SUPPLEMENT TO: ROBUST DISCRIMINATION DESIGNS
OVER HELLINGER NEIGHBOURHOODS

BY RUI HU AND DOUGLAS P. WIENS

MacEwan University and University of Alberta

We prove Theorem 2.1, and then show that it applies to normal and log-normal densities.

THEOREM 2.1. Given a design space \( S = \{ x_i \}_{i=1}^N \), assume that the experiment has \( n_i \) replicates at each covariate \( x_i \), with \( \sum_{i=1}^N n_i = n \). Define the integrated Kullback-Leibler (KL) divergence \( D \):

(S.1) \[ D(f_0, f_1, \xi | \mu_0, \mu_1) = \int_S \mathcal{I}\{ f_0, f_1 | x, \mu_0(x), \mu_1(x) \} \xi(dx), \]

with \( \xi \) being the design measure placing mass \( \xi_i = n_i/n \) at \( x_i \), and

(S.2) \[ \mathcal{I}\{ f_0, f_1 | x, \mu_0(x), \mu_1(x) \} = \int_{-\infty}^{\infty} f_1(y|\mathbf{x}, \mu_1(\mathbf{x})) \log \left( \frac{f_1(y|\mathbf{x}, \mu_1(\mathbf{x}))}{f_0(y|\mathbf{x}, \mu_0(\mathbf{x}))} \right) dy \]

being the Kullback-Leibler divergence measuring the information lost when \( f_0(y|\mathbf{x}, \mu_0(\mathbf{x})) \) is used to approximate \( f_1(y|\mathbf{x}, \mu_1(\mathbf{x})) \). For any two densities \( f_0, f_1 \) define

\[ \mathcal{R}(y|x_i; f_0, f_1) = \frac{f_1(y|x_i; \mu_1(x_i))}{f_0(y|x_i; \mu_0(x_i))}. \]

Assume that the densities \( f_0, f_1 \) satisfy

(a) for the KL-divergence,

(S.3) \[ nD = O(1), \]

(b) for all \( \delta > 0 \)

(S.4) \[ \lim_{n \to \infty} \sum_{i=1}^N n_i \int_{|\log \mathcal{R}(y|x_i; f_0, f_1)| \geq \delta} f_0(y|x_i; \mu_0(x)) \left( \sqrt{\mathcal{R}(y|x_i; f_0, f_1)} - 1 \right)^2 dy = 0, \]

(c) there is a \( \tau > 0 \) such that

(S.5) \[ \lim_{n \to \infty} \sum_{i=1}^N n_i \int_{|\log \mathcal{R}(y|x_i; f_0, f_1)| \geq \tau} f_1(y|x_i; \mu_1(x)) \log \mathcal{R}(y|x_i; f_0, f_1) dy = 0. \]
Then $\mathcal{R} = 2 \sum_{i=1}^{N} \sum_{l=1}^{n_i} \log r(y|x_i; f_0, f_1)$ and:

(i) under the null hypothesis,

$$\frac{\mathcal{R} + 2nD}{\sqrt{8nD}} \xrightarrow{L} N(0, 1),$$

(ii) under the alternative hypothesis,

$$\frac{\mathcal{R} - 2nD}{\sqrt{8nD}} \xrightarrow{L} N(0, 1).$$

**Proof of Theorem 2.1.** Denote by $F_{1i}, F_{0i}$ the distribution functions with densities $f_1(y|x_i)$ and $f_0(y|x_i)$, respectively. We denote the probability of an event $A$, under $F$, as $F(A)$. To prove Theorem 2.1, we need several preliminary results. In this section we abbreviate $r(y|x_i; f_0, f_1)$ by $r(y|x_i)$.

**Lemma 2.2.** Condition (S.4) holds iff each of

(S.6) \[ \lim_{n \to \infty} \sum_{i=1}^{N} n_i F_{1i}(\log r(y|x_i) \geq \delta) = 0, \]

(S.7) \[ \lim_{n \to \infty} \sum_{i=1}^{N} n_i F_{0i}(\log r(y|x_i) \leq -\delta) = 0, \]

holds for all $\delta > 0$.

**Proof.** This is proven in Theorem 2 of Oosterhoff and van Zwet (2012).

**Lemma 2.3.** Define $F_{0}^{(n)} = \Pi_{i=1}^{N} (F_{0i})^{n_i}$ and $F_{1}^{(n)} = \Pi_{i=1}^{N} (F_{1i})^{n_i}$. Then (S.3) and (S.4) imply that the sequences $\{F_{0}^{(n)}\}$ and $\{F_{1}^{(n)}\}$ are contiguous with respect to each other, i.e. for any sequence $\{A_n\}$ of measurable sets, $\lim_{n \to \infty} F_{0}^{(n)}(A_n) = 0$ iff $\lim_{n \to \infty} F_{1}^{(n)}(A_n) = 0$. Moreover, $\{F_{0}^{(n)}\}$ and $\{F_{1}^{(n)}\}$ being contiguous with respect to each other also implies that

(S.8) \[ \lim_{n \to \infty} \sum_{i=1}^{N} n_i F_{0i}(A_{n_i}) = 0 \Leftrightarrow \lim_{n \to \infty} \sum_{i=1}^{N} n_i F_{1i}(A_{n_i}) = 0 \]

for any collection of measurable sets $A_{n_i}$. 


Proof. Corollary 1 in Oosterhoff and van Zwet (2012) shows that \( \{F_0^{(n)}\} \) and \( \{F_1^{(n)}\} \) are contiguous with respect to each other under \((S.6)\), \((S.7)\) and the condition

\[
\lim_{n \to \infty} \sup_{i=1}^{N} n_i d^2_i(f_1(y|x_i), f_0(y|x_i)) < \infty. \tag{S.9}
\]

Lemma 2.2 shows that condition \((S.4)\) is equivalent to \((S.6)\) and \((S.7)\). Therefore, to establish contiguity it suffices to show that \((S.3)\) implies \((S.9)\). For this we note that the KL distance between \( f_0(y|x_i) \), \( f_1(y|x_i) \) is larger than twice their squared Hellinger distance, i.e.

\[
\int_{-\infty}^{\infty} f_1(y|x_i) \log (r(y|x_i)) dy \geq \int_{-\infty}^{\infty} (f_1^{1/2}(y|x_i) - f_0^{1/2}(y|x_i))^2 dy = 2d^2_1(f_1(y|x_i), f_0(y|x_i));
\]

now \((S.3)\) gives \( \sum_{i=1}^{N} n_i d^2_i(f_1(y|x_i), f_0(y|x_i)) = O(1) \). The equivalence in \((S.8)\) follows from (2.1) in Oosterhoff and van Zwet (2012).

Lemma 2.4. Under conditions \((S.3)\) and \((S.4)\), for all \( \delta > 0 \)

\[
\lim_{n \to \infty} \sum_{i=1}^{N} n_i F_{1i}(\delta | \log r(y|x_i)) \geq 0; \tag{S.10}
\]

\[
\lim_{n \to \infty} \sum_{i=1}^{N} n_i F_{0i}(\delta | \log r^{-1}(y|x_i)) \geq 0. \tag{S.11}
\]

Proof. Under \((S.3)\) and \((S.4)\) we have both \((S.7)\) and \((S.8)\), hence

\[
\lim_{n \to \infty} \sum_{i=1}^{N} n_i F_{1i}(\log r(y|x_i) \leq -\delta) = 0;
\]

this with \((S.6)\) gives \((S.10)\). The proof of \((S.11)\) is completely analogous.

The following lemma shows two different expansions of \( \log r(y|x_i) \). Both expansions are used in the proof of the asymptotic distribution of the test statistic \( R \) under the null and alternative hypotheses.

Lemma 2.5. Assume that \( |\log r(y|x_i)| < \delta \) for some \( \delta \in (0, 1) \). Then \( r(y|x_i) \) can be expanded as

\[
\log r(y|x_i) = -2(r^{-1/2}(y|x_i) - 1) + (r^{-1/2}(y|x_i) - 1)^2(1 + \rho_{1i\delta}). \tag{S.12}
\]
and

\[(S.13) \quad \log r(y|x_i) = 2(r^{1/2}(y|x_i) - 1) - (1 - r^{1/2}(y|x_i))^2(1 + \rho_{1i\delta}),\]

where \(|\rho_{ji\delta}| < 3\delta, j = 1, 2\).

**Proof.** Expansion (S.13) is (3.11) in Oosterhoff and van Zwet (2012); expansion (S.12) is obtained in a similar manner.

The following lemma gives two expansions of \((\log r(y|x_i))^2\) under the conditions of Lemma 2.5. It is a rather immediate consequence of Lemma 2.5 and so is presented without proof. The first expansion (S.14) is used in the determination of the asymptotic distribution of the test statistic \(R\) under the alternative hypothesis \(H_1\); the second expansion (S.15) is used in the proof of the asymptotic distribution of the test statistic \(R\) under the null hypothesis \(H_0\).

**Lemma 2.6.** Assume \(|\log r(y|x_i)| < \delta\). Let \(0 < \delta < 1\). Then

\[(S.14) \quad (\log r(y|x_i))^2 = 4(r^{-1/2}(y|x_i) - 1)^2 + (1 + \rho_{1i\delta})(r^{-1/2}(y|x_i) - 1)^2 \varpi_{1i\delta}\]

and

\[(S.15) \quad (\log r(y|x_i))^2 = 4(1 - r^{1/2}(y|x_i))^2 + (1 + \rho_{2i\delta})(1 - r^{1/2}(y|x_i))^2 \varpi_{2i\delta}\]

where \(|\varpi_{ji\delta}| < 6\delta, j = 1, 2\) and \(i = 1, ..., N\).

We now give the proof of the main result Theorem 2.1. Notice that for the test statistic \(R = 2 \sum_{i=1}^{N} n_i \log r(y|x_i)\), the means and variances under the two hypotheses are

\[E_{H_0}(R) = 2 \sum_{i=1}^{N} n_i \int (\log r(y|x_i)) f_0(y|x_i) dy,\]

\[\text{VAR}_{H_0}(R) = 4 \sum_{i=1}^{N} n_i \int (\log r(y|x_i))^2 f_0(y|x_i) dy - (E_{H_0}(R))^2;\]

and

\[E_{H_1}(R) = 2 \sum_{i=1}^{N} n_i \int (\log r(y|x_i)) f_1(y|x_i) dy,\]

\[\text{VAR}_{H_1}(R) = 4 \sum_{i=1}^{N} n_i \int (\log r(y|x_i))^2 f_1(y|x_i) dy - (E_{H_1}(R))^2.\]
According to the Normal Convergence Criterion (Loeve 1963, p. 316) and an equivalent form of this result, to prove Theorem 2.1 it is sufficient to prove the following:

(1) under $H_0$,

\begin{equation}
\lim_{n \to \infty} \sum_{i=1}^{N} n_i F_{0i}(|\log r(Y|x_i)| \geq \delta) = 0, \quad \text{for every } \delta > 0,
\end{equation}

\begin{equation}
\lim_{\delta \to 0^+} \lim_{n \to \infty} \sum_{i=1}^{N} n_i \left( \int_{|\log r(y|x_i)| \leq \delta} (\log r(y|x_i)) f_0(y|x_i) dy + D \right) = 0,
\end{equation}

\begin{equation}
\lim_{\delta \to 0^+} \lim_{n \to \infty} \sum_{i=1}^{N} n_i \left\{ \int_{|\log r(y|x_i)| \leq \delta} (\log r(y|x_i))^2 f_0(y|x_i) dy - \left( \int_{|\log r(y|x_i)| \leq \delta} (\log r(y|x_i)) f_0(y|x_i) dy \right)^2 - 2D \right\} = 0.
\end{equation}

(2) under $H_1$,

\begin{equation}
\lim_{n \to \infty} \sum_{i=1}^{N} n_i F_{1i}(|\log r(Y|x_i)| \geq \delta) = 0, \quad \text{for every } \delta > 0,
\end{equation}

\begin{equation}
\lim_{\delta \to 0^+} \lim_{n \to \infty} \sum_{i=1}^{N} n_i \left( \int_{|\log r(y|x_i)| \leq \delta} (\log r(y|x_i)) f_1(y|x_i) dy - D \right) = 0,
\end{equation}

\begin{equation}
\lim_{\delta \to 0^+} \lim_{n \to \infty} \sum_{i=1}^{N} n_i \left\{ \int_{|\log r(y|x_i)| \leq \delta} (\log r(y|x_i))^2 f_1(y|x_i) dy - \left( \int_{|\log r(y|x_i)| \leq \delta} (\log r(y|x_i)) f_1(y|x_i) dy \right)^2 - 2D \right\} = 0.
\end{equation}

Conditions (S.16) and (S.19) are direct results of Lemma 2.4. We prove (S.17); the proof of (S.20) is similar. For $\tau > \delta > 0$ we first split $nD$ into three terms

\[ nD = 3 \sum_{j=1}^{3} \sum_{i=1}^{N} n_i \int_{E_j} (\log r(y|x_i)) f_j(y|x_i) dy, \]

where $E_1 = \{|\log r(y|x_i)| \leq \delta\}$, $E_2 = \{\delta < |\log r(y|x_i)| \leq \tau\}$, $E_3 = \{|\log r(y|x_i)| > \tau\}$. Then in (S.17) we have that

\begin{equation}
\left| \sum_{i=1}^{N} n_i \left( \int_{|\log r(y|x_i)| \leq \delta} (\log r(y|x_i)) f_0(y|x_i) dy + D \right) \right| \leq \sum_{i=1}^{N} n_i \int_{E_1} \left( f_0(y|x_i) + f_1(y|x_i) \right) dy
\end{equation}

\begin{equation}
\sum_{j=2,3} \sum_{i=1}^{N} n_i \int_{E_j} (\log r(y|x_i)) f_1(y|x_i) dy \right|.
\end{equation}
We first prove that the terms in (S.23) tend to zero as $n \to \infty$. For $j = 2$, using (S.10) we have

$$
\left| \sum_{i=1}^{N} n_i \int_{E_2} (\log r(y|x_i)) f_1(y|x_i) dy \right| \leq \tau \sum_{i=1}^{N} n_i F_{1i}(\log r(y|x_i) \geq \delta) \to 0.
$$

Now consider $j = 3$. Using (S.10) we have

$$
\sum_{i=1}^{N} n_i \int_{\log r(y|x_i) < -\tau} (\log r(y|x_i)) f_1(y|x_i) dy < -\tau \sum_{i=1}^{N} n_i F_{1i}(\log r(Y|x_i) < -\tau) \to 0
$$

and since $-\log u > 1 - u$

$$
\sum_{i=1}^{N} n_i \int_{\log r(y|x_i) < -\tau} (\log r(y|x_i)) f_1(y|x_i) dy
$$

$$
= \sum_{i=1}^{N} n_i \int_{\log r(y|x_i) < -\tau} (-\log r^{-1}(y|x_i)) f_1(y|x_i) dy
$$

$$
\geq \sum_{i=1}^{N} n_i \int_{\log r(y|x_i) < -\tau} (1 - r^{-1}(y|x_i)) f_1(y|x_i) dy
$$

$$
= \sum_{i=1}^{N} n_i F_{1i}(\log r(y|x_i) < -\tau) - \sum_{i=1}^{N} n_i F_{0i}(\log r(y|x_i) < -\tau) \to 0.
$$

Therefore,

$$
\left| \sum_{i=1}^{N} n_i \int_{\log r(y|x_i) < -\tau} (\log r(y|x_i)) f_1(y|x_i) dy \right| \to 0.
$$

Combining this with (S.5),

$$
\left| \sum_{i=1}^{N} n_i \int_{E_3} (\log r(y|x_i)) f_1(y|x_i) dy \right| \to 0.
$$

Finally, the sum in (S.22) is, by Lemma 2.5, equal to

$$
\sum_{i=1}^{N} n_i \int_{E_1} \left[ (f_1^{1/2}(y|x_i) - f_0^{1/2}(y|x_i))^2 (\rho_{1i\delta} - \rho_{2i\delta}) \right] dy
$$

$$
+ 2 \sum_{i=1}^{N} n_i \int_{E_1} (f_1(y|x_i) - f_0(y|x_i)) dy.
$$

(S.24)
The second sum in (S.24) tends to 0 as $n \to \infty$ by (S.8) and (S.10). For the first we have, again using Lemma 2.5, that

$$
\lim_{\delta \to -0+} \lim_{n \to \infty} \sum_{i=1}^{N} n_i \int_{E_1} \left( \left( f_1^{1/2}(y|x_i) - f_0^{1/2}(y|x_i) \right)^2 (\rho_{1i\delta} - \rho_{2i\delta}) \right) dy
\leq \lim_{\delta \to -0+} \lim_{n \to \infty} \sum_{i=1}^{N} 6\delta n_i \int_{E_1} \left( \left( f_1^{1/2}(y|x_i) - f_0^{1/2}(y|x_i) \right)^2 \right) dy
\leq \lim_{\delta \to -0+} \lim_{n \to \infty} (6\delta nD) = 0.
$$

This completes the proof of (S.17).

The proofs of (S.18) and (S.21) run along the same lines, but using the bounds of Lemma 2.6 rather than those of Lemma 2.5. ■

**Applying (S.3)-(S.5) to Normal densities.** As in Remark 1 of the main paper, assume that $f_j(y|x_i), \mu_j(x_i), \sigma = \frac{1}{2} \phi \left( \frac{y - \mu_j(x_i)}{\sigma} \right)$, $j = 0, 1$, $i = 1, \ldots, N$, satisfy $\mu_1(x_i) = \mu_0(x_i) + \Delta(x_i)/\sqrt{n}$. Without loss of generality we take $\mu_0(x) = 0$, $\sigma = 1$. Condition (S.3) follows from

$$
\int f_1(y|x_i, \mu_1(x_i)) \log \frac{f_1(y|x_i, \mu_1(x_i))}{f_0(y|x_i, \mu_0(x_i))} dy = \frac{\Delta^2(x_i)}{2n}.
$$

For condition (S.4) we note that

$$
\log \frac{f_1(y|x_i, \mu_1(x_i))}{f_0(y|x_i, \mu_0(x_i))} = \mu_1(x_i) (y - \mu_1(x_i)/2),
$$

so that for any $\delta > 0$

$$
|\log r(y|x_i)| \geq \frac{\Delta^2(x_i)}{2\sqrt{n}} - \sqrt{n}\delta > y\Delta(x_i) \text{ or } \frac{\Delta^2(x_i)}{2\sqrt{n}} + \sqrt{n}\delta < y\Delta(x_i).
$$

For sufficiently large $n$ there exists $\tilde{\delta}$ satisfying

$$
\min \left( \left| \frac{\Delta(x_i)}{2\sqrt{n}} - \sqrt{n}\delta \right|, \left| \frac{\Delta(x_i)}{2\sqrt{n}} + \sqrt{n}\delta \right| \right) > \sqrt{n}\tilde{\delta}.
$$

Then

$$
\int_{\{|\log r(y|x_i)| \geq \delta\}} \left( \sqrt{f_0(y|x_i, \mu_0(x_i))} - \sqrt{f_1(y|x_i, \mu_1(x_i))} \right)^2 dy
\leq \int_{\{|y| \geq \sqrt{n}\tilde{\delta}\}} \left( \sqrt{f_0(y|x_i, \mu_0(x_i))} - \sqrt{f_1(y|x_i, \mu_1(x_i))} \right)^2 dy
= \Phi \left( \mu_1(x_i) - \sqrt{n}\tilde{\delta} \right) + \Phi \left( -\mu_1(x_i) - \sqrt{n}\tilde{\delta} \right) + 2\Phi \left( -\sqrt{n}\tilde{\delta} \right)
-2e^{-\mu_1^2(x_i)/8} \left( \Phi \left( \frac{\mu_1(x_i)}{2} - \sqrt{n}\tilde{\delta} \right) + \Phi \left( -\frac{\mu_1(x_i)}{2} - \sqrt{n}\tilde{\delta} \right) \right) .
$$
Now (S.4) follows from the observation that
\[
n \Phi \left( \frac{\Delta(x_i) n^{-1/2} - \sqrt{n} \delta}{\sqrt{n}} \right) \to 0 \quad \text{as} \quad n \to \infty,
\]
for each of \( \Delta(x_i) = \pm \Delta(x_i) \), \( \pm \Delta(x_i) \)/2 or 0.

Finally, we show that condition (S.5) holds for any \( \tau > 0 \). First take \( \Delta(x_i) > 0 \). For \( \tau > 0 \) we have, by calculations similar to those above, that
\[
0 \leq n \int_{\log r \geq \tau} f_1(y|x_i, \mu_1(x_i)) \log \frac{f_1(y|x_i, \mu_1(x_i))}{f_0(y|x_i, \mu_0(x_i))} dy
\]
\[
\leq n \int_{y \geq \tau \sqrt{n}/\Delta(x_i)} f_1(y|x_i, \mu_1(x_i)) \log \frac{f_1(y|x_i, \mu_1(x_i))}{f_0(y|x_i, \mu_0(x_i))} dy
\]
\[
= \sqrt{n} \Delta(x_i) \phi \left( \frac{\Delta(x_i)}{\sqrt{n}} - \frac{\tau \sqrt{n}}{\Delta(x_i)} \right) + \frac{\Delta^2(x_i)}{2n} \Phi \left( \frac{\Delta(x_i)}{\sqrt{n}} - \frac{\tau \sqrt{n}}{\Delta(x_i)} \right)
\]
\[
\to 0 \quad \text{as} \quad n \to \infty.
\]
For \( \Delta(x_i) < 0 \), we reach the same conclusion. \( \blacksquare \)

**Applying (S.3)-(S.5) to log-Normal densities.** Under the conditions of Remark 1 of the main paper, the densities are

(S.25) \( f_j(y|x, \mu_j(x), v^2) = y^{-1} \left( 2\pi \sigma^2_j(x) \right)^{-1/2} e^{-\frac{(\log y - \vartheta_j(x))^2}{2\sigma^2_j(x)}} I(y > 0) \), where
\[
\vartheta_j(x) = \log \left[ \frac{\mu_j(x)}{\left(1 + v^2_j(x)/\mu_j^2(x)\right)^{1/2}} \right] \quad \text{and} \quad \sigma^2_j(x) = \log \left[ 1 + \frac{v^2}{\mu_j^2(x)} \right].
\]

Using \( \mu_1(x_i) = \mu_0(x_i) + \Delta(x_i)/\sqrt{n} \) we have that

(S.26) \( \vartheta_1(x_i) = \vartheta_0(x_i) + \kappa(x_i)/\sqrt{n} + o(n^{-1/2}) \), for

(S.27) \( \kappa(x_i) = \frac{1}{\mu_0(x_i)} + \frac{v^2}{\mu_0^3(x_i) + v^2 \mu_0(x_i)} \),

and that \( \sigma^2_j(x) = \sigma^2_0(x) + o(1) \). One can now substitute (S.26) and (S.27) into (S.25) and proceed as above to conclude that \( f_0(y|x, \mu_0(x), v^2) \) and \( f_1(y|x, \mu_1(x), v^2) \) satisfy all the conditions in Theorem 2.1. Moreover, condition (S.5) holds for any \( \tau > 0 \). \( \blacksquare \)

**References.**
