Supplement to "On a Projective Ensemble Approach to Two Sample Test for Equality of Distributions"

Zhimei Li [1] **Yaowu Zhang** [1]

S.1. Proof of Theorem 1

The assertion that T is nonnegative is straightforward because $\{F_{\boldsymbol\beta}(t) - G_{\boldsymbol\beta}(t)\}^2$ and the weight function are both nonnegative when $H(\boldsymbol\beta, t)$ is the cumulative distribution function of a $p+1$ dimensional multivariate joint normal random vector with mean $\mathbf{0}$ and covariance $\mathbf{I}_{p+1}$. In addition, T equals zero if and only if $F = G$ because the weight function is positive for almost all $\boldsymbol\beta$ and t.

We now show that $T = T_1 - 2T_2 + T_3$. For simplicity, we only show that

$$\begin{aligned} & \iint F_{\boldsymbol\beta}^2(t)dH(\boldsymbol\beta, t) \\ = \quad & \frac{1}{4} + \frac{1}{2\pi} E \arcsin\left(\frac{1 + \mathbf{x}_1^{\mathrm T}\mathbf{x}_2}{\sqrt{1 + \mathbf{x}_1^{\mathrm T}\mathbf{x}_1}\sqrt{1 + \mathbf{x}_2^{\mathrm T}\mathbf{x}_2}}\right). \end{aligned}$$

By applying the Fubini's theorem, and treating $\mathbf{x}_1$ and $\mathbf{x}_2$ as constants, $(\boldsymbol\beta, t)^{\mathrm T}$ as a $p+1$ dimensional multivariate joint normal random vector with cumulative distribution function $H(\boldsymbol\beta, t)$,

$$\begin{aligned} & \iint F_{\boldsymbol\beta}^2(t)dH(\boldsymbol\beta, t) \\ = \quad & E \iint I(\boldsymbol\beta^{\mathrm T}\mathbf{x}_1 \le t, \boldsymbol\beta^{\mathrm T}\mathbf{x}_2 \le t) dH(\boldsymbol\beta, t) \\ = \quad & E\left\{\mathrm{P}\left(t - \boldsymbol\beta^{\mathrm T}\mathbf{x}_1 \ge 0, t - \boldsymbol\beta^{\mathrm T}\mathbf{x}_2 \ge 0 \mid \mathbf{x}_1, \mathbf{x}_2\right)\right\}. \end{aligned}$$

For each $\mathbf{x}_1$ and $\mathbf{x}_2$, $t - \boldsymbol\beta^{\mathrm T}\mathbf{x}_1$ and $t - \boldsymbol\beta^{\mathrm T}\mathbf{x}_2$ are joint normal with mean vector zero and correlation $\frac{1+\mathbf{x}_1^{\mathrm T}\mathbf{x}_2}{\sqrt{1+\mathbf{x}_1^{\mathrm T}\mathbf{x}_1}\sqrt{1+\mathbf{x}_2^{\mathrm T}\mathbf{x}_2}}$. Therefore, by applying Lemma 1, we have

$$\begin{aligned} & \left\{\mathrm{P}\left(t - \boldsymbol\beta^{\mathrm T}\mathbf{x}_1 \ge 0, t - \boldsymbol\beta^{\mathrm T}\mathbf{x}_2 \ge 0 \mid \mathbf{x}_1, \mathbf{x}_2\right)\right\} \\ = & \frac{1}{4} + \frac{1}{2\pi} \arcsin\left(\frac{1 + \mathbf{x}_1^{\mathrm T}\mathbf{x}_2}{\sqrt{1 + \mathbf{x}_1^{\mathrm T}\mathbf{x}_1}\sqrt{1 + \mathbf{x}_2^{\mathrm T}\mathbf{x}_2}}\right). \end{aligned}$$

With similar arguments for dealing with $\iint G_{\boldsymbol\beta}^2(t)dH(\boldsymbol\beta, t)$ and $\iint F_{\boldsymbol\beta}(t)G_{\boldsymbol\beta}(t)dH(\boldsymbol\beta, t)$, the proof is completed.

□

S.2. Proof of Theorem 2

Define the empirical processes

$$\zeta_{m,n}(\boldsymbol\beta, t) = \sqrt{mn/(m+n)}\{U_m(\boldsymbol\beta, t) - V_n(\boldsymbol\beta, t)\}$$

where

$$U_m(\boldsymbol\beta, t) = m^{-1}\sum_{i=1}^m I(\boldsymbol\beta^{\mathrm T}\mathbf{x}_i \le t),$$

$$V_n(\boldsymbol\beta, t) = n^{-1}\sum_{i=1}^n I(\boldsymbol\beta^{\mathrm T}\mathbf{y}_i \le t).$$

Then it can be verified that

$$\begin{aligned} \frac{mn}{m+n}\widehat{T} \quad &= \quad \frac{2\pi mn}{m+n}\iint \left\{\widehat{F}_{\boldsymbol\beta}(t) - \widehat{G}_{\boldsymbol\beta}(t)\right\}^2 dH(\boldsymbol\beta, t) \\ &= \quad 2\pi \iint \{\zeta_{m,n}(\boldsymbol\beta, t)\}^2 dH(\boldsymbol\beta, t). \end{aligned}$$

Under the null hypothesis, $\mathbf{x}$ and $\mathbf{y}$ are equally distributed, then we have

$$\begin{aligned} & E\{\zeta_{m,n}(\boldsymbol\beta, t)\} \\ = \quad & \sqrt{mn/(m+n)}E\left\{U_m(\boldsymbol\beta, t) - V_n(\boldsymbol\beta, t)\right\} \\ = \quad & 0. \end{aligned}$$

In addition,

$$\begin{aligned} & \mathrm{cov}\{U_m(\boldsymbol\beta, t) - V_n(\boldsymbol\beta, t), U_m(\boldsymbol\alpha, t) - V_n(\boldsymbol\alpha, s)\} \\ = \quad & \mathrm{cov}\left[\frac{1}{m}\sum_{i=1}^m \left\{I\left(\boldsymbol\beta^{\mathrm T}\mathbf{x}_i \le t\right)\right\} - \frac{1}{n}\sum_{i=1}^n \left\{I\left(\boldsymbol\beta^{\mathrm T}\mathbf{y}_i \le t\right)\right\},\right. \\ & \left.\frac{1}{m}\sum_{i=1}^m \left\{I\left(\boldsymbol\alpha^{\mathrm T}\mathbf{x}_i \le s\right)\right\} - \frac{1}{n}\sum_{i=1}^n \left\{I\left(\boldsymbol\alpha^{\mathrm T}\mathbf{y}_i \le s\right)\right\}\right] \\ = \quad & \frac{1}{m^2}\mathrm{cov}\left\{\sum_{i=1}^m I(\boldsymbol\beta^{\mathrm T}\mathbf{x}_i \le t), \sum_{i=1}^m I(\boldsymbol\alpha^{\mathrm T}\mathbf{x}_i \le s)\right\} \\ & + \frac{1}{n^2}\mathrm{cov}\left\{\sum_{i=1}^n I(\boldsymbol\beta^{\mathrm T}\mathbf{y}_i \le t), \sum_{i=1}^n I(\boldsymbol\alpha^{\mathrm T}\mathbf{y}_i \le s)\right\} \end{aligned}$$

[1]Research Institute for Interdisciplinary Sciences, School of Information Management and Engineering, Shanghai University of Finance and Economics, Shanghai, China. Correspondence to: Yaowu Zhang <zhang.yaowu@mail.shufe.edu.cn>.

Proceedings of the 37th International Conference on Machine Learning, Online, PMLR 119, 2020.

$$
\begin{aligned}
&= \frac{1}{m}\operatorname{cov}\left\{I(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}\le t), I(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}\le s)\right\} \\
&\quad +\frac{1}{n}\operatorname{cov}\left\{I(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{y}\le t), I(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{y}\le s)\right\} \\
&= \frac{m+n}{mn}\big\{\mathrm{P}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}\le t, \boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}\le s) \\
&\qquad -\mathrm{P}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}\le t)\mathrm{P}(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}\le s)\big\}.
\end{aligned}
$$

Therefore, the covariance function of $\zeta_{m,n}(\boldsymbol{\beta},t)$ can be written as

$$
\begin{aligned}
&\operatorname{cov}\left\{\zeta_{m,n}(\boldsymbol{\beta},t), \zeta_{m,n}(\boldsymbol{\alpha},s)\right\} \\
=\;& \frac{mn}{m+n}\operatorname{cov}\{U_m(\boldsymbol{\beta},t)-V_n(\boldsymbol{\beta},t), \\
& U_m(\boldsymbol{\alpha},s)-V_n(\boldsymbol{\alpha},s)\} \\
=\;& \mathrm{P}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}\le t, \boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}\le s)-\mathrm{P}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}\le t)\mathrm{P}(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}\le s).
\end{aligned}
$$

Consequently, by noting that $I(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}\le s)$ belongs to the VC class, and according to (Van Der Vaart & Wellner, 1996), it follows that the empirical processes $\zeta_{m,n}(\boldsymbol{\beta},t)$ converges in distribution to a Gaussian process $\zeta(\boldsymbol{\beta},t)$, where the mean function is zero and the covariance function $\operatorname{cov}\{\zeta(\boldsymbol{\beta},t),\zeta(\boldsymbol{\alpha},s)\}$ is given by

$$
\mathrm{P}\left(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}\le t, \boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}\le s\right)-\mathrm{P}\left(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}\le t\right)\mathrm{P}\left(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}\le s\right).
$$

Then we have

$$
\begin{aligned}
\frac{mn}{m+n}\widehat{T} &= 2\pi\iint\{\zeta_{m,n}(\boldsymbol{\beta},t)\}^2 dH(\boldsymbol{\beta},t) \\
&\xrightarrow{d} 2\pi\iint\{\zeta(\boldsymbol{\beta},t)\}^2 dH(\boldsymbol{\beta},t),
\end{aligned}
$$

which completes the proof.

□

S.3. Proof of Theorem 3

Under the global alternative, there exists some $\boldsymbol{\beta}$ and t, such that $\mathrm{P}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}\le t)\ne\mathrm{P}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{y}\le t)$. Therefore, we have

$$
\begin{aligned}
&\{U_m(\boldsymbol{\beta},t)-V_n(\boldsymbol{\beta},t)\}^2-\left\{\mathrm{P}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}\le t)-\mathrm{P}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{y}\le t)\right\}^2 \\
&= 2\left\{\mathrm{P}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}\le t)-\mathrm{P}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{y}\le t)\right\}\left\{U_m(\boldsymbol{\beta},t)-\mathrm{P}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}\le t)\right. \\
&\qquad \left.-V_n(\boldsymbol{\beta},t)+\mathrm{P}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{y}\le t)\right\}+o_p(m^{-1/2}+n^{-1/2}).
\end{aligned}
$$

With Fubini's theorem, it is easy to show that

$$
\begin{aligned}
&\iint\left\{\mathrm{P}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}\le t)-\mathrm{P}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{y}\le t)\right\}I(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}_i\le t)dH(\boldsymbol{\beta},t) \\
&\qquad = \frac{1}{4}+\frac{1}{2\pi}Z_{1i},
\end{aligned}
$$

where Z_{1i} is the independent copy of Z_1 defined as

$$
\begin{aligned}
E\Bigg\{&\arcsin\left(\frac{1+\widetilde{\mathbf{x}}^{\mathrm{T}}\mathbf{x}}{\sqrt{1+\widetilde{\mathbf{x}}^{\mathrm{T}}\widetilde{\mathbf{x}}}\sqrt{1+\mathbf{x}^{\mathrm{T}}\mathbf{x}}}\right) \\
&-\arcsin\left(\frac{1+\mathbf{x}^{\mathrm{T}}\widetilde{\mathbf{y}}}{\sqrt{1+\mathbf{x}^{\mathrm{T}}\mathbf{x}}\sqrt{1+\widetilde{\mathbf{y}}^{\mathrm{T}}\widetilde{\mathbf{y}}}}\right)\Bigg|\,\mathbf{x}\Bigg\} \quad \text{(S.3.1)}
\end{aligned}
$$

and $(\widetilde{\mathbf{x}},\widetilde{\mathbf{y}})$ is the independent copy of $(\mathbf{x},\mathbf{y})$. Similarly, we have

$$
\begin{aligned}
&\iint\left\{\mathrm{P}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}\le t)-\mathrm{P}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{y}\le t)\right\}I(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{y}_i\le t)dH(\boldsymbol{\beta},t) \\
&\qquad = \frac{1}{4}+\frac{1}{2\pi}Z_{2i},
\end{aligned}
$$

where Z_{2i} is the independent copy of Z_2 given by

$$
\begin{aligned}
E\Bigg\{&\arcsin\left(\frac{1+\widetilde{\mathbf{x}}^{\mathrm{T}}\mathbf{y}}{\sqrt{1+\widetilde{\mathbf{x}}^{\mathrm{T}}\widetilde{\mathbf{x}}}\sqrt{1+\mathbf{y}^{\mathrm{T}}\mathbf{y}}}\right) \\
&-\arcsin\left(\frac{1+\widetilde{\mathbf{y}}^{\mathrm{T}}\mathbf{y}}{\sqrt{1+\widetilde{\mathbf{y}}^{\mathrm{T}}\widetilde{\mathbf{y}}}\sqrt{1+\mathbf{y}^{\mathrm{T}}\mathbf{y}}}\right)\Bigg|\,\mathbf{y}\Bigg\} \quad \text{(S.3.2)}
\end{aligned}
$$

Combining the above results, we have

$$
\begin{aligned}
&\widehat{T}-T \\
=\;& 2\pi\iint\Big[\{U_m(\boldsymbol{\beta},t)-V_n(\boldsymbol{\beta},t)\}^2 \\
& -\left\{\mathrm{P}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}\le t)-\mathrm{P}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{y}\le t)\right\}^2\Big]dH(\boldsymbol{\beta},t) \\
=\;& 2\left\{m^{-1}\sum_{i=1}^{m}Z_{1i}-E(Z_{1i})-n^{-1}\sum_{i=1}^{n}Z_{2i}+E(Z_{2i})\right\} \\
& +o_p(m^{-1/2}+n^{-1/2}),
\end{aligned}
$$

which entails the desired result according to the central limit theorem and slutsky theorem.

□

S.4. Proof of Theorem 4

Under the local alternative, we have

$$
\mathrm{P}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}\le t)=\mathrm{P}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{y}\le t)+(m+n)^{-1/2}\ell(\boldsymbol{\beta},t).
$$

Then it can be shown that

$$
\begin{aligned}
& E\{\zeta_{m,n}(\boldsymbol{\beta},t)\} \\
=\;& \sqrt{mn/(m+n)}E\left\{U_m(\boldsymbol{\beta},t)-V_n(\boldsymbol{\beta},t)\right\} \\
=\;& \sqrt{mn}/(m+n)\ell(\boldsymbol{\beta},t),
\end{aligned}
$$

which converges in probability to $\sqrt{\tau(1-\tau)}\ell(\boldsymbol{\beta},t)$ as $\min(m,n)\to\infty$. In addition, similar to the proof of Theorem 2, the covariance function of $\zeta_{m,n}(\boldsymbol{\beta},t)$ can be calculated as

$$
\begin{aligned}
& \operatorname{cov}\left\{\zeta_{m,n}(\boldsymbol{\beta},t),\zeta_{m,n}(\boldsymbol{\alpha},s)\right\} \\
=\;& \mathrm{P}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}\le t, \boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}\le s)-\mathrm{P}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}\le t)\mathrm{P}(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}\le s).
\end{aligned}
$$

Therefore, it follows that the empirical processes $\zeta_{m,n}(\boldsymbol{\beta},t)$ converges in distribution to a Gaussian process with mean function $\sqrt{\tau(1-\tau)}\ell(\boldsymbol{\beta},t)$, and the covariance function given by (4). That is, under the local alternative, $\zeta_{m,n}(\boldsymbol{\beta},t)$

converges in distribution to $\zeta(\boldsymbol{\beta}, t) + \sqrt{\tau(1-\tau)}\ell(\boldsymbol{\beta}, t)$. Hence we have

$$\begin{aligned} & \frac{mn}{m+n}\widehat{T} \\ = \quad & 2\pi \iint \{\zeta_{m,n}(\boldsymbol{\beta}, t)\}^2 dH(\boldsymbol{\beta}, t) \\ \xrightarrow{d} \quad & 2\pi \iint \{\zeta(\boldsymbol{\beta}, t) + \sqrt{\tau(1-\tau)}\ell(\boldsymbol{\beta}, t)\}^2 dH(\boldsymbol{\beta}, t), \end{aligned}$$

which completes the proof.

□

S.5. Proof of Theorem 5

Since $\{\mathbf{z}_1^*, \mathbf{z}_2^*, \ldots, \mathbf{z}_{m+n}^*\}$ is a random permutation of $\{\mathbf{x}_1, \ldots, \mathbf{x}_n, \mathbf{y}_1, \ldots, \mathbf{y}_n\}$, conditional on the original sample, $\mathbf{x}_1^*, \ldots, \mathbf{x}_m^*, \mathbf{y}_1^*, \ldots, \mathbf{y}_n^*$ are asymptotically independently and identically distributed. The pooled distribution function is given by $m/(m+n)F_m + m/(m+n)G_n$. We define the empirical processes

$$\zeta_{m,n}^*(\boldsymbol{\beta}, t) = \sqrt{mn/(m+n)}\{U_m^*(\boldsymbol{\beta}, t) - V_n^*(\boldsymbol{\beta}, t)\}$$

where

$$\begin{aligned} U_m^*(\boldsymbol{\beta}, t) &= m^{-1}\sum_{i=1}^{m} I(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}_i^* \le t), \\ V_n^*(\boldsymbol{\beta}, t) &= n^{-1}\sum_{i=1}^{n} I(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{y}_i^* \le t). \end{aligned}$$

Therefore, according to the proof of Theorem 2, conditional on the original sample, the expectation of the empirical processes $\zeta_{m,n}^*(\boldsymbol{\beta}, t)$ is zero and the covariance function is given by

$$\begin{aligned} & \mathrm{cov}\left\{\zeta_{m,n}^*(\boldsymbol{\beta}, t), \zeta_{m,n}^*(\boldsymbol{\alpha}, s) \mid \mathbf{x}_1, \ldots, \mathbf{x}_m, \mathbf{y}_1, \ldots, \mathbf{y}_n\right\} \\ = \quad & \mathrm{P}_{m+n}\left(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{z} \le t, \boldsymbol{\alpha}^{\mathrm{T}}\mathbf{z} \le s\right) - \\ & \mathrm{P}_{m+n}\left(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{z} \le t\right) \mathrm{P}_{m+n}\left(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{z} \le s\right), \end{aligned}$$

where P_{m+n} is the pooled empirical probability, i.e.,

$$\begin{aligned} & \mathrm{P}_{m+n}\left(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{z} \le t, \boldsymbol{\alpha}^{\mathrm{T}}\mathbf{z} \le s\right) \\ = \quad & \frac{1}{m+n}\left\{\sum_{i=1}^{m} I(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}_i \le t, \boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x}_i \le s)\right. \\ & \left. + \sum_{i=1}^{n} I(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{y}_i \le t, \boldsymbol{\alpha}^{\mathrm{T}}\mathbf{y}_i \le s)\right\}. \end{aligned}$$

With the slutsky theorem, the empirical probability $\mathrm{P}_{m+n}\left(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{z} \le t, \boldsymbol{\alpha}^{\mathrm{T}}\mathbf{z} \le s\right)$ converges in probability to $\tau\mathrm{P}\left(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x} \le t, \boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x} \le s\right) + (1-\tau)\mathrm{P}\left(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{y} \le t, \boldsymbol{\alpha}^{\mathrm{T}}\mathbf{y} \le s\right)$. Subsequently, we have

$$\begin{aligned} & \mathrm{cov}\left\{\zeta_{m,n}^*(\boldsymbol{\beta}, t), \zeta_{m,n}^*(\boldsymbol{\alpha}, s) \mid \mathbf{x}_1, \ldots, \mathbf{x}_m, \mathbf{y}_1, \ldots, \mathbf{y}_n\right\} \\ = \; & \tau\mathrm{P}\left(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x} \le t, \boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x} \le s\right) + (1-\tau)\mathrm{P}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{y} \le t, \boldsymbol{\alpha}^{\mathrm{T}}\mathbf{y} \le s) \\ & - \left\{\tau\mathrm{P}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x} \le t) + (1-\tau)\mathrm{P}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{y} \le t)\right\} \\ & \quad \left\{\tau\mathrm{P}(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x} \le s) + (1-\tau)\mathrm{P}(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{y} \le s)\right\} + o_p(1). \end{aligned}$$

Therefore, by denoting $\zeta^*(\boldsymbol{\beta}, t)$ the Gaussian process with mean function zero and the covariance function $\mathrm{cov}\left\{\zeta^*(\boldsymbol{\beta}, t), \zeta^*(\boldsymbol{\alpha}, s)\right\}$ is given by

$$\begin{aligned} & \tau\mathrm{P}\left(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x} \le t, \boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x} \le s\right) + (1-\tau)\mathrm{P}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{y} \le t, \boldsymbol{\alpha}^{\mathrm{T}}\mathbf{y} \le s) \\ & - \left\{\tau\mathrm{P}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x} \le t) + (1-\tau)\mathrm{P}(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{y} \le t)\right\}\left\{\tau\mathrm{P}(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{x} \le s)\right. \\ & \qquad \left. + (1-\tau)\mathrm{P}(\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{y} \le s)\right\}. \end{aligned}$$

we have conditional on the original sample, the empirical processes $\zeta_{m,n}^*(\boldsymbol{\beta}, t)$ converges in distribution to a Gaussian process whose mean function is zero and covariance function is asymptotically the same as $\zeta^*(\boldsymbol{\beta}, t)$. According to (Zhu & Neuhaus, 2003), we have the conditional distribution of $mn/(m+n)\widehat{T}^*$ and $2\pi \iint \{\zeta^*(\boldsymbol{\beta}, t)\}^2 dH(\boldsymbol{\beta}, t)$ are asymptotically the same. This further yields the assertion of this theorem because the limiting distribution is continuous.

□

References

Van Der Vaart, A. W. and Wellner, J. A. *Weak convergence and empirical processes*. Springer, 1996.

Zhu, L.-X. and Neuhaus, G. Conditional tests for elliptical symmetry. *Journal of Multivariate Analysis*, 84(2):284–298, 2003.