Supplement to Estimation of a Two-component Mixture Model with Applications to Multiple Testing

Rohit Kumar Patra and Bodhisattva Sen

Columbia University, New York, USA

12. Identifiability of F_s

In this section we continue the discussion on the identifiability of F_s. First, we give some remarks to illustrate Lemmas 2.3 and 2.4.

Remark 12.1. *We consider mixtures of Poisson and binomial distributions to illustrate Lemma 2.3. If F_s is Poisson(λ_s) and F_b is Poisson(λ_b), then*

$$\inf_{x \in d(F_b)} \frac{J_{F_s(x)}}{J_{F_b(x)}} = \inf_{k \in \mathbb{N} \cup \{0\}} \frac{\lambda_s^k \exp(-\lambda_s)}{\lambda_b^k \exp(-\lambda_b)} = \exp(\lambda_b - \lambda_s) \inf_{k \in \mathbb{N} \cup \{0\}} \left(\frac{\lambda_s}{\lambda_b}\right)^k.$$

By an application of Lemma 2.3, we have if $\lambda_s < \lambda_b$ then $\alpha_0 = \alpha$; otherwise $\alpha_0 = \alpha(1 - \exp(\lambda_b - \lambda_s))$.

In the case of a binomial mixture, i.e., $F_s = Bin(n, p_s)$ and $F_b = Bin(n, p_b)$,

$$\alpha_0 = \begin{cases} \alpha \left[1 - (\frac{1-p_s}{1-p_b})^n\right], & p_s \geq p_b, \\ \alpha \left[1 - (\frac{p_s}{p_b})^n\right], & p_s < p_b. \end{cases}$$

Remark 12.2. *If F_s is $N(\mu_s, \sigma_s^2)$ and F_b $(\neq F_s)$ is $N(\mu_b, \sigma_b^2)$ then it can be easily shown that the problem is identifiable if and only if $\sigma_s \leq \sigma_b$. When $\sigma_s > \sigma_b$, the model is not identifiable, an application of Lemma 2.4 gives $\alpha_0 = \alpha\big[1 - (\sigma_b/\sigma_s)\exp\big(-\sigma_s\sigma_b(\mu_b - \mu_s)^2/2\big)\big]$. Thus, α_0 increases to α as $|\mu_s - \mu_b|$ tends to infinity. It should be noted that the problem is actually identifiable if we restrict ourselves to the parametric family of a two-component Gaussian mixture model.*

Remark 12.3. *Now consider a mixture of exponential random variables, i.e., F_s is $E(a_s, \sigma_s)$ and F_b $(\neq F_s)$ is $E(a_b, \sigma_b)$, where $E(a, \sigma)$ is the distribution that has the density $(1/\sigma)\exp(-(x-a)/\sigma)\mathbf{1}_{(a,\infty)}(x)$. In this case, the problem is identifiable if $a_s > a_b$, as this implies the support of F_s is a proper subset of the support of F_b. But when $a_s \leq a_b$, the problem is identifiable if and only if $\sigma_s \leq \sigma_b$.*

Remark 12.4. *It is also worth pointing out that even in cases where the problem is not identifiable the difference between the true mixing proportion α and the estimand α_0 may be very small. Consider the hypothesis test $H_0 : \theta = 0$ versus $H_1 : \theta \neq 0$ for the model $N(\theta, 1)$ with test statistic $\bar{X}$. The density of the p-values under θ is*

$$f_\theta(p) = \frac{1}{2} e^{-m\theta^2/2} \big[e^{-\sqrt{m}\theta^2 \Phi^{-1}(1-p/2)} + e^{\sqrt{m}\theta^2 \Phi^{-1}(1-p/2)}\big],$$

where m is the sample size. Here $f_\theta(1) = e^{-m\theta^2/2} > 0$, so the model is not identifiable. As F_b is uniform, it can be easily verified that $\alpha_0 = \alpha - \alpha \inf_p f_\theta(p)$. However, as the value of f_θ decreases exponentially with m, in many practical situations, where m is not too small, the difference between α and α_0 will be negligible.

In the following lemma, we try to find the relationship between α and α_0 when F is a general CDF.

LEMMA 12.5. *Suppose that*

$$F = \kappa F^{(a)} + (1-\kappa)F^{(d)}, \tag{101}$$

where $F^{(a)}$ is an absolutely continuous CDF and $F^{(d)}$ is a piecewise constant CDF, for some $\kappa \in (0,1)$. Then

$$\alpha_0 = \alpha - \min\left\{\frac{\alpha\kappa_s - \alpha_0^{(a)}\kappa}{\kappa_b}, \frac{\alpha(1-\kappa_s) - \alpha_0^{(d)}(1-\kappa)}{(1-\kappa_b)}\right\},$$

where $\alpha_0^{(a)}$ and $\alpha_0^{(d)}$ are defined as in (4), but with $\{F^{(a)}, F_b^{(a)}\}$ and $\{F^{(d)}, F_b^{(d)}\}$, respectively (instead of $\{F, F_b\}$). Similarly, κ_s and κ_b are defined as in (101), but for F_s and F_b, respectively.

PROOF. From the definition of κ_s and κ_b, we have $F_s = \kappa_s F_s^{(a)} + (1-\kappa_s)F_s^{(d)}$, and $F_b = \kappa_b F_b^{(a)} + (1-\kappa_b)F_b^{(d)}$. Thus from (1), we get

$$F = \alpha\kappa_s F_s^{(a)} + (1-\alpha)\kappa_b F_b^{(a)} + \alpha(1-\kappa_s)F_s^{(d)} + (1-\alpha)(1-\kappa_s)F_b^{(d)}.$$

Now using the definition of κ, we see that $\kappa = \alpha\kappa_s + (1-\alpha)\kappa_b$, $1-\kappa = \alpha(1-\kappa_s) + (1-\alpha)(1-\kappa_b)$. If we write

$$F^{(a)} = \alpha^{(a)} F_s^{(a)} + (1-\alpha^{(a)})F_b^{(a)},$$

it can easily seen that $\alpha^{(a)} = \frac{\alpha\kappa_s}{\kappa}$; and similarly, $\alpha^{(d)} = \frac{\alpha(1-\kappa_s)}{1-\kappa}$. Then, we can find $\alpha_0^{(d)}$ and $\alpha_0^{(a)}$ as in Lemmas 2.3 and 2.4, respectively. Note that

$$\begin{aligned}
& \sup\{0 \le \epsilon \le 1 : \alpha F_s - \epsilon F_b \text{ is a sub-CDF}\} \\
= \; & \sup\left\{0 \le \epsilon \le 1 : \alpha(\kappa_s F_s^{(a)} + (1-\kappa_s)F_s^{(d)}) - \epsilon(\kappa_b F_b^{(a)} + (1-\kappa_b)F_b^{(d)}) \text{ is a sub-CDF}\right\} \\
= \; & \sup\left\{0 \le \epsilon \le 1 : \text{both } \alpha\kappa_s F_s^{(a)} - \epsilon\kappa_b F_b^{(a)}, \alpha(1-\kappa_s)F_s^{(d)} - \epsilon(1-\kappa_b)F_b^{(d)} \text{ are sub-CDFs}\right\} \\
= \; & \min\left(\sup\left\{0 \le \epsilon \le 1 : \alpha\kappa_s F_s^{(a)} - \epsilon\kappa_b F_b^{(a)} \text{ is a sub-CDF}\right\},\right. \\
& \qquad \left.\sup\left\{0 \le \epsilon \le 1 : \alpha(1-\kappa_s)F_s^{(d)} - \epsilon(1-\kappa_b)F_b^{(d)} \text{ is a sub-CDF}\right\}\right) \\
= \; & \min\left(\frac{\alpha\kappa_s}{\kappa_b} \operatorname{essinf} \frac{f_s^{(a)}}{f_b^{(a)}}, \frac{\alpha(1-\kappa_s)}{(1-\kappa_b)} \inf_{x \in d(F_b^{(d)})} \frac{J_{F_s^{(d)}}(x)}{J_{F_b^{(d)}}(x)}\right) \\
= \; & \min\left(\frac{(\alpha\kappa_s - \alpha_0^{(a)}\kappa)}{\kappa_b}, \frac{(\alpha(1-\kappa_s) - \alpha_0^{(d)}(1-\kappa))}{(1-\kappa_b)}\right),
\end{aligned}$$

where J_G and $d(J_G)$ are defined before Lemma 2.3 and we use the notion that $\frac{0}{0} = 1$. Hence, by (5) the result follows.

Lemma 2.5 is now a corollary of this result.

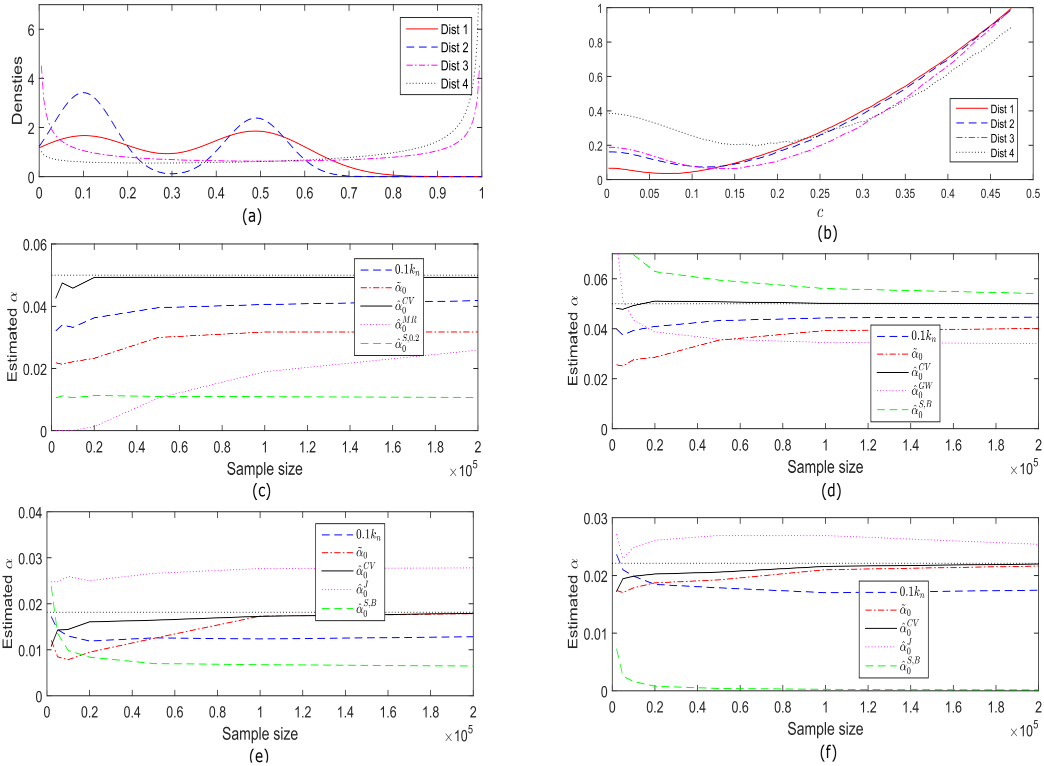

Fig. 7. Plots comparing the performance of $\hat{\alpha}_0^{c_n}$, $\hat{\alpha}_0^{CV}$, and $\tilde{\alpha}_0$; (a) density functions for four different choices of F_s; (b) plot of the average of $\sum_{k=1}^{K}\int(\mathbb{F}_n^k - \hat{F}^k)^2 d\mathbb{F}_n^k$ (see (13) of the main paper), as a function of c, computed over 500 independent samples of size 50000 corresponding to Dist 1-4; (c)-(f) gives the means of different competing estimators of α_0, computed over 500 independent samples for Dist 1-4 respectively (in each figure the horizontal dotted black line denotes the true α_0).

13. Performance comparison of $\hat{\alpha}_0^{c_n}$, $\hat{\alpha}_0^{CV}$, and $\tilde{\alpha}_0$

In Figs. 7 and 8 we present further simulation experiments to investigate the finite sample performance of $\hat{\alpha}_0^{c_n}$, $\hat{\alpha}_0^{CV}$, and $\tilde{\alpha}_0$ across different simulation scenarios. In each setting we also include the performance of the best performing competing estimators discussed in Section 8.2.

14. Detection of sparse heterogeneous mixtures

In this section we draw a connection between the lower confidence bound developed in Section 4 and the *Higher Criticism* method of Donoho and Jin (2004) for detection of sparse heterogeneous mixtures. The detection of heterogeneity in sparse models arises in many applications, e.g., detection of a disease outbreak (see Kulldorff et al. (2005)) or early detection of bioweapons use (see Donoho and Jin (2004)). Generally, in large scale multiple testing problems, when the non-null effect is sparse it is important to detect the existence of non-null effects (see Cai et al. (2007)).

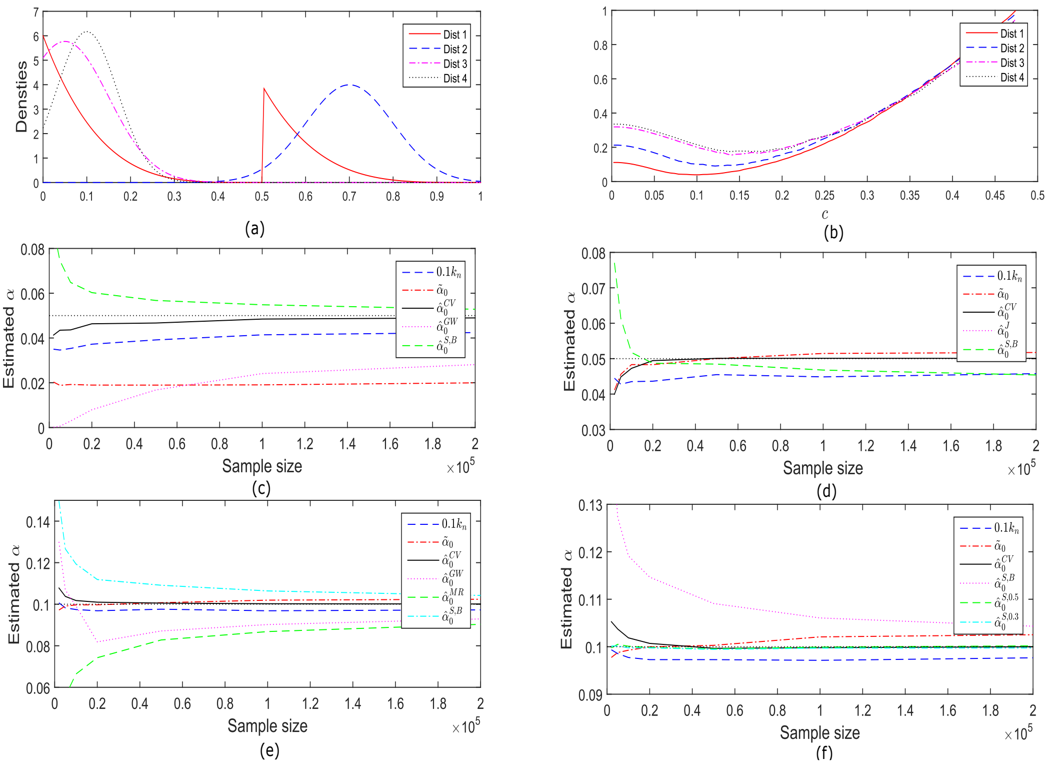

Fig. 8. Plots comparing the performance of $\hat{\alpha}_0^{c_n}$, $\hat{\alpha}_0^{CV}$, and $\tilde{\alpha}_0$; (a) density functions for four different choices of F_s; (b) plot of the average of $\sum_{k=1}^{K} \int (\mathbb{F}_n^k - \hat{F}^k)^2 d\mathbb{F}_n^k$ (see (13) of the main paper), as a function of c, computed over 500 independent samples of size 50000 corresponding to Dist 1-4; (c)-(f) gives the means of different competing estimators of α_0, computed over 500 independent samples for Dist 1-4 respectively (in each figure the horizontal dotted black line denotes the true α_0).

Donoho and Jin (2004) consider n i.i.d. data from one of the two possible situations:

$$H_0 : X_i \sim F_b, \quad 1 \le i \le n,$$
$$H_1^{(n)} : X_i \sim F^n := \alpha_n F_{n,s} + (1-\alpha_n) F_b, \quad 1 \le i \le n,$$

where $\alpha_n \sim n^{-\lambda}$ and $F_{n,s}$ is such that $d(F_{n,s}, F_b)$ is bounded away from 0. In Donoho and Jin (2004) the main focus is on testing H_0, i.e., $\alpha_n = 0$. We can test this hypothesis by rejecting H_0 when $\hat{\alpha}_L > 0$. The following lemma shows that indeed this yields a valid testing procedure for $\lambda < 1/2$.

THEOREM 14.1. *If $\alpha_n \sim n^{-\lambda}$, for $\lambda < 1/2$, then $P_{H_0}(\text{Reject } H_0) = \beta$ and $P_{H_1^{(n)}}(\hat{\alpha}_L > 0) \to 1$ as $n \to \infty$.*

PROOF. Note that $\{\hat{\alpha}_L > 0\}$ is equivalent to $\{c_n \le \sqrt{n} d_n(\mathbb{F}_n, F_b)\}$ which shows that

$$\begin{aligned} c_n &\le \sqrt{n} d_n(\mathbb{F}_n, (1-\alpha_n)F_b + \alpha_n F_{n,s}) + \sqrt{n} d_n(\alpha_n F_b, \alpha_n F_{n,s}) \\ &= \sqrt{n} d_n(\mathbb{F}_n, F^n) + \alpha_n \sqrt{n} d_n(F_{n,s}, F_b), \end{aligned}$$

where c_n is chosen as in Theorem 4.1. It is easy to see that $\sqrt{n} d_n(\mathbb{F}_n, F^n)$ is $O_P(1)$ and $\alpha_n \sqrt{n} d_n(F_{n,s}, F_b) \to \infty$, for $\lambda < 1/2$, which shows that $P_{H_1^{(n)}}(\hat{\alpha}_L > 0) \to 1$. It can be easily seen that $P_{H_0}(\hat{\alpha}_L > 0) = P_{H_0}(\text{Reject } H_0) = \beta$.

15. Proofs of theorems and lemmas in the main paper

15.1. Proof of Lemma 2.3

When $d(F_b) \not\subset d(F_s)$, there exists a $x \in d(F_b) - d(F_s)$, i.e., there exists a x which satisfies $F_b(x) - F_b(x-) > 0$ and $F_s(x) - F_s(x-) = 0$. Then for all $\epsilon > 0$, $F_s(x-) - \epsilon F_b(x-) > F_s(x) - \epsilon F_b(x)$. This shows that $F_s - \epsilon F_b$ cannot be a sub-CDF, and hence by Lemma 2.2 the model is identifiable. Now let us assume that $d(F_b) \subset d(F_s)$.

$$\begin{aligned} \{0 \le \epsilon \le 1 : \alpha F_s - \epsilon F_b \text{ is a sub-CDF}\} &= \{0 \le \epsilon \le 1 : \alpha J_{F_s}(x) - \epsilon J_{F_b}(x) \ge 0, \ \forall x \in d(J_{F_b})\} \\ &= \left\{0 \le \epsilon \le 1 : \frac{J_{F_s}(x)}{J_{F_b}(x)} \ge \frac{\epsilon}{\alpha}, \ \forall x \in d(J_{F_b})\right\} \\ &= \left\{0 \le \epsilon \le 1 : \inf_{x \in d(F_b)} \frac{J_{F_s}(x)}{J_{F_b}(x)} \ge \frac{\epsilon}{\alpha}\right\}. \end{aligned}$$

Therefore, using (5), we get the desired result.

15.2. Proof of Lemma 2.4

From (5), we have

$$\begin{aligned} \alpha_0 &= \alpha - \sup\{0 \le \epsilon \le 1 : \alpha F_s - \epsilon F_b \text{ is a sub-CDF}\} \\ &= \alpha - \sup\{0 \le \epsilon \le 1 : \alpha f_s(x) - \epsilon f_b(x) \ge 0 \text{ almost every } x\} \\ &= \alpha - \sup\left\{0 \le \epsilon \le 1 : \alpha \frac{f_s}{f_b}(x) \ge \epsilon \text{ almost every } x\right\} \\ &= \alpha \left\{1 - \operatorname{essinf} \frac{f_s}{f_b}\right\}. \end{aligned}$$

15.3. Proof of Theorem 3.1

Without loss of generality, we can assume that F_b is the uniform distribution on $(0,1)$ and, for clarity, in the following we write U instead of F_b. Let us define

$$\begin{aligned} A &:= \left\{\gamma \in (0,1] : \frac{F-(1-\gamma)U}{\gamma} \text{ is a valid CDF}\right\}, \\ A^Y &:= \left\{\gamma \in (0,1] : \frac{G-(1-\gamma)U\circ\Psi}{\gamma} \text{ is a valid CDF}\right\}. \end{aligned}$$

Since $\alpha_0 = \inf A$, and $\alpha_0^Y = \inf A^Y$ for the first part of the theorem it is enough to show that $A = A^Y$. Let us first show that $A^Y \subset A$. Suppose $\eta \in A^Y$. We first show that $(F-(1-\eta)U)/\eta$ is a non-decreasing function. For all $t_1 \le t_2$, we have that

$$\frac{G(t_1)-(1-\eta)U(\Psi(t_1))}{\eta} \le \frac{G(t_2)-(1-\eta)U(\Psi(t_2))}{\eta}.$$

Let $y_1 \le y_2$. Then,

$$\frac{G(\Psi^{-1}(y_1))-(1-\eta)U(\Psi(\Psi^{-1}(y_1)))}{\eta} \le \frac{G(\Psi^{-1}(y_2))-(1-\eta)U(\Psi(\Psi^{-1}(y_2)))}{\eta},$$

since $y_1 \le y_2 \Rightarrow \Psi^{-1}(y_1) \le \Psi^{-1}(y_2)$. However, as Ψ is continuous, $\Psi(\Psi^{-1}(y)) = y$ and $G(\Psi^{-1}(y)) = \alpha F_s(y) + (1-\alpha)U(y) = F(y)$. Hence, we have

$$\frac{F(y_1)-(1-\eta)U(y_1)}{\eta} \le \frac{F(y_2)-(1-\eta)U(y_2)}{\eta}.$$

As F and U are CDFs, it is easy to see that $\lim_{x\to-\infty}(F(x)-(1-\eta)U(x))/\eta = 0$, $\lim_{x\to\infty}(F(x)-(1-\eta)U(x))/\eta = 1$ and $(F-(1-\eta)U)/\eta$ is a right continuous function. Hence, for $\eta \in A^Y$, $(F-(1-\eta)U)/\eta$ is a CDF and thus, $\eta \in A$. We can similarly prove $A \subset A^Y$. Therefore, $A = A^Y$ and $\alpha_0 = \alpha_0^Y$.

Note that

$$\gamma d_n(\hat{F}_{s,n}^\gamma, \check{F}_{s,n}^\gamma) = \min_{W\in\mathcal{F}} \frac{1}{n}\sum_{i=1}^n \{W(X_i) - \hat{F}_{s,n}^\gamma(X_i)\}^2,$$

where $\mathcal{F}$ is the class of all CDFs. For the second part of theorem it is enough to show that

$$\min_{W\in\mathcal{F}} \frac{1}{n}\sum_{i=1}^n \{W(X_i) - \hat{F}_{s,n}^\gamma(X_i)\}^2 = \min_{B\in\mathcal{F}} \frac{1}{n}\sum_{i=1}^n \{B(Y_i) - \hat{G}_{s,n}^\gamma(Y_i)\}^2.$$

First note that

$$\begin{aligned} \mathbb{G}_n(y) &= \frac{1}{n}\sum_{i=1}^n \mathbf{1}\{\Psi^{-1}(X_i) \le y\} \\ &= \frac{1}{n}\sum_{i=1}^n \mathbf{1}\{X_i \le \Psi(y)\} \\ &= \mathbb{F}_n(\Psi(y)). \end{aligned}$$

Thus, from the definition of $\hat{G}^{\gamma}_{s,n}$, we have

$$\begin{aligned} \hat{G}^{\gamma}_{s,n}(Y_i) &= \frac{\mathbb{F}_n(\Psi(Y_i)) - (1-\gamma)U(\Psi(Y_i))}{\gamma} \\ &= \frac{\mathbb{F}_n(X_i) - (1-\gamma)U(X_i)}{\gamma} = \hat{F}^{\gamma}_{s,n}(X_i). \end{aligned}$$

Therefore,

$$\begin{aligned} \frac{1}{n}\sum_{i=1}^{n}\{B(Y_i) - \hat{G}^{\gamma}_{s,n}(Y_i)\}^2 &= \frac{1}{n}\sum_{i=1}^{n}\{B(Y_i) - \hat{F}^{\gamma}_{s,n}(X_i)\}^2 \\ &= \frac{1}{n}\sum_{i=1}^{n}\{B(\Psi^{-1}(X_i)) - \hat{F}^{\gamma}_{s,n}(X_i)\}^2 \\ &= \frac{1}{n}\sum_{i=1}^{n}\{W(X_i) - \hat{F}^{\gamma}_{s,n}(X_i)\}^2, \end{aligned}$$

where $W(x) := B(\Psi^{-1}(x))$. W is a valid CDF as Ψ^{-1} is non-decreasing.

15.4. Proof of Theorem 3.4

We need to show that $P(|\hat{\alpha}_0^{c_n} - \alpha_0| > \epsilon) \to 0$ for any $\epsilon > 0$. Let us first show that

$$P(\hat{\alpha}_0^{c_n} - \alpha_0 < -\epsilon) \to 0.$$

The statement is obviously true if $\alpha_0 \le \epsilon$. So let us assume that $\alpha_0 > \epsilon$. Suppose $\hat{\alpha}_0^{c_n} - \alpha_0 < -\epsilon$, i.e., $\hat{\alpha}_0^{c_n} < \alpha_0 - \epsilon$. Then by the definition of $\hat{\alpha}_0^{c_n}$ and the convexity of A_n, we have $(\alpha_0 - \epsilon) \in A_n$ (as A_n is a convex set in $[0,1]$ with $1 \in A_n$ and $\hat{\alpha}_0^{c_n} \in A_n$), and thus

$$d_n(\hat{F}^{\alpha_0-\epsilon}_{s,n}, \check{F}^{\alpha_0-\epsilon}_{s,n}) \le \frac{c_n}{\sqrt{n}(\alpha_0 - \epsilon)}. \tag{102}$$

But by (10) the left-hand side of (102) goes to a non-zero constant in probability. Hence, if $c_n/\sqrt{n} \to 0$,

$$P(\hat{\alpha}_0^{c_n} - \alpha_0 < -\epsilon) \le P\left(d_n(\hat{F}^{\alpha_0-\epsilon}_{s,n}, \check{F}^{\alpha_0-\epsilon}_{s,n}) \le \frac{c_n}{\sqrt{n}(\alpha_0 - \epsilon)}\right) \to 0.$$

This completes the proof of the first part of the claim.

Now suppose that $\hat{\alpha}_0^{c_n} - \alpha_0 > \epsilon$. Then,

$$\begin{aligned} \hat{\alpha}_0^{c_n} - \alpha_0 > \epsilon \quad &\Rightarrow \quad \sqrt{n} d_n(\hat{F}^{\alpha_0+\epsilon}_{s,n}, \check{F}^{\alpha_0+\epsilon}_{s,n}) \ge \frac{c_n}{\alpha_0 + \epsilon} \\ &\Rightarrow \quad \sqrt{n} d_n(\mathbb{F}_n, F) \ge c_n. \end{aligned}$$

The first implication follows from the definition of $\hat{\alpha}_0^{c_n}$, while the second implication is true by Lemma 3.2. The right-hand side of the last inequality is (asymptotically similar to) the Cramér–von Mises statistic for which the asymptotic distribution is well-known and thus if $c_n \to \infty$ the result follows.

15.5. Proof of Theorem 3.6

As the proof of this result is slightly involved we break it into a number of lemmas (whose proofs are provided later in this sub-section) and give the main arguments below.

We need to show that given any $\epsilon > 0$, we can find an $M > 0$ and $n_0 \in \mathbb{N}$ (depending on ϵ) for which $\sup_{n>n_0} P(r_n|\hat{\alpha}_0^{c_n} - \alpha_0| > M) \leq \epsilon$.

LEMMA 15.1. *If $c_n \to \infty$, then for any $M > 0$, $\sup_{n>n_0} P\left(r_n(\hat{\alpha}_0^{c_n} - \alpha_0) > M\right) < \epsilon$, for large enough $n_0 \in \mathbb{N}$.*

Finding an r_n such that $P\left(r_n(\hat{\alpha}_0^{c_n} - \alpha_0) < -M\right) < \epsilon$ for large enough n is more complicated. We start with some notation. Let $\mathcal{F}$ be the class of all CDFs and $\mathbb{H}$ be the Hilbert space $L_2(F) := \{f : \mathbb{R} \to \mathbb{R} | \int f^2 dF < \infty\}$. For a closed convex subset $\mathcal{K}$ of $\mathbb{H}$ and $h \in \mathbb{H}$, we define the projection of h onto $\mathcal{K}$ as

$$\Pi(h|\mathcal{K}) := \operatorname*{arg\,min}_{f \in \mathcal{K}} d(f, h), \tag{103}$$

where d stands for the $L_2(F)$ distance, i.e., if $g, h \in \mathbb{H}$, then $d^2(g, h) = \int (g - h)^2 dF$. We define the tangent cone of $\mathcal{F}$ at $f_0 \in \mathcal{F}$, as

$$T_{\mathcal{F}}(f_0) := \{\lambda(f - f_0) : \lambda \geq 0, f \in \mathcal{F}\}. \tag{104}$$

For any $H \in \mathcal{F}$ and $\gamma > 0$, let us define

$$\hat{H}^\gamma := \frac{H - (1-\gamma)F_b}{\gamma}, \quad \check{H}_n^\gamma := \operatorname*{arg\,min}_{G \in \mathcal{F}} \gamma d_n(\hat{H}^\gamma, G), \quad \text{and} \quad \bar{H}_n^\gamma := \operatorname*{arg\,min}_{G \in \mathcal{F}} \gamma d(\hat{H}^\gamma, G).$$

For $H = \mathbb{F}_n$ and $\gamma = \alpha_0$ we define the three quantities above and call them $\hat{F}_{s,n}^{\alpha_0}$, $\check{F}_{s,n}^{\alpha_0}$, and $\bar{F}_{s,n}^{\alpha_0}$ respectively. Note that

$$P\left(r_n(\hat{\alpha}_0^{c_n} - \alpha_0) < -M\right) = P(\sqrt{n}\gamma_n \; d_n(\hat{F}_{s,n}^{\gamma_n}, \check{F}_{s,n}^{\gamma_n}) < c_n), \tag{105}$$

where $\gamma_n = \alpha_0 - M/r_n$. To study the limiting behavior of $d_n(\hat{F}_{s,n}^{\gamma_n}, \check{F}_{s,n}^{\gamma_n})$ we break it as the sum of $d_n(\hat{F}_{s,n}^{\gamma_n}, \check{F}_{s,n}^{\gamma_n}) - d(\hat{F}_{s,n}^{\gamma_n}, \bar{F}_{s,n}^{\gamma_n})$ and $d(\hat{F}_{s,n}^{\gamma_n}, \bar{F}_{s,n}^{\gamma_n})$. The following two lemmas (proved in Sections 15.5.2 and 15.5.3 respectivley) give the asymptotic behavior of the two terms. The proof of Lemma 15.3 uses the functional delta method (cf. Theorem 20.8 of Van der Vaart (1998)) for the projection operator; see Theorem 1 of Fils-Villetard et al. (2008).

LEMMA 15.2. *If $\sqrt{n}/r_n^2 \to 0$, then $U_n := \sqrt{n}\gamma_n d_n(\hat{F}_{s,n}^{\gamma_n}, \check{F}_{s,n}^{\gamma_n}) - \sqrt{n}\gamma_n d(\hat{F}_{s,n}^{\gamma_n}, \bar{F}_{s,n}^{\gamma_n}) \xrightarrow{P} 0$.*

LEMMA 15.3. *If $c_n \to \infty$, then*

$$\frac{\sqrt{n}\gamma_n}{c_n M} d(\hat{F}_{s,n}^{\gamma_n}, \bar{F}_{s,n}^{\gamma_n}) \xrightarrow{P} \left\{\int V^2 dF\right\}^{1/2} > 0$$

where

$$V := (F_s^{\alpha_0} - F_b) - \Pi\left(F_s^{\alpha_0} - F_b | T_{\mathcal{F}}(F_s^{\alpha_0})\right) \neq 0$$

and

$$F_s^{\alpha_0} := \frac{F - (1-\alpha_0)F_b}{\alpha_0}. \tag{106}$$

Using (105), and the notation introduced in the above two lemmas we see that

$$P\left(r_n(\hat{\alpha}_0^{c_n} - \alpha_0) < -M\right) \;=\; P\left(\frac{1}{c_n}U_n + \frac{\sqrt{n}\gamma_n}{c_n} d(\hat{F}_{s,n}^{\gamma_n}, \bar{F}_{s,n}^{\gamma_n}) < 1\right). \tag{107}$$

However, $U_n \xrightarrow{P} 0$ (by Lemma 15.2) and $\frac{\sqrt{n}\gamma_n}{c_n M} d(\hat{F}_{s,n}^{\gamma_n}, \bar{F}_{s,n}^{\gamma_n}) \xrightarrow{P} \int V^2 dF$ (by Lemma 15.3). The result now follows from (107), by taking a large enough M.

15.5.1. Proof of Lemma 15.1

Note that

$$\begin{aligned} P(r_n(\hat{\alpha}_0^{c_n} - \alpha_0) > M) \le P\left(\hat{\alpha}_0^{c_n} > \alpha_0\right) &= P\left(\sqrt{n}\alpha_0 d_n(\hat{F}_{s,n}^{\alpha_0}, \check{F}_{s,n}^{\alpha_0}) > c_n\right) \\ &\le P\left(\sqrt{n}\alpha_0 d_n(\hat{F}_{s,n}^{\alpha_0}, F_s^{\alpha_0}) > c_n\right) \\ &= P\left(\sqrt{n} d_n(\mathbb{F}_n, F) > c_n\right) \to 0, \end{aligned}$$

as $c_n \to \infty$, since $\sqrt{n} d_n(\mathbb{F}_n, F) = O_P(1)$. Therefore, the result holds for sufficiently large n.

15.5.2. Proof of Lemma 15.2

It is enough to show that

$$W_n := n\gamma_n^2 d_n^2(\hat{F}_{s,n}^{\gamma_n}, \check{F}_{s,n}^{\gamma_n}) - n\gamma_n^2 d^2(\hat{F}_{s,n}^{\gamma_n}, \bar{F}_{s,n}^{\gamma_n}) \;\xrightarrow{P}\; 0, \tag{108}$$

since $U_n^2 \le |W_n|$. Note that

$$\begin{aligned} \check{F}_{s,n}^{\gamma_n} &= \arg\min_{G \in \mathcal{F}} d_n(\mathbb{F}_n, \gamma_n G + (1-\gamma_n)F_b), \\ \bar{F}_{s,n}^{\gamma_n} &= \arg\min_{G \in \mathcal{F}} d(\mathbb{F}_n, \gamma_n G + (1-\gamma_n)F_b). \end{aligned}$$

For each positive integer n and $c > 0$, we introduce the following classes of functions:

$$\begin{aligned} \mathcal{G}_c(n) &= \left\{\sqrt{n}(G - (1-\gamma_n)F_b - \gamma_n \check{G}_n^{\gamma_n})^2 : G \in \mathcal{F},\ \|G - F\| < \frac{c}{\sqrt{n}}\right\}, \\ \mathcal{H}_c(n) &= \left\{\sqrt{n}(H - (1-\gamma_n)F_b - \gamma_n \bar{H}_n^{\gamma_n})^2 : H \in \mathcal{F},\ \|H - F\| < \frac{c}{\sqrt{n}}\right\}. \end{aligned}$$

Let us also define

$$D_n := \sup_{t \in \mathbb{R}} \sqrt{n}|\mathbb{F}_n(t) - F(t)| = \|\mathbb{F}_n - F\|.$$

From the definition of the minimisers $\check{F}_{s,n}^{\gamma_n}$ and $\bar{F}_{s,n}^{\gamma_n}$, we see that

$$\begin{aligned} \gamma_n^2\ |d_n^2(\hat{F}_{s,n}^{\gamma_n}, \check{F}_{s,n}^{\gamma_n}) - d^2(\hat{F}_{s,n}^{\gamma_n}, \bar{F}_{s,n}^{\gamma_n})| \le \max\big\{&|(d_n^2 - d^2)(\mathbb{F}_n, \gamma_n \check{F}_{s,n}^{\gamma_n} + (1-\gamma_n)F_b)|\,, \\ &|(d_n^2 - d^2)(\mathbb{F}_n, \gamma_n \bar{F}_{s,n}^{\gamma_n} + (1-\gamma_n)F_b)|\big\}. \end{aligned} \tag{109}$$

Observe that

$$n\gamma_n^2\ [(d_n^2 - d^2)(\mathbb{F}_n, \gamma_n \check{F}_{s,n}^{\gamma_n} + (1-\gamma_n)F_b)] = \sqrt{n}(\mathbb{P}_n - P)[g_n] = \nu_n(g_n),$$

where $g_n := \sqrt{n}\{\mathbb{F}_n - \gamma_n \check{F}_{s,n}^{\gamma_n} - (1-\gamma_n)F_b\}^2$, $\mathbb{P}_n$ denotes the empirical measure of the data, and $\nu_n := \sqrt{n}(\mathbb{P}_n - P)$ denotes the usual empirical process. Similarly,

$$n\gamma_n^2\ [(d_n^2 - d^2)(\mathbb{F}_n, \gamma_n \bar{F}_{s,n}^{\gamma_n} + (1-\gamma_n)F_b)] = \sqrt{n}(\mathbb{P}_n - P)[h_n] = \nu_n(h_n),$$

where $h_n := \sqrt{n}\{\mathbb{F}_n - \gamma_n \bar{F}_{s,n}^{\gamma_n} - (1-\gamma_n)F_b\}^2$. Thus, combining (108), (109) and the above two displays, we get, for any $\delta > 0$,

$$P(|W_n| > \delta) \leq P\left(|\nu_n(g_n)| > \delta\right) + P\left(|\nu_n(h_n)| > \delta\right). \tag{110}$$

The first term in the right hand side of (110) can be bounded above as

$$\begin{aligned}
P(|\nu_n(g_n)| > \delta) &= P(|\nu_n(g_n)| > \delta, g_n \in \mathcal{G}_c(n)) + P(|\nu_n(g_n)| > \delta, g_n \notin \mathcal{G}_c(n)) \\
&\leq P(|\nu_n(g_n)| > \delta, g_n \in \mathcal{G}_c(n)) + P(g_n \notin \mathcal{G}_c(n)) \\
&\leq P\left(\sup_{g \in \mathcal{G}_c(n)} |\nu_n(g)| > \delta\right) + P(g_n \notin \mathcal{G}_c(n)) \\
&\leq \frac{1}{\delta} E\left(\sup_{g \in \mathcal{G}_c(n)} |\nu_n(g)|\right) + P(g_n \notin \mathcal{G}_c(n)) \\
&\leq J_{[\,]} \frac{P[G_{c,n}^2]}{\delta} + P(g_n \notin \mathcal{G}_c(n)),
\end{aligned} \tag{111}$$

where $G_{c,n} := 6c^2/\sqrt{n} + 16\sqrt{n}\frac{M^2}{r_n^2}\|F_s^{\alpha_0} - F_b\|^2$ is an envelope for $\mathcal{G}_c(n)$ and $J_{[\,]}$ is a constant. Note that to derive the last inequality, we have used the maximal inequality in Corollary (4.3) of Pollard (1989); the class $\mathcal{G}_c(n)$ is "manageable" in the sense of Pollard (1989) (as a consequence of equation (2.5) of Van de Geer (2000)).

To see that $G_{c,n}$ is an envelope for $\mathcal{G}_c(n)$, observe that for any $G \in \mathcal{F}$,

$$G - (1-\gamma_n)F_b \quad = \quad G - F + \frac{M}{r_n}(F_s^{\alpha_0} - F_b) + \gamma_n F_s^{\alpha_0}.$$

Hence,

$$F_s^{\alpha_0} - \frac{M}{r_n\gamma_n}\|F_s^{\alpha_0} - F_b\| - \frac{\|G-F\|}{\gamma_n} \leq \frac{G - (1-\gamma_n)F_b}{\gamma_n} \leq F_s^{\alpha_0} + \frac{M}{r_n\gamma_n}\|F_s^{\alpha_0} - F_b\| + \frac{\|G-F\|}{\gamma_n}.$$

As the two bounds are monotone, from the properties of isotonic estimators (see e.g., Theorem 1.3.4 of Robertson et al. (1988)), we can always find a version of $\check{G}_s^{\gamma_n}$ such that

$$F_s^{\alpha_0} - \frac{M}{r_n\gamma_n}\|F_s^{\alpha_0} - F_b\| - \frac{\|G-F\|}{\gamma_n} \leq \check{G}_s^{\gamma_n} \leq F_s^{\alpha_0} + \frac{M}{r_n\gamma_n}\|F_s^{\alpha_0} - F_b\| + \frac{\|G-F\|}{\gamma_n}.$$

Therefore,

$$-2\frac{M}{r_n}\|F_s^{\alpha_0} - F_b\| - \|G-F\| \leq \gamma_n \check{G}_s^{\gamma_n} - \gamma_n F_s^{\alpha_0} - \frac{M}{r_n}(F_s^{\alpha_0} - F_b) \leq 2\frac{M}{r_n}\|F_s^{\alpha_0} - F_b\| + \|G-F\|. \tag{112}$$

Thus, for $\sqrt{n}(G-(1-\gamma_n)F_b-\gamma_n\check{G}_s^{\gamma_n})^2 \in \mathcal{G}_c(n)$,

$$\begin{aligned}
(G-(1-\gamma_n)F_b-\gamma_n\check{G}_s^{\gamma_n})^2 &= \left[(G-F)+\left(\gamma_n\check{G}_s^{\gamma_n}-\gamma_n F_s^{\alpha_0}-\frac{M}{r_n}(F_b-F_s^{\alpha_0})\right)\right]^2 \\
&\le 2(G-F)^2+2\left(\gamma_n\check{G}_s^{\gamma_n}-\gamma_n F_s^{\alpha_0}-\frac{M}{r_n}(F_b-F_s^{\alpha_0})\right)^2 \\
&\le 2\|G-F\|^2+2\left(2\frac{M}{r_n}\|F_s^{\alpha_0}-F_b\|+\|G-F\|\right)^2 \\
&\le 6\|G-F\|^2+16\frac{M^2}{r_n^2}\|F_s^{\alpha_0}-F_b\|^2 \\
&\le 6c^2+16\frac{M^2}{r_n^2}\|F_s^{\alpha_0}-F_b\|^2=\frac{G_{c,n}}{\sqrt{n}},
\end{aligned}$$

where the second inequality follows from (112). From the definition of g_n and D_n^2, we have $|g_n(t)| \le \frac{6}{\sqrt{n}}D_n^2+16\sqrt{n}\frac{M^2}{r_n^2}\|F_s^{\alpha_0}-F_b\|^2$, for all $t\in\mathbb{R}$. As $D_n=O_P(1)$, for any given $\epsilon>0$, there exists $c>0$ (depending on ϵ) such that

$$P(g_n\notin\mathcal{G}_c(n)) = P\left(\|\mathbb{F}_n-F\|\ge\frac{c}{\sqrt{n}}\right)=P(D_n\ge c)\le\epsilon, \tag{113}$$

for all sufficiently large n.

Therefore, for any given $\delta>0$ and $\epsilon>0$, we can make both $J\{6\frac{c^2}{\sqrt{n}}+16\sqrt{n}\frac{M^2}{r_n^2}\|F_s^{\alpha_0}-F_b\|^2\}^2$ and $P(g_n\notin\mathcal{G}_c(n))$ less than ϵ for large enough n and $c(>0)$, using the fact that $\sqrt{n}/r_n^2\to 0$ and (113). Thus, $P(|\nu_n(g_n)|>\delta)\le 2\epsilon$ by (111).

A similar analysis can be done for the second term of (110). The result now follows.

15.5.3. Proof of Lemma 15.3

Note that

$$\frac{\sqrt{n}\gamma_n}{c_n}(\hat{F}_{s,n}^{\gamma_n}-\bar{F}_{s,n}^{\gamma_n})=\frac{\sqrt{n}\gamma_n}{c_n}(\hat{F}_{s,n}^{\gamma_n}-F_s^{\alpha_0})-\frac{\sqrt{n}\gamma_n}{c_n}(\bar{F}_{s,n}^{\gamma_n}-F_s^{\alpha_0}).$$

However, a simplification yields

$$\frac{\sqrt{n}\gamma_n}{c_n}(\hat{F}_{s,n}^{\gamma_n}-F_s^{\alpha_0})=\frac{1}{c_n}\sqrt{n}(\mathbb{F}_n-F)+\frac{\sqrt{n}M}{c_n r_n\alpha_0}(F-F_b).$$

Since $\sqrt{n}(\mathbb{F}_n-F)/c_n$ is $o_P(1)$, $\sqrt{n}=c_n r_n$, and $F-F_b=\alpha_0(F_s^{\alpha_0}-F_b)$, we have

$$\frac{\sqrt{n}\gamma_n}{c_nM}(\hat{F}_{s,n}^{\gamma_n}-F_s^{\alpha_0})\xrightarrow{P}F_s^{\alpha_0}-F_b \quad \text{in } \mathbb{H}. \tag{114}$$

By applying the functional delta method (see Theorem 20.8 of Van der Vaart (1998)) for the projection operator (see Theorem 1 of Fils-Villetard et al. (2008)) to (114), we have

$$\frac{\sqrt{n}\gamma_n}{c_nM}(\bar{F}_{s,n}^{\gamma_n}-F_s^{\alpha_0})\xrightarrow{P}\Pi\left(F_s^{\alpha_0}-F_b\,|\,T_{\mathcal{F}}(F_s^{\alpha_0})\right) \quad \text{in } \mathbb{H}. \tag{115}$$

By combining (114) and (115), we have

$$\frac{\sqrt{n}\gamma_n}{c_n M}(\hat{F}_{s,n}^{\gamma_n} - \bar{F}_{s,n}^{\gamma_n}) \xrightarrow{P} (F_s^{\alpha_0} - F_b) - \Pi\left(F_s^{\alpha_0} - F_b \middle| T_{\mathcal{F}}(F_s^{\alpha_0})\right) \quad \text{in } \mathbb{H}. \tag{116}$$

The result now follows by applying the continuous mapping theorem to (116). We prove $V \neq 0$ by contradiction. Suppose that $V = 0$, i.e., $(F_s^{\alpha_0} - F_b) \in T_{\mathcal{F}}(F_s^{\alpha_0})$. Therefore, for some distribution function G and $\eta > 0$, we have $V = (\eta+1)F_s^{\alpha_0} - F_b - \eta G$, by the definition of $T_{\mathcal{F}}(F_s^{\alpha_0})$. By the discussion leading to (5), it can be easily seen that ηG is a sub-CDF, while $(\eta+1)F_s^{\alpha_0} - F_b$ is not (as that would contradict (5)). Therefore, $V \neq 0$ and thus $\int V^2 dF > 0$.

15.6. Proof of Theorem 3.8

The constant c defined in the statement of the theorem can be explicitly expressed as

$$c = -\left\{\int V^2 dF\right\}^{-\frac{1}{2}},$$

where

$$V = (F_s - F_b) - \Pi(F_s - F_b | T_{\mathcal{F}}(F_s)),$$

and Π and $T_{\mathcal{F}}(\cdot)$ are defined in (103) and (104), respectively.

Let $x > 0$. Obviously,

$$P(r_n(\hat{\alpha}_0^{c_n} - \alpha_0) \leq x) = 1 - P(r_n(\hat{\alpha}_0^{c_n} - \alpha_0) > x).$$

By Lemma 15.1, we have that $P(r_n(\hat{\alpha}_0^{c_n} - \alpha_0) > x) \to 0$ if $c_n \to \infty$. Now let $x \leq 0$. In this case the left hand side of the above display equals $P(\sqrt{n}\gamma_n d_n(\hat{F}_{s,n}^{\gamma_n}, \check{F}_{s,n}^{\gamma_n}) \leq c_n)$, where $\gamma_n = \alpha_0 + x/r_n$. A simplification yields

$$\frac{\sqrt{n}}{c_n}\gamma_n(\hat{F}_{s,n}^{\gamma_n} - F_s^{\alpha_0}) \xrightarrow{P} -x(F_s^{\alpha_0} - F_b), \text{ in } \mathbb{H}, \tag{117}$$

since $\sqrt{n}(\mathbb{F}_n - F)/c_n$ is $o_P(1)$; see the proof of Lemma 15.3 (Section 15.5.3) for the details. By applying the functional delta method (cf. Theorem 20.8 of Van der Vaart (1998)) for the projection operator (see Theorem 1 of Fils-Villetard et al. (2008)) to (117), we have

$$\frac{\sqrt{n}}{c_n}\gamma_n(\bar{F}_{s,n}^{\gamma_n} - F_s^{\alpha_0}) \xrightarrow{d} \Pi\left(-x(F_s^{\alpha_0} - F_b) \middle| T_{\mathcal{F}}(F_s^{\alpha_0})\right) \quad \text{in } \mathbb{H}. \tag{118}$$

Adding (117) and (118), we get

$$\frac{\sqrt{n}}{c_n}\gamma_n(\hat{F}_{s,n}^{\gamma_n} - \bar{F}_{s,n}^{\gamma_n}) \to -x(F_s^{\alpha_0} - F_b) - \Pi\left(-x(F_s^{\alpha_0} - F_b) \middle| T_{\mathcal{F}}(F_s^{\alpha_0})\right) \quad \text{in } \mathbb{H}.$$

By the continuous mapping theorem, we get $\sqrt{n}/c_n\gamma_n d(\hat{F}_{s,n}^{\gamma_n}, \bar{F}_{s,n}^{\gamma_n}) \xrightarrow{P} |x| \left\{\int V^2 dF\right\}^{1/2}$. Hence, by Lemma 15.2,

$$P(r_n(\hat{\alpha}_0^{c_n} - \alpha_0) \leq x) \to \begin{cases} 1, & \text{if } x > 0, \\ 1, & \text{if } x \leq 0 \text{ and } |x| \leq \left\{\int V^2 dF\right\}^{-1/2}, \\ 0, & \text{otherwise.} \end{cases}$$

15.7. Proof of Theorem 4.2

It is enough to show that $\sup_x |H_n(x) - G(x)| \to 0$, where G is the limiting distribution of the Cramér-von Mises statistic, a continuous distribution. As $\sup_x |G_n(x) - G(x)| \to 0$, it is enough to show that

$$\sqrt{n} d_n(\mathbb{F}_n, F) - \sqrt{n} d(\mathbb{F}_n, F) \xrightarrow{P} 0. \tag{119}$$

We now prove (119). Observe that

$$n(d_n^2 - d^2)(\mathbb{F}_n, F) = \sqrt{n}(\mathbb{P}_n - P)[\hat{g}_n] = \nu_n(\hat{g}_n), \tag{120}$$

where $\hat{g}_n = \sqrt{n}(\mathbb{F}_n - F)^2$, $\mathbb{P}_n$ denotes the empirical measure of the data, and $\nu_n := \sqrt{n}(\mathbb{P}_n - P)$ denotes the usual empirical process. We will show that $\nu_n(\hat{g}_n) \xrightarrow{P} 0$, which will prove (120).

For each positive integer n, we introduce the following class of functions

$$\mathcal{G}_c(n) = \left\{ \sqrt{n}(H - F)^2 : H \in \mathcal{F} \text{ and } \sup_{t \in \mathbb{R}} |H(t) - F(t)| < \frac{c}{\sqrt{n}} \right\}.$$

Let us also define

$$D_n := \sup_{t \in \mathbb{R}} \sqrt{n} |\mathbb{F}_n(t) - F(t)|.$$

From the definition of $\hat{g}_n$ and D_n^2, we have $\hat{g}_n(t) \le \frac{1}{\sqrt{n}} D_n^2$, for all $t \in \mathbb{R}$. As $D_n = O_P(1)$, for any given $\epsilon > 0$, there exists $c > 0$ (depending on ϵ) such that

$$P(\hat{g}_n \notin \mathcal{G}_c(n)) = P(\sqrt{n} \sup_t |\hat{g}_n(t)| \ge c^2) = P(D_n^2 \ge c^2) \le \epsilon, \tag{121}$$

for all sufficiently large n. Therefore, for any $\delta > 0$, using the same sequence of steps as in (111),

$$P(|\nu_n(\hat{g}_n)| > \delta) \quad \le \quad J_{[\,]} \frac{E[G_c^2(n)]}{\delta} + P(\hat{g}_n \notin \mathcal{G}_c(n)), \tag{122}$$

where $G_c(n) := \frac{c^2}{\sqrt{n}}$ is an envelope for $\mathcal{G}_c(n)$ and $J_{[\,]}$ is a constant. Note that to derive the last inequality we have used the maximal inequality in Corollary (4.3) of Pollard (1989); the class $\mathcal{G}_c(n)$ is "manageable" in the sense of Pollard (1989) (as a consequence of equation (2.5) of Van de Geer (2000)).

Therefore, for any given $\delta > 0$ and $\epsilon > 0$, for large enough n and $c > 0$ we can make both $J_{[\,]} c^4/(\delta n)$ and $P(\hat{g}_n \notin \mathcal{G}_c(n))$ less than ϵ, using (121) and (122), and thus, $P(|\nu_n(\hat{g}_n)| > \delta) \le 2\epsilon$. The result now follows.

15.8. Proof of Theorem 4.3

The random variable U defined in the statement of the theorem can be explicitly expressed as

$$U := \left[\int \{\mathbb{G}_F - \Pi(\mathbb{G}_F | T_{\mathcal{F}}(F_s^{\alpha_0})\}^2 \, dF \right]^{1/2},$$

where $\mathbb{G}_F$ is the F-Brownian bridge.

By the same line of arguments as in the proof of Lemma 15.2 (see Section 15.5.2), it can be easily seen that $\sqrt{n}\alpha_0 \; d_n(\hat{F}^{\alpha_0}_{s,n}, \check{F}^{\alpha_0}_{s,n}) - \sqrt{n}\alpha_0 \; d(\hat{F}^{\alpha_0}_{s,n}, \bar{F}^{\alpha_0}_{s,n}) \xrightarrow{P} 0$. Moreover, by Donsker's theorem,

$$\sqrt{n}\alpha_0(\hat{F}^{\alpha_0}_{s,n} - F^{\alpha_0}_s) \xrightarrow{d} \mathbb{G}_F.$$

By applying the functional delta method for the projection operator, in conjunction with the continuous mapping theorem to the previous display, we have

$$\sqrt{n}\alpha_0(\bar{F}^{\alpha_0}_{s,n} - F^{\alpha_0}_s) \xrightarrow{d} \Pi(\mathbb{G}_F | T_{\mathcal{F}}(F^{\alpha_0}_s)) \quad \text{in} \quad \mathbb{H},$$

where Π, $T_{\mathcal{F}}(\cdot)$, and $F^{\alpha_0}_s$ are defined in (103), (104), and (106), respectively. Hence, by an application of the continuous mapping theorem, we have $\sqrt{n}\alpha_0 d(\hat{F}^{\alpha_0}_{s,n}, \bar{F}^{\alpha_0}_{s,n}) \xrightarrow{d} U$. The result now follows.

15.9. Proof of Theorem 6.1

The constant c and the function Q defined in the statement of the theorem can be explicitly expressed as

$$c = d(Q, \Pi\,(Q | T_{\mathcal{F}}(F_s))),$$

and

$$Q := (F_s - F_b)\left\{\alpha_0^2 \int V^2 dF\right\}^{-1/2},$$

where

$$r_n = \sqrt{n}/c_n, \quad V = (F_s - F_b) - \Pi(F_s - F_b | T_{\mathcal{F}}(F_s)),$$

and Π and $T_{\mathcal{F}}(\cdot)$ are defined in (103) and (104), respectively.

Recall the notation of Section 15.5. Note that from (2),

$$\hat{F}^{\check{\alpha}_n}_{s,n}(x) = \frac{\alpha_0}{\check{\alpha}_n} F_s(x) + \frac{\check{\alpha}_n - \alpha_0}{\check{\alpha}_n} F_b(x) + \frac{(\mathbb{F}_n - F)(x)}{\check{\alpha}_n},$$

for all $x \in \mathbb{R}$. Thus we can bound $\hat{F}^{\check{\alpha}_n}_{s,n}(x)$ as follows:

$$\frac{\alpha_0}{\check{\alpha}_n} F_s(x) - \frac{|\check{\alpha}_n - \alpha_0|}{\check{\alpha}_n} - \frac{D'_n}{\check{\alpha}_n} \le \hat{F}^{\check{\alpha}_n}_{s,n}(x) \le \frac{\alpha_0}{\check{\alpha}_n} F_s(x) + \frac{|\check{\alpha}_n - \alpha_0|}{\check{\alpha}_n} + \frac{D'_n}{\check{\alpha}_n},$$

where $D'_n = \sup_{x\in\mathbb{R}} |\mathbb{F}_n(x) - F(x)|$. As both the upper and lower bounds are monotone, we can always find a version of $\check{F}^{\check{\alpha}_n}_{s,n}$ such that

$$\frac{\alpha_0}{\check{\alpha}_n} F_s - \frac{|\check{\alpha}_n - \alpha_0|}{\check{\alpha}_n} - \frac{D'_n}{\check{\alpha}_n} \le \check{F}^{\check{\alpha}_n}_{s,n} \le \frac{\alpha_0}{\check{\alpha}_n} F_s + \frac{|\check{\alpha}_n - \alpha_0|}{\check{\alpha}_n} + \frac{D'_n}{\check{\alpha}_n}.$$

Therefore,

$$\begin{aligned} |\check{F}^{\check{\alpha}_n}_{s,n} - F_s| &\le \frac{|\alpha_0 - \check{\alpha}_n|}{\check{\alpha}_n} F_s + \frac{|\check{\alpha}_n - \alpha_0|}{\check{\alpha}_n} + \frac{D'_n}{\check{\alpha}_n} \\ &\le 2\frac{|\alpha_0 - \check{\alpha}_n|}{\check{\alpha}_n} + \frac{D'_n}{\check{\alpha}_n} \xrightarrow{P} 0, \end{aligned}$$

as $n \to \infty$, using the fact $\check{\alpha}_n \xrightarrow{P} \alpha_0 \in (0,1)$. Furthermore, if $q_n(\check{\alpha}_n - \alpha_0) = O_P(1)$, where $q_n/\sqrt{n} \to 0$, it is easy to see that $q_n|\check{F}_{s,n}^{\check{\alpha}_n} - F_s| = O_P(1)$, as $q_n D'_n = o_P(1)$. Note that

$$r_n \hat{\alpha}_0^{c_n}(\hat{F}_{s,n}^{\hat{\alpha}_0^{c_n}} - F_s) = r_n(\mathbb{F}_n - F) + r_n(\alpha_0 - \hat{\alpha}_0^{c_n})\ (F_s - F_b)$$

Thus

$$\sup_{x\in\mathbb{R}} |r_n(\hat{F}_{s,n}^{\hat{\alpha}_0^{c_n}} - F_s)(x) - Q(x)| \xrightarrow{P} 0.$$

Hence by an application of functional delta method for the projection operator, in conjunction with the continuous mapping theorem, we have

$$r_n d(\check{F}_{s,n}^{\hat{\alpha}_0^{c_n}}, F_s) \xrightarrow{P} d(Q, \Pi(Q|T_{\mathcal{F}}(F_s))).$$

References

Cai, T., J. Jin, and M. G. Low (2007). Estimation and confidence sets for sparse normal mixtures. *Ann. Statist. 35*(6), 2421–2449.

Donoho, D. and J. Jin (2004). Higher criticism for detecting sparse heterogeneous mixtures. *Ann. Statist. 32*(3), 962–994.

Fils-Villetard, A., A. Guillou, and J. Segers (2008). Projection estimators of Pickands dependence functions. *Canad. J. Statist. 36*(3), 369–382.

Kulldorff, M., J. Heffernan, R. Hartman, R. Assuncao, and F. Mostashari (2005). A space-time permutation scan statistic for disease outbreak detection. *PLoS Med. 2*(3), e59.

Pollard, D. (1989). Asymptotics via empirical processes. *Statist. Sci. 4*(4), 341–366. With comments and a rejoinder by the author.

Robertson, T., F. T. Wright, and R. L. Dykstra (1988). *Order restricted statistical inference.* Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. Chichester: John Wiley & Sons Ltd.

Van de Geer, S. A. (2000). *Applications of empirical process theory*, Volume 6 of *Cambridge Series in Statistical and Probabilistic Mathematics.* Cambridge: Cambridge University Press.

Van der Vaart, A. W. (1998). *Asymptotic statistics*, Volume 3 of *Cambridge Series in Statistical and Probabilistic Mathematics.* Cambridge: Cambridge University Press.