s|=M} U(s) = \sup_{|s|=M}\left(s'V^{1/2}Z_d - 2^{-1}s'Ws\right) \le aM|Z_d| - 2^{-1}bM^2,$$

where $a := \sup_{|u|,|v|=1} u'W^{1/2}v$ and $b := \inf_{|u|=1} u'W^{1/2}u$ satisfy that $0 < a, b < \infty$. Hence, we get that

$$\begin{aligned}
&\lim_{M\to 0}\limsup_{n\to\infty} \Pr\left\{\int_{|t|>M} e^{U_n(t)}\,dt \ge \epsilon\right\}\\
\le&\lim_{M\to 0} \Pr\left\{\int_{|t|>M} e^{M^{-1}|t|\left(aM|Z_d|-2^{-1}bM^2\right)}\,dt \ge \epsilon\right\} = 0.
\end{aligned}$$

and (3.4) follows.

Equation (3.5) follows by an argument similar to the one just done. Therefore, (3.1) follows from (3.3)–(3.5).

We have that

$$\begin{aligned}
&\int_{\Re^d} e^{V_n(t)}\,dt = \int_{\Re^d} e^{-t'Z_n-2^{-1}t'Wt}\,dt\\
=&\int_{\Re^d} e^{2^{-1}Z_n'W^{-1}Z_n-2^{-1}|W^{1/2}t+W^{-1/2}Z_n|^2}\,dt = e^{2^{-1}Z_n'W^{-1}Z_n}\det(W^{-1/2})(2\pi)^{d/2}
\end{aligned}$$

The theorem follows from the previous expression and (3.1). □

Proof of Proposition 2.1. We have that for each $\theta \in \Re^d$,

$$\begin{aligned}
&nE\left[\min\left(|r(X,n^{-1/2}\theta)|, |r(X,n^{-1/2}\theta)|^2\right)\right]\\
\le&nE[r(X,n^{-1/2}\theta)|^2] \le E\Big[\sup_{|t|\le|\theta|n^{-1/2}} |\theta|^2|\dot{g}(X,\theta_0+t) - \dot{g}(X,\theta_0)|^2\Big] \to 0, \text{as } n\to\infty.
\end{aligned}$$

□

Proof of Proposition 2.2. We have that

$$\begin{aligned}
&|E[g(X,\theta) - g(X,\theta_0) - (\theta-\theta_0)'\dot{g}(X,\theta_0) - 2^{-1}(\theta-\theta_0)'\ddot{g}(X,\theta_0)(\theta-\theta_0)]|\\
\le&E\Big[\sup_{|t|\le|\theta-\theta_0|} |\theta-\theta_0|^2|\ddot{g}(X,\theta_0+t) - \ddot{g}(X,\theta_0)|\Big] = o(|\theta-\theta_0|^2), \text{as } \theta\to\theta_0,
\end{aligned}$$

and for each $\theta \in \Re^d$,

$$\begin{aligned}
&nE\left[\min\left(|r(X,n^{-1/2}\theta)-E[r(X,n^{-1/2}\theta)]|,|r(X,n^{-1/2}\theta)-E[r(X,n^{-1/2}\theta)]|^2\right)\right]\\
\le&nE[|r(X,n^{-1/2}\theta)-E[r(X,n^{-1/2}\theta)]|]\\
\le&nE[|r(X,n^{-1/2}\theta)-E[n^{-1}\theta'\ddot{g}(X,\theta_0)\theta]|]\\
&+n|E[n^{-1}\theta'\ddot{g}(X,\theta_0)\theta)]-E[r(X,n^{-1/2}\theta)]|\\
\le&2nE[|r(X,n^{-1/2}\theta)-n^{-1}\theta'\ddot{g}(X,\theta_0)\theta|]\\
\le&2E\Big[\sup_{|t|\le|\theta|n^{-1/2}}|\theta|^2|\ddot{g}(X,\theta_0+t)-\ddot{g}(X,\theta_0)|\Big]\to 0,\text{as } n\to\infty.
\end{aligned}$$

□

Proof of Theorem 2.2. By the change of variables, $t=t_1+n^{-1/2}\theta$,

$$\begin{aligned}
&\int_{\Re^d}\exp\left(\sum_{j=1}^n\log(f_1(X_j-t))\right)\,dt\\
=&n^{-d/2}\int_{\Re^d}\exp\left(\sum_{j=1}^n\log(f_1(X_j-t_1-n^{-1/2}\theta))\right)\,d\theta.
\end{aligned}$$

By the change of variables, $t=t_0+n^{-1/2}\theta$,

$$\begin{aligned}
&\int_{\Re^d}\exp\left(\sum_{j=1}^n\log(f_0(X_j-t))\right)\,dt\\
=&n^{-d/2}\int_{\Re^d}\exp\left(\sum_{j=1}^n\log(f_0(X_j-t_0-n^{-1/2}\theta))\right)\,d\theta.
\end{aligned}$$

Hence, by Theorem 2.1

$$\begin{aligned}
&\log(T_{n,\text{loc}})\\
=&\log\left(\int_{\Re^d}\exp\left(\sum_{j=1}^n\log(f_1(X_j-t_1-n^{-1/2}\theta)\right)\,d\theta\right)\\
&-\log\left(\int_{\Re^d}\exp\left(\sum_{j=1}^n\log(f_0(X_j-t_0-n^{-1/2}\theta)\right)\,d\theta\right)\\
=&\sum_{j=1}^n\log f_1(X_j-t_1)+\log\left((2\pi)^{d/2}\det(W_1^{-1/2})\right)+2^{-1}|W_1^{-1/2}Z_{n,1}|^2\\
&-\sum_{j=1}^n\log f_0(X_j-t_0)-\log\left((2\pi)^{d/2}\det(W_0^{-1/2})\right)-2^{-1}|W_0^{-1/2}Z_{n,0}|^2+o_P(1).
\end{aligned}$$

□

We will need the following lemma, whose proof is omitted, since it is a simple calculus exercise.

Lemma 3.2. *There exists a universal constant c, depending only on p, such that:*
(i) If $p\ge 1$, then

$$||x-\theta|^p-|x|^p|\le c(|x|^{p-1}|\theta|\vee|\theta|^p),$$

for each $x,\theta\in\Re^d$.
(ii) If $p=1$ and $d=1$, then

$$||x-\theta|-|x|+|x|^{-1}\theta'x|\le 2|\theta|I_{|x|\le|\theta|},$$

for each $x, \theta \in \Re^d$.

(iii) If $2 > p \geq 1$, *then*

$$||x-\theta|^p - |x|^p + p|x|^{p-2}\theta' x| \leq c(|x|^{p-2}|\theta|^2 \wedge |\theta|^p),$$

for each $x, \theta \in \Re^d$.

(iv) If $p \geq 2$, *then*

$$||x-\theta|^p - |x|^p + p|x|^{p-2}\theta' x| \leq c(|x|^{p-2}|\theta|^2 \vee |\theta|^p),$$

for each $x, \theta \in \Re^d$.

(v) If $2 \geq p > 0$, *then*

$$||x-\theta|^p - |x|^p + p|x|^{p-2}\theta' x - 2^{-1}p(p-2)|x|^{p-4}(\theta' x)^2 - 2^{-1}p|x|^{p-2}|\theta|^2| \leq c(|x|^{p-3}|\theta|^3 \wedge |x|^{p-2}|\theta|^2),$$

for each $x, \theta \in \Re^d$.

(vi) If $3 \geq p \geq 2$, *then*

$$||x-\theta|^p - |x|^p + p|x|^{p-2}\theta' x - 2^{-1}p(p-2)|x|^{p-4}(\theta' x)^2 - 2^{-1}p|x|^{p-2}|\theta|^2| \leq c(|x|^{p-3}|\theta|^3 \wedge |\theta|^p),$$

for each $x, \theta \in \Re^d$.

(viii) If $p \geq 3$, *then*

$$||x-\theta|^p - |x|^p + p|x|^{p-2}\theta' x - 2^{-1}p(p-2)|x|^{p-4}(\theta' x)^2 - 2^{-1}p|x|^{p-2}|\theta|^2| \leq c(|x|^{p-3}|\theta|^3 \vee |\theta|^p),$$

for each $x, \theta \in \Re^d$.

Corollary 3.1. *Let* $p \geq 1$. *Let* X *be a* $\Re^d$*–valued r.v. Let* $\theta_0 \in \Re^d$. *Suppose that*

(i)

$$E[|X-\theta_0|^p] = \inf_{\theta \in \Re^d} E[|X-\theta|^p] < \infty.$$

(ii) There exists no $v \in \Re^d$ *such that* $P\{v'(X-\theta_0) = 0\} = 1$.

(iii) If $p > 1$, *suppose that* $E[|X|^{2p-2}] < \infty$ *and*

(iv) If $p = 1$ *and* $d > 1$, *suppose that* $E[|X|^{-1}] < \infty$.

(v) If $p = 1$ *and* $d = 1$, *suppose that* $F(x) = P\{X \leq x\}$, $x \in \Re$, *is differentiable at* θ_0 *and* $F'(\theta_0) > 0$.

Then, Theorem 2.1 applies to $g(x,\theta) = -|x-\theta|^p$, $x, \theta \in \Re$, $\phi(x) = p|x - \theta_0|^{p-2}(x-\theta_0)I(x-\theta_0 \neq 0)$ *and* $W = 2F'(\theta_0)$, *if* $p = 1$ *and* $d = 1$, *and*

$$W = \frac{p}{2}E[|X-\theta_0|^{p-2}I_d] + \frac{p(p-2)}{2}E[|X-\theta_0|^{p-2}(X-\theta_0)(X-\theta_0)'], \text{otherwise.}$$

Proof. Without loss of generality, we may assume that $\theta_0 = 0$. Hypotheses (i) and (iv) in Theorem 2.1 are obviously satisfied. Next, we will check the rest of hypotheses case by case. Suppose that $p > 1$. By Lemma 3.2 (iii) and (iv), $E[|X-\theta|^p]$, $\theta \in \Re^d$, is differentiable at zero and its derivative is $E[p|X|^{p-2}X]$. Since $E[|X-\theta|^p]$, $\theta \in \Re^d$, has a maximum at zero, $E[p|X|^{p-2}X] = E[\phi(X)] = 0$. Hence, by Lemma 3.2 (v)–(vii),

$$\begin{aligned} &E[g(X,\theta) - g(X,0)] + 2^{-1}\theta W \theta \\ =&E[-|X-\theta|^p + |X|^p - p|X|^{p-2}\theta' X + 2^{-1}p(p-2)|X|^{p-4}(\theta' X)^2 + 2^{-1}p|X|^{p-2}|\theta|^2] \\ =&o(|\theta|^2). \end{aligned}$$

Condition (iii) in Theorem 2.1 follows from Lemma 3.2 (iii) and (iv).

The case $p = 1$ and $d > 1$, follows similarly to the previous case.

Finally, suppose that $p = 1$ and $d = 1$. Since $E[|X - \theta|]$, $\theta \in \Re$, has a maximum at zero, $F(0) = 1/2$. Hence, $E[|X|^{-1}X] = 0$. For $\theta > 0$, we have that

$$\begin{aligned} & E[-|X-\theta| + |X| - |X|^{-1}X\theta] = E[2(X-\theta)I(0 < X \le \theta)] = \textstyle\int_0^\theta 2(x-\theta)\,dF(x) \\ = \quad & \textstyle\int_0^\theta 2(x-\theta)\,d(F(x)-F(0)) = -2\int_0^\theta (F(x)-F(0))\,dx = -\theta^2 F'(0) + o(\theta^2). \end{aligned}$$

A similar argument holds for $\theta < 0$. Lemma 3.2 (iii),

$$|r(X, n^{-1/2}\theta)| \le 2n^{-1/2}|\theta| I(|X| \le n^{-1/2}|\theta|).$$

Hence,

$$nE\left[\min\left(|r(X,n^{-1/2}\theta)|, |r(X,n^{-1/2}\theta)|^2\right)\right] \le 4|\theta|^2 \Pr\{|X| \le n^{-1/2}|\theta|\} \to 0. \qquad \square$$

Proof of Corollary 2.2. By symmetry, we have

$$E[|X-\theta_0|^{p_1}] = \inf_{\theta \in \Re^d} E[|X-\theta|^{p_1}]$$

and

$$E[|X-\theta_0|^{p_2}|] = \inf_{\theta \in \Re^d} E[|X-\theta|^{p_2}].$$

Corollary 3.1 and the arguments in the proof of Theorem 2.2 imply the result. $\square$

Proof of Theorem 2.3. This theorem follows from the one on location families by a transformation. Let $Y_j = \log X_j$, let $\tilde{f}_1(x) = f_1(e^x)e^x$ and let $\tilde{f}_0(x) = f_0(e^x)e^x$. By the change of variables $\lambda = e^\theta$,

$$\begin{aligned} (3.7) \qquad & \textstyle\int_0^\infty \left(\prod_{j=1}^n \lambda^{-1} f_i(X_j/\lambda)\right)\lambda^{-1}\,d\lambda = \int_{-\infty}^\infty \prod_{j=1}^n \left(f_i(X_j e^{-\theta})e^{-\theta}\right)\,d\theta \\ = \qquad & \textstyle\int_{-\infty}^\infty \prod_{j=1}^n \left(\tilde{f}_i(Y_j-\theta)e^{-Y_j}\right)\,d\theta. \end{aligned}$$

Hence,

$$T_{n,\text{scale}} = \frac{\int_{-\infty}^\infty \prod_{j=1}^n \tilde{f}_1(Y_j-\theta)\,d\theta}{\int_{-\infty}^\infty \prod_{j=1}^n \tilde{f}_0(Y_j-\theta)\,d\theta}$$

and the claim follows from Theorem 2.2. $\square$

Proof of Corollary 2.3. We apply Theorem 2.3. We need to show that Theorem 2.1 applies to $\log(f_0(e^{x-\theta})e^{x-\theta})$ and $\log X$. We have that

$$g_0(x,\theta) = \log(f_0(e^{x-\theta})e^{x-\theta}) = \alpha_0(x-\theta) - e^{x-\theta} - \log(\Gamma(\alpha_0)), x \in \Re,$$

is a concave function on θ, i.e. (i) in Theorem 2.1 holds. We have that

$$\begin{aligned} & E[g_0(\log X, \theta_0 + t) - g_0(\log X, \theta_0)] \\ = \quad & E[\alpha_1(\log X - \theta_0 - t) - Xe^{-\theta_0 - t} - \alpha_0(\log X_1 - \theta_0) + X_1 e^{-\theta_0}] \\ & = \alpha_0(-t) + \alpha_0\lambda_0(e^{-\theta_0} - e^{-\theta_0 - t}) = -\alpha_0(e^{-t} - 1 + t) = -2^{-1}\alpha_0 t^2 + o(|t|^2), \end{aligned}$$

as $t \to 0$, where $\theta_0 = \log\lambda_0$. Hence, (ii) in Theorem 2.1 holds. We take $\phi_0(x) = -\alpha_0 + e^{x-\theta_0}$, $x \in \Re$. Then,

$$E[\phi_0(\log X)] = E[-\alpha_0 + \lambda_0^{-1}X] = 0,$$

$$E[(\phi_0(\log X))^2] = E[(-\alpha_0 + \lambda_0^{-1}X)^2] = \alpha_0.$$

and

$$r_0(x,t) = g_0(x,\theta_0+t) - g_0(x,\theta_0+t) - t\phi_1(x)$$
$$= -e^{x-\theta_0-t} + e^{x-\theta_0} - te^{x-\theta_0} = -e^{x-\theta_0}(e^{-t} - 1 + t) =, x \in \Re.$$

Hence,

$$nE\left[\min\left(|r(\log X, n^{-1/2}t)|, (r(\log X, n^{-1/2}t))^2\right)\right] \le nE\left[(r(\log X, n^{-1/2}t))^2\right]$$
$$= \qquad n\lambda_0^{-2}\left(e^{-n^{-1/2}t} - 1 + n^{-1/2}t\right)^2 E[X^2] \to 0,$$

as $n \to \infty$. Therefore, condition (iii) in Theorem 2.1 holds. Hence, (i) in Theorem 2.3 holds.

We have that

$$E[g_1(\log X,\theta)] = E[\alpha_1(\log X - \theta) - Xe^{-\theta} - \log(\Gamma(\alpha_1))]$$
$$= E[\alpha_1 \log X] - \alpha_1\theta - \lambda_0\alpha_0 e^{-\theta} - \log(\Gamma(\alpha_1))], \theta \in \Re,$$

is minimized when $\alpha_1 = \lambda_0\alpha_0 e^{-\theta}$, i.e. $\theta_1 = \log\lambda_1 = \log(\lambda_0\alpha_0\alpha_1^{-1})$.

The proof that condition (iii) in Theorem 2.3 holds is similar to the proof of (i) and it is omitted. □

References

[1] Andersen, P. K. and Gill, R. (1982). Cox's regression model for counting processes: A large sample study. *Ann. Statist.* **10** 1100–1120. MR0673646

[2] Arcones, M. A. (1998). Asymptotic theory for M-estimators over a convex kernel. *Econometric Theory* **14** 387–422. MR1650029

[3] Atkinson, A. C. (1970). A method for discriminating between models (with discussion). *J. R. Statist. Soc. B* **32** 323–353.

[4] Barndorff-Nielsen, O. E. and Cox, D. R. (1989). *Asymptotic Techniques for Use in Statistics.* Chapman and Hall, New York. MR1010226

[5] Cox, D. R. (1961). Tests of separate families of hypotheses. In *Proc. 4th Berkeley Sympos. Math. Statist. Prob.* **I** 105–123. Univ. California Press, Berkeley, California. MR0131927

[6] Cox, D. R. (1962). Further results on tests of separate families of hypotheses. *J. Roy. Statist. Soc. Ser. B* **24** 406–424. MR0156409

[7] Daniels, H. E. (1961) The asymptotic efficiency of a maximum likelihood estimator. In *Proceedings of the Fourth Berkeley Simposium on Mathematical Statistics and Probability* **I** 151–163. MR0131924

[8] Ducharme, G. R. and Frichot, B. (2003). Quasi most powerful invariant goodness-of-fit tests. *Scand. J. Statist.* **30**(2) 399–414. MR1983133

[9] Haberman, S. J. (1989) Concavity and estimation. *Annals of Statistics* **17** 1631–1661. MR1026303

[10] Hájek, J., Sidák, Z. and Sen, P. K. (1999). *Theory of Rank Tests*, 2nd ed. Academic Press, Inc., San Diego, Califonia. MR1680991

[11] Hjort, N. L. and Pollard, D. (1993). Asymptotics for minimisers of convex processes. Unpublished Technical Report, Department of Statistics, Yale University.

[12] Lehmann, E. L. and Romano, J. P. (2005). *Testing Statistical Hypotheses. Third Edition.* Springer, New York. MR2135927

[13] Lindley, D. V. (1961). The use of prior probability distributions in statistical inference and decisions. In *Proc. 4–th Berkeley Sympos. Math. Statist. Prob.* **I** 453–468. Univ. California Press, Berkeley, CA. MR0156437

[14] Loh, W. Y. (1985). A new method for testing separate families of hypotheses. *J. Amer. Statist. Assoc.* **80**(390) 362–368. MR0792735

[15] Pace, L. and Salvan, A. (1990). Best conditional tests for separate families of hypotheses. *J. Roy. Statist. Soc. Ser. B* **52**(1) 125–134. MR1049305

[16] Pereira, B. de B. (1977). Discriminating among separate models: A bibliography. *Internat. Statist. Rev.* **45**(2) 163–172. MR0494638

[17] Pereira, B. de B. (1977). Separate families of hypothesis. In *MB Encyclopedia of Biostatistics*, 2nd ed. (P. Armitage and T. Colton, eds.) **7** 4881–4886.

[18] Rockafellar, R. T. (1970) *Convex Analysis.* Princeton University Press, Princeton, NJ. MR0274683

[19] White, H. (1982). Regularity conditions for Cox's test of nonnested hypothesis. *J. Econometrics* **19** 301–318. MR0672058