CHAPTER 6

GOODNESS–OF–FIT TESTS FOR THE ERRORS

6.1. INTRODUCTION

Consider the model (1.1.1) and the goodness–of–fit hypothesis

$$(1) \qquad H_0\colon \mathrm{F}_{ni} \equiv \mathrm{F}_0, \quad \mathrm{F}_0 \text{ a known continuous d.f..}$$

This is a classical problem yet not much is readily available in literature. Observe that even if F_0 is known, having an unknown $\boldsymbol{\beta}$ in the model poses a problem in constructing tests of H_0 that would be implementable, at least asymptotically.

One test of H_0 could be based on $\hat{\mathrm{D}}_1$ of (1.3.3). This test statistic is suggested by looking at the estimated residuals and mimicking the one sample location model technique. In general, its large sample distribution depends on the design matrix. In addition, it does not reduce to the Kiefer (1959) tests of goodness–of–fit in the k–sample location problem when (1.1.1) is reduced to this model. The test statistics that overcome these deficiencies are those that are based on the w.e.p.'s $\mathbf{V}$ of (1.1.2). For example, the two candidates that will be considered in this chapter are

$$(2) \qquad \hat{\mathrm{D}}_2 := \sup_y |\mathbf{W}^0(y, \hat{\boldsymbol{\beta}})|, \qquad \hat{\mathrm{D}}_3 := \sup_y \|\mathbf{W}^0(y, \hat{\boldsymbol{\beta}})\|,$$

where $\hat{\boldsymbol{\beta}}$ is an estimator of $\boldsymbol{\beta}$ and,

$$(3) \qquad \mathbf{W}^0(y, \mathbf{t}) := (\mathbf{X}'\mathbf{X})^{-1/2}\{\mathbf{V}(y, \mathbf{t}) - \mathbf{X}'\mathbf{1}\,\mathrm{F}_0(y)\}, \qquad y \in \mathbb{R},\ \mathbf{t} \in \mathbb{R}^p,$$

$$\mathbf{1}' := (1, \ldots, 1)_{1\times n}.$$

Other classes of tests are based on $\mathrm{K}^0_{\mathbf{X}}(\hat{\boldsymbol{\beta}}_{\mathbf{X}})$ and $\inf\{\mathrm{K}^0_{\mathbf{X}}(\mathbf{t}), \mathbf{t} \in \mathbb{R}^p\}$, where $\mathrm{K}^0_{\mathbf{X}}$ is equals to the $\mathrm{K}_{\mathbf{X}}$ of (1.3.2) with $\mathbf{W}$ replaced by $\mathbf{W}^0$ in there.

Section 6.2a discusses the asymptotic null distributions (a.n.d.'s) of the supremum distance test statistics for H_0 when $\boldsymbol{\beta}$ is estimated arbitrarily and asymptotically efficiently. Also discussed in this section are some asymptotically distribution free (a.d.f.) tests for H_0. Some comments about the asymptotic power of these tests appear at the end of this section. Section 6.2b discusses a smooth bootstrap distribution of $\hat{\mathrm{D}}_3$.

Analogous results for tests of H_0 based on L_2-distances involving the ordinary and weighted empirical processes appear in Section 6.3.

A closely related problem to H_0 is that of testing the composite hypothesis

(4) $$H_1: \ F_{ni}(\cdot) = F_0(\cdot/\sigma), \ \sigma > 0, \ F_0 \text{ a known d.f..}$$

Modifications of various tests of H_0 and their asymptotic null distributions are discussed in Section 6.4.

Another problem of interest is to test the composite hypothesis of symmetry of the errors:

(5) $$H_s\text{: } F_{ni} = F, \ 1 \le i \le n, \ n \ge 1; \ F \text{ a d.f. symmetric around } 0.$$

This is a more general hypothesis than H_0. In some situations it may be of interest to test H_s before testing, say, that the errors are normally distributed. Rejection of H_s would *a priori* exclude any possibility of normality of the errors. A test of H_s could be based on

(6) $$\hat{D}_{1s} := \sup_y |W_1^+(y, \hat{\beta})|,$$

where

(7) $$W_1^+(y, \mathbf{t}) := n^{-1/2} \sum_{i=1}^{n} [I(Y_{ni} \le y + \mathbf{x}_{ni}'\mathbf{t}) - I(-Y_{ni} < y - \mathbf{x}_{ni}'\mathbf{t})]$$
$$:= H_n(y, \mathbf{t}) - 1 + H_n(-y, \mathbf{t}), \qquad y \in \mathbb{R}, y \in \mathbb{R}^p,$$

with H_n as in (1.2.1). Other candidates are

(8) $$\hat{D}_{2s} := \sup_y |\mathbf{W}^+(y, \hat{\beta})|,$$
$$\hat{D}_{3s} := \sup_y \|\mathbf{W}^+(y, \hat{\beta})\| = \sup_y[\mathbf{V}^{+'}(y, \hat{\beta})(\mathbf{X}'\mathbf{X})^{-1}\mathbf{V}^+(y, \hat{\beta})]^{1/2},$$

where

(9) $$\mathbf{W}^+ := \mathbf{A}\mathbf{V}^+, \qquad \mathbf{V}^{+'} := (V_1^+, \ldots, V_p^+), \text{ with}$$
$$V_j^+(y, \mathbf{t}) := V_j(y, \mathbf{t}) - \sum_{i=1}^{n} x_{nij} + V_j(-y, \mathbf{t}), \qquad 1 \le j \le p, \quad y \in \mathbb{R}, \quad \mathbf{t} \in \mathbb{R}^p.$$

Yet other tests can be obtained by considering various L_2-norms involving W_1^+ and $\mathbf{W}^+$. The asymptotic null distribution of all of these test statistics is given in Section 6.5.

It will be observed that the tests based on the vectors $\mathbf{W}^0$ and $\mathbf{W}^+$ of w.e.p.'s will have asymptotic distributions similar to their counterparts in the k–sample location models. Consequently these tests can use, at least for the large samples, the null distribution tables that are available for such problems. For the sake of the completeness some of these table are reproduced in the following sections.

6.2. THE SUPREMUM DISTANCE TESTS

6.2a. Asymptotic Null Distributions.

To begin with, define, for $0 \le t \le 1$, $s \in \mathbb{R}^p$,

$$W_1(t, s) := n^{1/2}\{H_n(F_0^{-1}(t), s) - t\}, \qquad \mathbf{W}(t, s) := \mathbf{W}^0(F_0^{-1}(t), s). \tag{1}$$

Let

$$\hat{W}_1(t) := W_1(t, \hat{\beta}), \qquad \hat{\mathbf{W}}(t) := \mathbf{W}(t, \hat{\beta}), \qquad 0 \le t \le 1. \tag{2}$$

Clearly, if F_0 is continuous then the distribution of $\hat{D}_j$, $j = 1, 2, 3$, is the same as that of $\|\hat{W}_1\|_\infty$, $\sup\{|\hat{\mathbf{W}}(t)|;\ 0 \le t \le 1\}$, $\sup\{\|\hat{\mathbf{W}}(t)\|;\ 0 \le t \le 1\}$, respectively. Consequently, from Corollaries 2.3.3 and 2.3.5 one readily obtains the following Theorem 6.2a.1. Recall the conditions (**F**$_0$**1**) and (**NX**) from Corollary 2.3.1 and just after Corollary 2.3.2.

Theorem 6.2a.1. *Suppose that the model* (1.1.1) *and* H_0 *hold. In addition, assume that* X *and* F_0 *satisfy* (**NX**) *and* (**F**$_0$**1**), *and that* $\hat{\beta}$ *satisfies*

$$\|A^{-1}(\hat{\beta} - \beta)\| = O_p(1). \tag{3}$$

Then

$$\sup|W_1(t, \hat{\beta}) - \{W_1(t, \beta) + q_0(t) \cdot n^{1/2}\, \bar{x}_n' \mathbf{A} \cdot A^{-1}(\hat{\beta} - \beta)\}| = o_p(1), \tag{4}$$

$$\sup\|\mathbf{W}(t, \hat{\beta}) - \{\mathbf{W}(t, \beta) + q_0(t) \cdot A^{-1}(\hat{\beta} - \beta)\}\| = o_p(1), \tag{5}$$

where $q_0 := f_0(F_0^{-1})$ *and the supremum is over* $0 \le t \le 1$. □

Write $W_1(t)$, $\mathbf{W}(t)$ for $W_1(t, \beta)$, $\mathbf{W}(t, \beta)$, respectively. The following corollary gives the weak limits of W_1 and $\mathbf{W}$ under H_0.

Lemma 6.2a.2. *Suppose that the model* (1.1.1) *and* H_0 *hold. Then*

$$W_1 \Rightarrow B, \quad B \text{ a Brownian bridge in } \mathbb{C}[0, 1]. \tag{7}$$

In addition, if X *satisfies* (**NX**), *then,*

$$\mathbf{W} \Rightarrow B' := (B_1, \dots, B_p) \tag{8}$$

where $B_1, \dots, B_p$ *are independent Brownian bridges in* $\mathbb{C}[0, 1]$.

Proof. The result (7) is well known or may be deduced from Corollary 2.2a.2. The same corollary implies (8). To see this, rewrite

$$\text{(9)} \qquad \mathbf{W}(t) = \mathbf{A}\sum_i \mathbf{x}_{ni}\{I(e_{ni} \le F_0^{-1}(t)) - t\} = \mathbf{A}\mathbf{X}'\alpha_n(t),$$

where $\boldsymbol{\alpha}_n(t) := (\alpha_{n1}(t), ..., \alpha_{nn}(t))'$, with

$$\alpha_{ni}(t) := \{I(e_{ni} \le F_0^{-1}(t)) - t\}, \qquad 1 \le i \le n, \ 0 \le t \le 1.$$

Clearly, under H_0,

$$\text{(10)} \qquad E\mathbf{W} \equiv 0, \qquad \mathrm{Cov}(\mathbf{W}(s), \mathbf{W}(t)) = (s \wedge t - st)\mathbf{I}_{p\times p}, \qquad 0 \le s, t \le 1.$$

Now apply Corollary 2.2a.2 p times, jth time to the w.e.p. with the weights and r.v.'s given as in (11) below, $1 \le j \le p$, to conclude (8).

(11) weights $\mathbf{d}_{(j)} \equiv$ the j^{th} column of $\mathbf{XA}$, the r.v.'s $X_{ni} \equiv e_{ni}$, and $F \equiv F_0$, $1 \le j \le p$,

See (2.3.33) and (2.3.34) for ensuring the applicability of Corollary 2.2a.2 to this case. □

Remark 6.2a.1. From (5) it follows that if $\hat{\beta}$ is chosen so that the finite dimensional asymptotic distributions of $\{\mathbf{W}(t) + q_0(t)\ \mathbf{A}^{-1}(\hat{\beta} - \beta);\ 0 \le t \le 1\}$ do not depend on the design matrix then the a.n.d.'s of $\hat{D}_j$, $j = 2, 3$, will also not depend on the design matrix. The classes of estimators that satisfy this requirment include M-, R- and m.d. estimators. Consequently, in these cases, the a.n.d.'s of $\hat{D}_j$, $j = 2, 3$, are design free.

On the other hand, from (4), the a.n.d. of $\hat{D}_1$ depends on the design matrix through $n^{1/2}\bar{\mathbf{x}}_n'\mathbf{A}$. Of course, if $\bar{\mathbf{x}}_n$ equals to zero, then this distribution is free from F_0 and the design matrix. □

Remark 6.2a.2. *The effect of estimating the parameter* β *efficiently.* To describe this, assume that

(12) F_0 has an a.c. density f_0 with a.e. derivative $\dot{f}_0$ satisfying

$$0 < I_0 := \int (\dot{f}_0/f_0)^2\, dF_0 < \infty.$$

Define

$$\text{(13)} \qquad s_{ni} := -\dot{f}_0(e_{ni})/f_0(e_{ni}), \quad 1 \le i \le n; \qquad \mathbf{s}_n := (s_{n1}, ..., s_{nn})',$$

and assume that the estimator $\hat{\beta}$ satisfies

$$\text{(14)} \qquad \mathbf{A}^{-1}(\hat{\beta} - \beta) = I_0^{-1}\,\mathbf{A}\mathbf{X}'\mathbf{s}_n + o_p(1).$$

Then, the approximating processes in (4) and (5), respectively, become

$$(15)\qquad \mathcal{W}_1(t) := W_1(t) + q_0(t)\cdot n^{1/2}\,\bar{x}_n' A\cdot \mathcal{I}_0^{-1} AX' s_n,$$

$$\mathcal{W}(t) := \mathbf{W}(t) + q_0(t)\cdot \mathcal{I}_0^{-1} AX' s_n, \qquad 0 \le t \le 1.$$

Using the independence of the errors, one directly obtains

$$(16)\qquad E\,\mathcal{W}_1(s)\,\mathcal{W}_1(t) = \{s(1-t) - n\bar{x}_n'(X'X)^{-1}\bar{x}_n\, q_0(s)q_0(t)\mathcal{I}_0^{-1}\},$$

$$E\,\mathcal{W}(s)\,\mathcal{W}'(t) = \{s(1-t) - q_0(s)q_0(t)\mathcal{I}_0^{-1}\}\, I_{p\times p}, \qquad 0 \le s \le t \le 1.$$

The calculations in (16) use the facts that $Es_n \equiv \mathbf{0}$, $E\alpha_n(t)s_n' \equiv q_0(t)I_{n\times n}$.

From (16), Theorem 2.2a.1(i) applied to the quantities given in (11), and the uniform continuity of q_0, which is implied by (12), it readily follows that $\mathcal{W} \Rightarrow \mathbf{Z} := (Z_1, \ldots, Z_p)'$, where $Z_1, \ldots, Z_p$ are continuous independent Gaussian processes, each having the covariance function

$$(17)\qquad \rho(s, t) := s(1-t) - q_0(s)q_0(t)\mathcal{I}_0^{-1}, \qquad 0 \le s \le t \le 1.$$

Consequently,

$$(18)\qquad \hat{D}_2 \Rightarrow \sup\{|\mathbf{Z}(t)|;\ 0\le t\le 1\}, \quad \hat{D}_3 \Rightarrow \sup\{\|\mathbf{Z}(t)\|;\ 0\le t\le 1\}.$$

This shows that the a.n.d.'s of $\hat{D}_j$, $j = 2, 3$, are design free when an asymptotically efficient estimator of β is used in constructing the residuals while the same can not be said about $\hat{D}_1$.

Moreover, recall, say from Durbin (1975), that when testing for H_0 in the one sample location model, the Gaussian process Z_1 with the covariance function ρ appears as the limiting process for the analogue of $\hat{D}_1$. Note also that in this case, $\hat{D}_1 = \hat{D}_2 = \hat{D}_3$. However, it is the test based on $\hat{D}_3$ that provides the right extension of the one sample Kolmogorov goodness–of–fit test to the linear regression model (1.1.1) for testing H_0 in the sense that it includes the k–sample goodness–of–fit Kolmogorov type test of Kiefer (1959). That is, if we specialize (1.1.1) to the k–sample location model, then $\hat{D}_3$ reduces to the T_N' of Section 2 of Kiefer modulo the fact that we have to estimate β.

The distribution of $\sup\{|Z_1(t)|;\ 0\le t\le 1\}$ has been studied by Durbin (1976) when F_0 equals N(0, 1) and some other distributions. Consequently, one can use these results together with the independence of $Z_1, \ldots, Z_p$ to implement the tests based on $\hat{D}_2$, $\hat{D}_3$ in a routine fashion. □

Remark 6.2a.3. *Asymptotically distribution free (a.d.f.) tests.* Here we shall construct estimators of β such that the above tests become a.d.f. for testing H_0. To that effect, write $\mathbf{X}_n$ and $\mathbf{A}_n$ for $\mathbf{X}$ and $\mathbf{A}$ to emphasize their dependence on n. Recall that n is the number of rows in $\mathbf{X}_n$. Let m $= m_n$ be a sequence of positive integers, $m_n \le n$. Let $\mathbf{X}_m$ be $m_n \times p$ matrix obtained from some m_n rows of $\mathbf{X}_n$. A way to choose m_n and these rows will be discussed later on. Relable the rows of $\mathbf{X}_n$ so that its first m_n rows are the rows of $\mathbf{X}_m$ and let $\{e_{ni}^*,\ 1 \le i \le m_n\}$, $\{Y_{ni}^*;\ 1 \le i \le m_n\}$ denote the corresponding errors and observations, respectively. Define

$$(19)\qquad s_{ni}^* := -\dot{f}_0(e_{ni}^*)/f_0(e_{ni}^*),\ 1 \le i \le m_n; \quad \mathbf{s}_m^* := (s_{ni}^*,\ 1 \le i \le m_n)',$$
$$\mathbf{T}_m := I_0^{-1}\,\mathbf{A}_m\mathbf{X}_m'\,\mathbf{s}_m^*, \qquad \mathbf{A}_m = (\mathbf{X}_m'\mathbf{X}_m)^{-1/2}.$$

Observe that under (12),

$$(20)\qquad E\mathbf{T}_m \equiv \mathbf{0}, \qquad E\mathbf{T}_m\mathbf{T}_m' \equiv I_0^{-1}\,\mathbf{I}_{p\times p}.$$

Consider the assumption

$$(21)\qquad m_n \le n,\ m_n \longrightarrow \infty \text{ such that}$$
$$(\mathbf{X}_n'\mathbf{X}_n)^{1/2}(\mathbf{X}_m'\mathbf{X}_m)^{-1}(\mathbf{X}_n'\mathbf{X}_n)^{1/2} \longrightarrow 2\mathbf{I}_{p\times p}.$$

The assumptions (21) and (NX) together imply

$$(22)\qquad \max_{1 \le i \le m} \mathbf{x}_{ni}'\mathbf{A}_m\mathbf{A}_m\,\mathbf{x}_{ni} = o(1).$$

Consequently one obtains, with the aid of the Cramer–Wold LF–CLT, that

$$(23)\qquad \mathbf{T}_m \underset{d}{\longrightarrow} N(\mathbf{0},\ I_0^{-1}\mathbf{I}_{p\times p}).$$

Now use $\{(\mathbf{x}_{ni}',\ Y_{ni}^*);\ 1 \le i \le m_n\}$ to construct an estimator $\hat{\beta}_m$ of β such that

$$(24)\qquad \mathbf{A}_m^{-1}(\hat{\beta}_m - \beta) = \mathbf{T}_m + o_p(1).$$

Note that, by (21) and (23), $\|\mathbf{A}_n^{-1}\mathbf{A}_m\|_\infty = O(1)$ and, hence

$$(25)\qquad \mathbf{A}_n^{-1}(\hat{\beta}_m - \beta) = \mathbf{A}_n^{-1}\mathbf{A}_m\mathbf{T}_m + o_p(1).$$

Therefore it follows that $\hat{\beta}_m$ satisfies (3). Define

$$\mathbf{K}^*(t) := \mathbf{W}(t) + \mathbf{A}_n^{-1}\mathbf{A}_m\mathbf{T}_m\,q_0(t), \qquad 0 \le t \le 1.$$

From (5) and (25) it now readily follows that

$$(26)\qquad \sup_{0\le t\le 1} \|\mathbf{W}(t,\hat{\beta}) - \mathbf{K}^*(t)\| = o_p(1).$$

We shall now show that

$$(27)\qquad \mathbf{K}^* \Rightarrow B, \text{ with } B \text{ as in (8).}$$

First, consider the covariance function of $\mathbf{K}^*$. By the independence of the errors and by (12) one obtains that

$$\begin{aligned} E\{I(e_{ni} \le F_0^{-1}(t)) - t\}\dot{f}_0(e_{nj}^*)/f_0(e_{nj}^*) &= 0, \qquad i\ne j,\ 1\le i\le n,\ 1\le j\le m_n,\\ &= q_0(t), \quad 1\le i=j\le m_n, \quad 0\le t\le 1. \end{aligned}$$

Use this and direct calculations to obtain that

$$\begin{aligned}(28)\quad E\mathbf{K}^*(s)\mathbf{K}^*(t) &= s(1-t)\mathbf{I}_{p\times p}\\ &\quad -I_0^{-1}q_0(s)q_0(t)[2\mathbf{I}_{p\times p}-(\mathbf{X}_n'\mathbf{X}_n)^{1/2}(\mathbf{X}_m'\mathbf{X}_m)^{-1}(\mathbf{X}_n'\mathbf{X}_n)^{1/2}],\\ &\qquad 0\le s\le t\le 1.\end{aligned}$$

Thus (21) implies that

$$(29)\qquad E\mathbf{K}^*(s)\mathbf{K}^*(t) \longrightarrow s(1-t)\mathbf{I}_{p\times p}, \qquad \forall\ 0\le s\le t\le 1.$$

Because of (8) and the uniform continuity of q_0, the relative compactness of the sequence $\{\mathbf{K}^*\}$ is *a priori* established, thereby completing the proof of (27). Consequently, we obtain the following

Corollary 6.2a.1. *Under* (1.1.1), H_0, (**NX**), (12), (21) *and* (24),

$$\hat{D}_{2m} \xrightarrow[d]{} \sup_{0\le t\le 1}\max_{1\le j\le p} |B_j(t)|, \qquad \hat{D}_{3m} \xrightarrow[d]{} \sup_{0\le t\le 1}\{\sum_{j=1}^{p} B_j^2(t)\}^{1/2},$$

where $\hat{D}_{jm}$ *stand for the* $\hat{D}_j$ *with* $\hat{\beta} = \hat{\beta}_m$, $j = 2, 3$. □

It thus follows, from the independence of the Brownian bridges $\{B_j,\ 1 \le j \le p\}$ and Theorem V.3.6.1 of Hajek and Sidak (1967), that the test that rejects H_0 when $\hat{D}_{2m} \ge d$ is of the asymptotic size α, provided d is determined from the relation

$$(30)\qquad 2\sum_{j=1}^{\infty}(-1)^{j+1}e^{-2j^2d^2} = 1-(1-\alpha)^{1/p}.$$

Let T_p stand for the the limiting r.v. of $\hat{D}_{3m}$. The distribution of T_p has been tabulated by Kiefer (1959) for $1 \le p \le 5$. Delong (1983) has also computed these tables for $1 \le p \le 7$. The following table is obtained from

Kiefer for $1 \le p \le 5$ and Delong for $p = 6, 7$, for the sake of completeness. The last place digit is rounded from their entries.

$\alpha \backslash p$	1	2	3	4	5	6	7
.001	1.9495	2.1516	2.3030	2.4301	2.5422	2.6437	2.7373
.005	1.7308	1.9417	2.0977	2.2280	2.3424	2.445	2.540
.01	1.6276	1.8427	2.0009	2.1326	2.2480	2.3525	2.4525
.02	1.5174	1.7370	1.8974	2.0305	2.1470	2.252	2.350
.025	1.480	1.702	1.8625	1.9961	2.116	2.217	2.315
.05	1.3581	1.5838	1.7473	1.8823	2.0001	2.1053	2.2031
.10	1.2239	1.4540	1.6196	1.7559	1.8746	1.981	2.0788
.15	1.1380	1.3703	1.5370	1.6740	1.7930	1.900	1.9977
.20	1.0728	1.3061	1.4734	1.6107	1.730	1.8352	1.9349
.25	1.0192	1.2530	1.4205	1.5579	1.6773	1.785	1.8825

Table 1: Values d *such that* $P(T_p \ge d) \simeq \alpha$ *for* $1 \le p \le 7$. *Obtained from Kiefer* (1959) *& Delong (personal communication).*

Note that for $p = 1$, $\hat{D}_{2m}$ and $\hat{D}_{3m}$ are the same tests and d of (30) is the same as the d of column 1 of Table 1 for various values of α.

The entries in Table 1 can be used to get the asymptotic critical level of $\hat{D}_{3m}$ for $1 \le p \le 7$. Thus for $p = 5$, $\alpha = .05$, the test that rejects H_0 when $\hat{D}_{3m} \ge 2.0001$ is of the asymptotic size .05, no matter what F_0 is within the class of d.f.'s satisfying (12).

Next, to make $\hat{D}_1$-test a.d.f., let $r = r_n$ be a sequence of positive integers, $r_n \le n$, $r_n \to \infty$. Let X_r denote the $r_n \times p$ matrix obtain from some r_n rows of X_n. Relable the rows of X_n so that the first r_n rows are in X_r and let Y_i^o, e_i^o denote the corresponding Y_i's and e_i's. Let $A_r = (X_r' X_r)^{-1/2}$. Assume that

(31) (i) $\|n^{1/2}\bar{\mathbf{x}}_n'\mathbf{A}_r\| = O(1)$, and

(ii) $|n\bar{\mathbf{x}}_n'(\mathbf{X}_r'\mathbf{X}_r)^{-1}\bar{\mathbf{x}}_n - 2r_n\bar{\mathbf{x}}_n'(\mathbf{X}_r'\mathbf{X}_r)^{-1}\bar{\mathbf{x}}_r| = o(1)$.

Let $\hat{\boldsymbol{\beta}}_r$ be an estimator of $\boldsymbol{\beta}$ based on $\{(\mathbf{x}_{ni}', Y_{ni}^o),\ 1 \le i \le r_n\}$ such that

(32) $$\mathbf{A}_r^{-1}(\hat{\boldsymbol{\beta}}_r - \boldsymbol{\beta}) = \mathbf{T}_r + o_p(1), \qquad \mathbf{T}_r := I_0^{-1}\mathbf{A}_r\mathbf{X}_r'\mathbf{s}_r^o$$

where $s_{ni}^o = -\dot{f}(e_{ni}^o)/f(e_{ni}^o)$, $1 \le i \le r_n$, and $\mathbf{s}_r^o = (s_{ni}^o,\ 1 \le i \le r_n)'$. Define

$$K_1^*(t) := W_1(t) + n^{1/2}\bar{\mathbf{x}}_n'\mathbf{A}_r \cdot \mathbf{T}_r q_0(t), \qquad 0 \le t \le 1.$$

Similar to (28), we obtain, for $s \le t$, that

$$EK_1^*(s)K_1^*(t) = s(1-t) - I_0^{-1}q_0(s)q_0(t)\{\bar{\mathbf{x}}_n'(\mathbf{X}_r'\mathbf{X}_r)^{-1}[n\bar{\mathbf{x}}_n - 2r_n\bar{\mathbf{x}}_r]\}$$

Argue as for Corollary 6.2a.1 to conclude

Corollary 6.2a.2. *Under* (1.1.1), H_0, (NX), (12), (31) *and* (32),

(33) $$\hat{D}_{1r} \xrightarrow[d]{} \sup_{0 \le t \le 1} |B(t)|,$$

where $\hat{D}_{1r}$ *is the* $\hat{D}_1$ *with* $\hat{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}_r$. □

Remark 6.2a.4. *Assumptions* (21) *and* (31). To begin with note that if

(34) $\lim_n n^{-1}(\mathbf{X}_n'\mathbf{X}_n)$ exists and is positive definite,

then (21) is equivalent to

(35) $nm_n^{-1} \longrightarrow 2$.

If, in addition to (34), one also assumes

(36) $\lim_n \bar{\mathbf{x}}_n$ exists and is finite,

then (31) is equivalent to

(37) $nr_n^{-1} \longrightarrow 2$.

There are many designs that satisfy (34) and (36). These include the one way classification, randomized block and the factorial designs, among others.

The choice of m_n and r_n rows is, of course, crucial, and obviously, depends on the design matrix. In the one way classification design with p treatments, n_j observations from the jth treatment, it is recommended to choose the first $m_{nj} = [n_j/2]$ observations from the jth treatment, $1 \le j \le p$, to estimate β. Here $m_n = m_{n1} + \dots + m_{np} = [n/2]$. One chooses $r_{nj} = m_{nj}$, $1 \le j \le p$, $r_n = \Sigma_j r_{nj} = [n/2]$. The choice of m_n and r_n is made similarly in the randomized block design and other similar designs. If one had several replications of a design, where the design matrix satisfies (34) and (36), then one could use the first half of the replications to estimate β and all replications to carry out the test.

Thus, in those cases where designs satisfy (34) and (36), the above construction of the a.d.f. tests is similar to the half sample technique in the one sample problem as found in Rao (1972) or Durbin (1976).

Of course there are designs of interest where (34) and (36) do not hold. An example is $p = 1$, $x_{ni} \equiv i$. Here, $X_n'X_n = O(n^3)$. If one decides to choose the first $m_n(r_n)$ x_i's, then (21) and (31) are equivalent to requiring $(m_n/n)^3 \longrightarrow 1/2$ and $(r_n/n)^2 \longrightarrow 1/2$. Thus, here $\hat{D}_{2m}$ or $\hat{D}_{3m}$ would use 79% of the observations to estimate β while $\hat{D}_{1r}$ would use 71%. On the other hand, if one decides to use the last $m_n(r_n)$ x_i's, then $\hat{D}_2$, $\hat{D}_3$ will use the last 21% observations while $\hat{D}_1$ will use the last 29% observations to estimate β. Of course all of these tests would be based on the entire sample.

In general, to avoid the above kind of problem, one may wish to use, from the practical point of view, some other characteristics of the design matrix in deciding which m_n, r_n rows to choose. One criterion to use may be to choose those $m_n(r_n)$ rows that will approximately maximize (m_n/n) $((r_n/n))$ subject to (21) ((31)). □

Remark 6.2a.5. *Construction of* $\hat{\beta}_m$ and $\hat{\beta}_r$. If F_0 is a d.f. for which the maximum likelihood estimator (m.l.e.) of β has a limiting distribution under (NX) and (12) then one should use this estimator based on r_n (m_n) observations $\{(x_i', Y_i)\}$ for $\hat{D}_1$ $(\hat{D}_2$ or $\hat{D}_3)$. For example, if F_0 is the N(0,1) d.f., then the obvious choice for $\hat{\beta}_r$ and $\hat{\beta}_m$ are the least squares estimators:

$$\hat{\beta}_r := (X_r'X_r)^{-1}X_r'Y_r^o; \quad \hat{\beta}_m := (X_m X_m)^{-1}X_m'Y_m^*$$

Of course there are many d.f.'s F_0 that satisfy the above conditions, but for which the computation of m.l.e. is not easy. One way to proceed in such cases is to use one step linear approximation. To make this precise, let $\bar{\beta}_m$ be an estimator of β based on $\{(x_{ni}', Y_{ni}),\ 1 \le i \le m_n\}$ such that

$$A_m^{-1}(\bar{\beta}_m - \beta) = O_p(1). \tag{38}$$

Define

$$(39)\qquad \psi_0(y) := -\dot{f}_0(y)/f_0(y), \qquad\qquad y \in \mathbb{R};$$

$$\bar{s}_{ni} := \psi_0(Y_{ni} - x_{ni}'\bar{\beta}_m),\ \ 1 \le i \le m_n; \qquad \bar{s}_m := (\bar{s}_{ni},\ 1 \le i \le m_n)';$$

$$\hat{\beta}_m := \bar{\beta}_m + I_0\, A_m A_m\, X_m' s_m;$$

$$V_m^*(y, t) = A_m \sum_{i=1}^{m_n} x_{ni}\, I(Y_{ni} \le y + x_{ni}'t), \qquad\qquad y \in \mathbb{R},\ t \in \mathbb{R}^p.$$

Then

$$A_m X_m' \bar{s}_m = \int \psi_0(y)\, V_m^*(dy, \bar{\beta}_m).$$

From this and (2.3.37), applied to $\{(x_{ni}', Y_{ni}),\ 1 \le i \le m_n\}$, one readily obtains

Corollary 6.2a.3. *Assume that* (1.1.1) *and* H_0 *hold. In addition, assume that* F_0 *is strictly increasing, satisfies* (12) *and is such that* ψ_0 *is a finite linear combination of nondecreasing bounded functions,* X *and* $\{\bar{\beta}_m\}$ *satisfy* (NX) *and* (38). *Then* $\{\hat{\beta}_m\}$ *of* (39) *satisfies* (24) *for any sequence* $m_n \longrightarrow \infty$, *as* $n \longrightarrow \infty$.

Proof. Clearly,

$$A_m^{-1}(\hat{\beta}_m - \beta) = A_m^{-1}(\bar{\beta}_m - \beta) + I_0^{-1}\, A_m X_m' \bar{s}_m.$$

But, integration by parts and (2.3.37) yield

$$\begin{aligned} A_m X_m'\{\bar{s}_m - s_m\} &= \int \psi_0(y)\{V_m^*(dy, \bar{\beta}) - V_m^*(dy, \beta)\} \\ &= -\int \{V_m^*(y, \bar{\beta}) - V_m^*(y, \beta)\}\, d\psi_0(y) \\ &= -A_m^{-1}(\bar{\beta}_m - \beta) \int f_0(y) d\psi_0(y) + o_p(1) \\ &= -A_m^{-1}(\bar{\beta}_m - \beta) I_0 + o_p(1). \end{aligned}$$

□

The above result is useful, e.g., when F_0 is logistic, Cauchy or double exponential. In the first case m.l.e. is not easy to compute but F_0 has finite second moment. So take $\bar{\beta}_m$ to be the l.s.e. and then use (39) to obtain the final estimator to be used for testing. In the case of Cauchy, $\bar{\beta}_m$ may be chosen to be an R-estimator.

Clearly, there is an analogue of the above corollary involving $\{\hat{\beta}_r\}$ that would satisfy (31). □

6.2b. Bootstrap Distributions

In this subsection we shall obtain a weak convergence result about a bootstrapped w.e.p.'s and then apply this to yield bootstrap distributions of some of the above tests.

Let (1.1.1) with $e_{ni} \equiv e_i$ and H_0 hold. Let E_0 and P_0 denote the expectation and probability, respectively, under these assumptions. In addition, throughout this section we shall *assume that* ($\mathbf{F_0 1}$), ($\mathbf{F_0 2}$) and (**NX**) hold.

Recall the definition of $\mathbf{W}$, $\hat{\mathbf{W}}$ from (6.2a.1), (6.2a.2). Let $\hat{\boldsymbol{\beta}}$ be an M-estimators of $\boldsymbol{\beta}$ corresponding to a bounded nondecreasing right continuous score function ψ such that

$$(1) \qquad \int \psi \, dF_0 = 0, \qquad \int f_0 \, d\psi > 0.$$

Upon specializing (4.2a.8) to the current setup one readily obtains

$$(2) \qquad \mathbf{A}^{-1}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) = -\kappa \, \Sigma_i \, \mathbf{A}\mathbf{x}_{ni} \, \psi(e_i) + o_p(1), \qquad (P_0).$$

where $\kappa := 1/\int f_0 \, d\psi$.

Let the approximating process obtained from (6.2a.5) and (2) be denoted by $\overline{\mathbf{W}}$, i.e.,

$$(3) \qquad \overline{\mathbf{W}}(t) := \Sigma_i \, \mathbf{A}\mathbf{x}_{ni}\{I(e_i \le F_0^{-1}(t)) - t - \kappa \, q_0(t) \, \psi(e_i)\}, \qquad 0 \le t \le 1.$$

Define

$$(4) \qquad \sigma^2 := E_0 \psi^2(e_1),$$

$$g_0(t) := E_0\{I(e_1 \le F_0^{-1}(t)) - t\} \, \psi(e_1)$$

$$= \int I(x \le F_0^{-1}(t)) \, \psi(x) \, dF_0(x), \qquad 0 \le t \le 1,$$

and, for $0 \le t \le u \le 1$,

$$(5) \qquad \rho_0(t, u) := t(1-u) - \kappa \, [q_0(t)g_0(u) + g_0(t)q_0(u)] + \kappa^2 q_0(t)q_0(u)\sigma^2.$$

Note that

$$(6) \qquad \mathcal{C}_0(t, u) := E_0\{\overline{\mathbf{W}}(t)\overline{\mathbf{W}}(u)'\} = \rho_0(t, u)\mathbf{I}_{p\times p}, \qquad 0 \le t \le u \le 1.$$

Let $\mathcal{G}_0 := (\mathcal{G}_{01}, \ldots, \mathcal{G}_{0p})'$ be a p-vector of independent Gaussian processes each having the covariance function ρ_0. Thus, $E\mathcal{G}_0(t)\mathcal{G}_0(u)' \equiv \mathcal{C}_0(t, u)$. Since ρ_0 is continuous, $\mathcal{G}_0 \in \{\mathbb{C}[0, 1]\}^p$. Moreover, from Corollary 2.2a.1 applied p time, j^{th} time to the entities $X_{ni} \equiv e_i$, $F_{ni} \equiv F_0$ and $d_{ni} \equiv (i,j)^{th}$

entry of $\mathbf{AX}$, $1 \le j \le p$, $1 \le i \le n$, and from the uniform continuity of q_0 it readily follows that

$$(7) \qquad \overline{\mathbf{W}} \Rightarrow \mathcal{G}_0 \text{ in } [\{\mathbb{D}[0, 1]\}^p, \mathscr{d}].$$

Now, let $\hat{f}_n$ be a density estimator based on $\{\hat{e}_{ni} := Y_{ni} - \mathbf{x}_{ni}'\hat{\boldsymbol{\beta}};\ 1 \le i \le n\}$ and $\hat{F}_n$ be the corresponding d.f.. Let $\{e^*_{ni};\ 1 \le i \le n\}$ represent i.i.d. $\hat{F}_n$ r.v.'s, i.e., $\{e^*_{ni};\ 1 \le i \le n\}$ is a random sample from the population $\hat{F}_n$. Because $\hat{F}_n$ is continuous, the resampling procedures based on it are usually called *smooth* bootstrap procedures. Let

$$(8) \qquad Y^*_{ni} := \mathbf{x}_{ni}'\hat{\boldsymbol{\beta}} + e^*_{ni}, \qquad 1 \le i \le n.$$

Define the bootstrap estimator $\boldsymbol{\beta}^*$ to be a solution $\mathbf{s} \in \mathbb{R}^p$ of the equation

$$(9) \qquad \Sigma_i\, \mathbf{A}\mathbf{x}_{ni}\, \{\psi(Y^*_{ni} - \mathbf{x}_{ni}'\mathbf{s}) - \hat{E}_n\psi(e^*_{n1})\} = 0.$$

where $\hat{E}_n$ is the expectation under $\hat{F}_n$. Let $\hat{P}_n$ denote the the bootstrap probability under $\hat{F}_n$. Finally, define

$$(10) \qquad \mathcal{S}^*(t, \mathbf{u}) := \Sigma_i\, \mathbf{A}\mathbf{x}_{ni}\, I(e^*_{ni} \le \hat{F}_n^{-1}(t) + \mathbf{x}_{ni}'\mathbf{A}\mathbf{u}), \qquad 0 \le t \le 1,\ \mathbf{u} \in \mathbb{R}^p,$$

and the vector of bootstrap w.e.p.'s

$$(11) \qquad \hat{\mathbf{W}}^*(t) := \Sigma_i\, \mathbf{A}\mathbf{x}_{ni}\, \{I(Y^*_{ni} - \mathbf{x}_{ni}'\boldsymbol{\beta}^* \le \hat{F}_n^{-1}(t)) - t\}, \qquad 0 \le t \le 1.$$

We also need

$$(12) \qquad \mathbf{W}^*(t) := \Sigma_i\, \mathbf{A}\mathbf{x}_{ni}\{I(e^*_{ni} \le \hat{F}_n^{-1}(t)) - t\}, \qquad 0 \le t \le 1.$$

Our goal is to show that $\hat{\mathbf{W}}^*$ converges weakly to $\mathcal{G}_0$ in $[\{\mathbb{D}[0, 1]\}^p, \mathscr{d}]$, a.s.. Here a.s. refers to almost all error sequences $\{e_i;\ i \ge 1\}$. We in fact have the following

Theorem 6.2b.1. *In addition to* (1.1.1), H_0, $(\mathbf{F}_0\mathbf{1})$, $(\mathbf{F}_0\mathbf{2})$, (**NX**) *and* (1), *assume that* ψ *is a bounded nondecreasing right continuous score function and that the following hold.*

(13) *For almost all error sequences* $\{e_i;\ i \ge 1\}$, $\hat{f}_n(x) > 0$ *for almost all* $x \in \mathbb{R}$, $n \ge 1$.

$$(14) \qquad \|\hat{f}_n - f_0\|_\infty \longrightarrow 0, \text{ a.s., } (P_0).$$

Then, $\forall\ 0 < B < \infty$,

$$(15) \qquad \sup \|\mathcal{S}^*(t, \mathbf{u}) - \mathcal{S}^*(t, 0) - \mathbf{u}\hat{f}_n(\hat{F}_n^{-1}(t))\| = o_p(1),\ (\hat{P}_n),\ \text{ a.s.},$$

where the supremum is over $0 \le t \le 1$, $\|\mathbf{u}\| \le B$.

Moreover, for almost all error sequences $\{e_i;\ i \ge 1\}$,

$$\text{(16)} \qquad \mathbf{A}^{-1}(\beta^* - \hat{\beta}) = -\hat{\kappa}_n \Sigma_i \mathbf{A}\mathbf{x}_{ni}\{\psi(e^*_{ni}) - \hat{E}_n\psi(e^*_{n1})\} + o_p(1), \quad (\hat{P}_n),$$

and

$$\text{(17)} \qquad \hat{\mathbf{W}}^* \Rightarrow \mathcal{G}_0 \ \text{ in } [\{\mathbb{D}[0,1]\}^p, d],$$

where $\hat{\kappa}_n := 1 / \int \hat{f}_n \, d\psi$.

Proof. Fix an error sequences $\{e_i;\ i \ge 1\}$ for which

$$\text{(14*)} \qquad \hat{f}_n(x) > 0, \text{ for almost all } x \in \mathbb{R}, \text{ and } \|\hat{f}_n - f_0\|_\infty \longrightarrow 0.$$

The following arguments are carried out conditional on this sequence.

Observe that $\mathcal{S}^*(t, \mathbf{u})$ is a p-vector of w.e.p.'s $S_d(t, \mathbf{u})$ of (2.3.1) whose j^{th} component has various underlying entities as follows:

$$\text{(18)} \qquad X_{ni} = e^*_{ni}, \ F_{ni} = \hat{F}_n, \ \mathbf{c}_{ni} = \mathbf{A}\mathbf{x}_{ni}, \ d_{ni} = \mathbf{a}'_{(j)}\mathbf{x}_{ni}, \ 1 \le i \le n$$

where, as usual, $\mathbf{a}_{(j)} = j^{th}$ column of $\mathbf{A}$, $1 \le j \le p$.

Thus, (15) follows from p applications of Theorem 2.3.1, j^{th} time applied to the above entities, provided we ensure the validity of the assumptions of that theorem. But, f_0 uniformly continuous and (14) readily imply that $\{\hat{f}_n,\ n \ge 1\}$ satisfies (2.3.3a,b). In view of (2.3.33), (2.3.34) and (**NX**), it follows that all other assumptions of Theorem 2.3.1 are satisfied. Hence, (15) follows from (2.3.6). In view of (13) we also obtain, from (2.3.7),

$$\text{(19)} \qquad \sup \|\mathcal{S}^{0*}(x, \mathbf{u}) - \mathcal{S}^{0*}(x, \mathbf{0}) - \mathbf{u}\hat{f}_n(x)\| = o_p(1), \quad (\hat{P}_n),$$

where $\mathcal{S}^{0*}(x, \mathbf{u}) \equiv \mathcal{S}^*(\hat{F}_n(x), \mathbf{u})$ and where the supremum is over $x \in \mathbb{R}$, $\|\mathbf{u}\| \le B$. Now, (16) follows from (19) in precisely the same fashion as does (4.2a.8) from (2.3.7).

From (11), (15), (16) and (31) below, we readily obtain that, under $\hat{P}_n$,

$$\text{(20)} \qquad \hat{\mathbf{W}}^*(t) = \Sigma_i \mathbf{A}\mathbf{x}_{ni}\{I(e^*_{ni} \le \hat{F}_n^{-1}(t)) - t - \hat{\kappa}_n \hat{q}_n(t)[\psi(e^*_{ni}) - \hat{E}_n\psi(e^*_{n1})]\} + o_p(1),$$

where $\hat{q}_n := \hat{f}_n(\hat{F}_n^{-1})$.

In analogy to (4) and (5), let $\hat{g}_n$, $\hat{\rho}_n$ stand for g_0, ρ_0 after F_0 is replaced by $\hat{F}_n$ in these entities. Thus

$$\text{(21)} \qquad \hat{g}_n(t) := \hat{E}_n\{I(e^*_{n1} \le \hat{F}_n^{-1}(t)) - t\}\ \psi(e^*_{n1})$$
$$= \int I(x \le \hat{F}_n^{-1}(t))\ \psi(x)\ d\hat{F}_n(x), \qquad 0 \le t \le 1,$$

and, for $0 \le t \le u \le 1$,

(22) $$\hat{\rho}_n(t, u) := t(1-u) - \hat{\kappa}_n[\hat{q}_n(t)\hat{g}_n(u)+\hat{g}_n(t)\hat{q}_n(u)] + \hat{\kappa}_n^2\hat{q}_n(t)\hat{q}_n(u)\hat{\sigma}_n^2.$$

where $\hat{\sigma}_n^2 := \hat{E}_n[\psi(e_{n1}^*) - \hat{E}_n\psi(e_{n1}^*)]^2$.

Let $\tilde{W}^*(t)$ denote the leading r.v. in the r.h.s. of (20). Observe that,

(23) $$\tilde{\mathcal{C}}_n(t, u) := \hat{E}_n\{\tilde{W}^*(t)\tilde{W}^*(u)'\} = \hat{\rho}_n(t, u)\, I_{p\times p}, \qquad 0 \le t \le u \le 1.$$

(24) **Claim:** $\hat{\rho}_n(t, u) \longrightarrow \rho_0(t, u), \quad \forall\ 0 \le t \le u \le 1.$

To prove (24), note that (14*) and Scheffé's Theorem (Lehmann, 1986, p573) imply that for the given error sequence $\{e_i\,;\ i \ge 1\}$,

(25) $$\delta_n := \|\hat{F}_n - F_0\|_\infty \longrightarrow 0,$$

which, together with the continuity of $\hat{F}_n$, yields

(26) $$\sup_{0\le t\le 1} |F_0(\hat{F}_n^{-1}(t)) - t| \longrightarrow 0.$$

Also, observe that

$$\sup_{0\le t\le 1} |\hat{f}_n(\hat{F}_n^{-1}(t)) - f_0(\hat{F}_n^{-1}(t))| \le \|\hat{f}_n - f_0\|_\infty \longrightarrow 0,$$

by (14*), and that,

$$|f_0(\hat{F}_n^{-1}(t)) - f_0(F_0^{-1}(t))| \equiv |q_0(F_0(\hat{F}_n^{-1}(t))) - q_0(t)|, \qquad \forall\ 0 \le t \le 1.$$

Hence, by (26) and the uniform continuity of q_0, which is implied by ($\mathbf{F}_0\mathbf{1}$),

(27) $$\sup_{0\le t\le 1}|\hat{q}_n(t) - q_0(t)| \longrightarrow 0.$$

Next, let $g_n(t) = \int I(\hat{F}_n(x) \le t)\psi(x)f_0(x)\,dx, \quad 0 \le t \le 1$. Upon rewriting $\hat{g}_n(t) = \int I(\hat{F}_n(x) \le t)\psi(x)\hat{f}_n(x)\,dx$, from (14*), Scheffé's Theorem (Lehmann: 1986, p 573) and the boundedness of ψ, we readily obtain that

$$\sup_{0\le t\le 1}|\hat{g}_n(t) - g_n(t)| \le \int |\hat{f}_n(x) - f_0(x)|\,dx \longrightarrow 0.$$

But, the inequality $F_0(x) - \delta_n \le \hat{F}_n(x) \le F_0(x) + \delta_n$ for all x, implies that

$$\begin{aligned} |g_n(t) - g_0(t)| &\le \|\psi\|_\infty \int I(F_0(x) - \delta_n \le t \le F_0(x) + \delta_n)\,dF_0(x), \\ &\le \|\psi\|_\infty\, 2\delta_n, \qquad \forall\ 0 \le t \le 1. \end{aligned}$$

Hence, by (25),

(28) $$\sup_{0\le t\le 1}|\hat{g}_n(t) - g_0(t)| \longrightarrow 0.$$

Again by the boundedness of ψ, (14*) and (25), one readily concludes that

(29) $$\hat{\kappa}_n \longrightarrow \kappa, \quad \hat{\sigma}_n^2 \longrightarrow \sigma^2.$$

Claim (24) now readily follows from (27) – (29).

Now recall (12) and rewrite $\tilde{\mathbf{W}}^*$ as

(30) $$\tilde{\mathbf{W}}^*(t) = \mathbf{W}^*(t) - \hat{\kappa}_n \, \hat{q}_n(t) \, \Sigma_i \, \mathbf{A}\mathbf{x}_{ni} \, [\psi(e^*_{ni}) - \hat{E}_n\psi(e^*_{n1})].$$

Observe that because

$$\hat{E}_n\|\Sigma_i \, \mathbf{A}\mathbf{x}_{ni} \, [\psi(e^*_{ni}) - \hat{E}_n\psi(e^*_{n1})]\|^2 = p \, \hat{\sigma}_n^2,$$

by (29) and the Markov inequality it follow that

(31) $$\|\Sigma_i \, \mathbf{A}\mathbf{x}_{ni} \, [\psi(e^*_{ni}) - \hat{E}_n\psi(e^*_{n1})]\| = O_p(1), \ (\hat{P}_n).$$

Apply Corollary 2.2a.1 p times, j^{th} time to the entities given at (18), to conclude that

$$\lim_{\eta\to 0} \limsup_n \hat{P}_n\Big(\sup_{|t-s|<\eta} |\mathbf{W}^*(t) - \mathbf{W}^*(s)| > \eta\Big) = 0.$$

This together with (31), (30), (27) and the uniform continuity of F_0 implies that the sequence of processes $\{\tilde{\mathbf{W}}^*\}$ is tight in the uniform metric $\mathscr{d}$ and all its subsequential limits must be in $\{\mathbb{C}[0, 1]\}^p$. Now, (17) follows from this, Claim (24), (20), (13), (14) and (6). □

Remark 6.2b.1. One of the main consequences of (17) is that one can use the bootstrap analogue of $\hat{D}_3$, v.i.z., $\hat{D}_3^* := \sup\{\|\hat{\mathbf{W}}^*(t)\|, \ 0\le t\le 1\}$ to carry out the test H_0. Thus an approximation to the the null distribution of $\hat{D}_3$ is obtained by the distribution of $\hat{D}_3^*$ under $\hat{P}_n$. In practice it means to obtain repeated random samples of size n from $\hat{F}_n$, compute the frequency distribution of $\hat{D}_3^*$ from these samples and use that to approximate the null distribution of $\hat{D}_3$. At least asymptotically this converges to the right distribution. Obviously the smooth bootstrap distributions for $\hat{D}_1$, $\hat{D}_2$ can be obtained similarly.

Reader might have realized that the conclusion (17) is true for any sequence of estimators $\{\hat{\beta}\}$, $\{\beta^*\}$ satisfying (2) and (16). □

6.3. L_2-DISTANCE TESTS

Let K_1^o and K_2^o, respectively, stand for the K_1 and $K_{\mathbf{X}}$ of (5.2.5) and (5.2.7) after the d.f.'s $\{H_{ni}\}$ there are replaced by F_0. Thus, for $G \in \mathbb{D}\mathcal{I}(\mathbb{R})$,

$$\text{(1)} \qquad K_1^o(\mathbf{t}) := \int \{W_1^o(y, \mathbf{t})\}^2 dG(y),$$

$$K_2^o(\mathbf{t}) := \int \|\mathbf{W}^o(y, \mathbf{t})\|^2 dG(y), \qquad \mathbf{t} \in \mathbb{R}^p,$$

where $\mathbf{W}^o$ is as in (6.1.3) and

$$\text{(2)} \qquad W_1^o(y, \mathbf{t}) := n^{1/2}[H_n(y, \mathbf{t}) - F_0(y)], \qquad y \in \mathbb{R},\ \mathbf{t} \in \mathbb{R}^p.$$

Let $\hat{\boldsymbol{\beta}}$ be an estimator of $\boldsymbol{\beta}$ and define the four test statistics

$$\text{(3)} \qquad K_j^* := \inf\{K_j(\mathbf{t});\ \mathbf{t} \in \mathbb{R}^p\}, \qquad \hat{K}_j := K_j(\hat{\boldsymbol{\beta}}), \qquad j = 1, 2.$$

The large values of these statistics are significant for testing H_0.

We shall first discuss the a.n.d.'s of K_j^*, $j = 1, 2$. Let $W_1^o(\cdot)$, $\mathbf{W}^o(\cdot)$ stand for $W_1^o(\cdot, \boldsymbol{\beta})$ and $\mathbf{W}^o(\cdot, \boldsymbol{\beta})$.

Theorem 6.3.1. *Assume that* (1.1.1), H_0, (NX), (5.5.68) – (5.5.70) *with* $F \equiv F_0$ *hold.*

(a) *If, in addition,* (5.6a.10) *and* (5.6a.11) *hold, then*

$$\text{(4)} \qquad K_1^* = \int \Big\{W_1^o(y) - f_0(y) \frac{\int W_1^o f_0\, dG}{\int f_0^2 dG}\Big\}^2 dG + o_p(1).$$

(b) *Under no additional assumptions,*

$$\text{(5)} \qquad K_2^* = \int \Big\|\mathbf{W}^o(y) - f_0(y) \frac{\int \mathbf{W}^o f_0\, dG}{\int f_0^2 dG}\Big\|^2 dG + o_p(1).$$

Proof. Apply Theorems 5.5.1 and 5.5.3 twice, once with $\mathbf{D} = n^{-1/2}(1, 0, \ldots, 0]$ and once with $\mathbf{D} = \mathbf{XA}$, and the rest of the entities as follows:

$$\text{(6)} \qquad Y_{ni} \equiv e_{ni}, \qquad H_{ni} \equiv F_0 \equiv F_{ni}, \qquad G_n \equiv G.$$

The theorem then follows from (5.5.28), (5.6a.5), (5.6a.12) and some algebra. See also Claim 5.5.2. □

Remark 6.3.1. Perhaps it is worthwhile repeating that (5) holds without any extra conditions on the design matrix $\mathbf{X}$. Thus, at least in this sense, K_2^* is a more natural statistic to use than K_1^* for testing H_0.

A consequence of (4) is that even if $\hat{\beta}_1$ of (5.2.4) is asymptotically non–unique, K_1^* asymptotically behaves like a unique sequence of r.v.'s. Moreover, unlike the $\hat{D}_1$–statistic, the asymptotic null distribution of K_1^* does not depend on the design matrix among all those designs that satisfy the given conditions.

The assumptions (5.6a.10) and (5.6a.11) are restrictive. For example, in the case $p = 1$, (5.6a.10) translates to requiring that either $x_{i1} \geq 0$ for all i or $x_{i1} \leq 0$ for all i. The assumption (5.6a.11) says that $\bar{\mathbf{x}} \neq \mathbf{0}$ or can not converge to $\mathbf{0}$. Compare this with the fact that if $\bar{\mathbf{x}} \approx 0$ then the asymptotic distribution of $\hat{D}_1$ does not depend on the preliminary estimator $\hat{\beta}$. □

Next, we need a result that will be useful in deriving the limiting distributions of certain quadratic forms involving w.e.p.'s. To that effect, let $L_2^p(\mathbb{R}, G)$ be the equivalence classes of measurable functions h: $\mathbb{R}$ to $\mathbb{R}^p$ such that $|h|_G^2 := \int \|h\|^2 dG < \infty$. The equivalence classes are defined in terms of the norm $|\cdot|_G^2$. In the following lemma, $\{\mathbf{a}_i;\ i \geq 1\}$ is a fixed orthonormal basis in $L_2^p(\mathbb{R}, G)$.

Lemma 6.3.1. *Let* $\{\mathbf{Z}_n,\ n \geq 1\}$ *be a sequence of p–vector stochastic processes with* $E\mathbf{Z}_n = \mathbf{0}$, $\mathrm{Cov}(\mathbf{Z}_n(x),\ \mathbf{Z}_n(y)) := \mathbf{K}_n(x, y) = ((K_{nij}(x, y)))$, $1 \leq i, j \leq p$, $x, y \in \mathbb{R}$. *In addition, assume the following:*

There is a covariance matrix function $\mathbf{K}(x, y) = ((K_{ij}(x, y)))$, *and a p–vector mean zero covariance–*$\mathbf{K}$ *Gaussian process* $\mathbf{Z}$ *such that*

(i) (a) $\quad \sum_{j=1}^{p} \int K_{njj}(x, x)dG(x) < \infty,\ n \geq 1.$ (b) $\sum_{j=1}^{p} \int K_{jj}(x, x)dG(x) < \infty.$

(ii) $\quad \sum_{j=1}^{p} \int K_{njj}(x, x)dG(x) \longrightarrow \sum_{j=1}^{p} \int K_{jj}(x, x)dG(x).$

(iii) *For every* $m \geq 1$,

$$\left(\int \mathbf{Z}_n' \mathbf{a}_1 dG, \ldots, \int \mathbf{Z}_n' \mathbf{a}_m dG\right) \xrightarrow[d]{} \left(\int \mathbf{Z}' \mathbf{a}_1 dG, \ldots, \int \mathbf{Z}' \mathbf{a}_m dG\right);$$

(iv) *For each* $i \geq 1$,

$$E\left(\int \mathbf{Z}_n' \mathbf{a}_i dG\right)^2 \longrightarrow E\left(\int \mathbf{Z}' \mathbf{a}_i dG\right)^2.$$

Then, $\mathbf{Z}_n$, $\mathbf{Z}$ *belong to* $L_2^p(\mathbb{R}, G)$, *and*

(7) $$\mathbf{Z}_n \Rightarrow \mathbf{Z} \ \text{ in } L_2^p(\mathbb{R}, G).$$

Proof: In view of Theorem VI.2.2 of Parthasarthy (1967) and in view of (iii), it suffices to show that for any $\epsilon > 0$, there is an $N\ (= N\epsilon)$ such that

$$\text{(8)} \qquad \sup_n E \sum_{i\geq N} \left(\int Z_n' a_i dG\right)^2 \leq \epsilon.$$

Because of the properties of $\{a_i\}$, Fubini and (i),

$$\text{(9)} \qquad \sum_{j=1}^{p} \int K_{njj}(x, x) dG(x) = E|Z_n|_G^2 = \sum_{i\geq 1} E\left(\int Z_n' a_i dG\right)^2,$$

$$\text{(10)} \qquad \sum_{j=1}^{p} \int K_{jj}(x, x) dG(x) = E|Z|_G^2 = \sum_{i\geq 1} E\left(\int Z' a_i dG\right)^2.$$

Thus, to prove (8), it suffices to exhibit an N such that

$$\text{(11)} \qquad \sup_n \sum_{i\geq N} E\left(\int Z_n' a_i dG\right)^2 \leq \epsilon.$$

By (ii), (9) and (10), there exists $N_{1\epsilon}$ such that

$$\text{(12)} \qquad \sum_{i\geq 1} E\left(\int Z_n' a_i dG\right)^2 \leq \sum_{i\geq 1} E\left(\int Z' a_i dG\right)^2 + \epsilon/3, \qquad n \geq N_{1\epsilon}.$$

By (i)(b) and (10), there exists $N(= N_\epsilon)$ such that

$$\text{(13)} \qquad \sum_{i\geq N} E\left(\int Z' a_i dG\right)^2 \leq \epsilon/3.$$

By (iv), there exists $N_{2\epsilon}$ such that

$$\text{(14)} \qquad \sum_{i<N} E\left(\int Z' a_i dG\right)^2 \leq \sum_{i<N} E\left(\int Z_n' a_i dG\right)^2 + \epsilon/3, \qquad n \geq N_{2\epsilon}.$$

Therefore, from (12) – (14), with $N = N_\epsilon := N_{1\epsilon} \vee N_{2\epsilon}$,

$$\begin{aligned} &\sup_{n\geq N} \sum_{i<N} E\left(\int Z_n' a_i dG\right)^2 \\ &\qquad \leq \sup_{n\geq N} \Big[\sum_{i\geq 1} E\left(\int Z' a_i dG\right)^2 - \sum_{i<N} \left(\int Z_n' a_i dG\right)^2\Big] + \epsilon/3 \leq \epsilon. \end{aligned}$$

Use (i)(a) to take care of the case $n < N_\epsilon$. This proves the result. □

Remark 6.3.2. Millar (1981) contains a special case of the above lemma where $p = 1$, Z_n is the standardized ordinary e.p. and Z is the Brownian bridge. The above lemma is an extension of Millar's result to cover more general processes like the w.e.p.'s under general independent setting. In applications of the above lemma, one may choose $\{\mathbf{a}_i\}$ to be such that the support S_i of $\mathbf{a}_i$ has $G(S_i) < \infty$, $i \geq 1$ and such that $\{\mathbf{a}_i\}$ are bounded. □

Corollary 6.3.1. (a) *Under the conditions of Theorem* 6.3.1(a),

$$K_1^* \xrightarrow[d]{} \int \Big\{B(F_0) - f_0 \cdot \frac{\int B(F_0) f_0 \, dG}{\int f_0^2 dG}\Big\}^2 dG =: \mathcal{G}_1, \qquad \text{(say)}. \tag{15}$$

(b) *Under the conditions of Theorem* 6.3.1(b),

$$K_2^* \xrightarrow[d]{} \int \Big\| \mathbf{B}(F_0) - f_0 \cdot \frac{\int \mathbf{B}(F_0) f_0 \, dG}{\int f_0^2 dG} \Big\|^2 =: \mathcal{G}_2, \qquad \text{(say)}. \tag{16}$$

Here B, $\mathbf{B}$ *are is as in* (6.2a.7), (6.2a.8).

Proof: (b) Apply Lemma 6.3.1, with $\mathbf{a}_i$ as in the Remark 6.3.2 above, to

$$\mathbf{Z}_n = \mathbf{W}^o - \frac{\int \mathbf{W}^o f_0 \, dG}{\int f_0^2 dG} \cdot f_0, \qquad \mathbf{Z} = \mathbf{B}(F_0) - \frac{\int \mathbf{B}(F_0) f_0 dG}{\int f_0^2 dG} \cdot f_0.$$

Direct calculations show that $E\mathbf{Z}_n = \mathbf{0} = E\mathbf{Z}$, and $\forall\ x, y \in \mathbb{R}$,

$$\mathbf{K}_n(x, y) := E\mathbf{Z}_n(x)\mathbf{Z}_n'(y) = \mathbf{I}_{p\times p}\, \ell(x, y) = \mathbf{K}(x, y) =: E\mathbf{Z}(x)\mathbf{Z}'(y),$$

where, for $x, y \in \mathbb{R}$,

$$\ell(x, y) := k(x, y) - a^{-1} f_0(y) \int k(x, s)\, d\psi(s) - a^{-1} f_0(y) \int k(y, s)\, d\psi(s) + \\ + a^{-2} \int\int k(s, t)\, d\psi(s) d\psi(t),$$

$$k(x, y) := F_0(x \wedge y) - F_0(x)F_0(y), \quad \psi(x) = \int_{-\infty}^{x} f_0 dG, \quad a = \psi(\infty).$$

Therefore, (5.5.68), (5.5.69) imply (i), (ii) and (iv). To prove (iii), let $\lambda_1, \ldots, \lambda_m$ be real numbers. Then,

$$\sum_{j=1}^{m} \lambda_j \int \mathbf{Z}_n' \mathbf{a}_j dG = \int \mathbf{W}^{o\prime} \mathbf{b}\, dG - \frac{\int \mathbf{W}^{o\prime} d\psi}{\int f_0 d\psi} \cdot \int \mathbf{b}\, d\psi =: h(\mathbf{W}^o), \qquad \text{(say)},$$

where $\mathbf{b} := \sum_{j=1}^{m} \lambda_j \mathbf{a}_j$. Because ψ and $\mathbf{b}dG$ are finite measures, $h(\mathbf{W}^o)$ is a uniformly continuous function of $\mathbf{W}^o$. Thus by Lemma 6.2a.2 and Theorem

5.1 of Billingsley (1968), $h(\mathbf{W}^o) \underset{d}{\rightarrow} h(B(F_0))$, under H_0 and (NX). This then verifies all conditions of Lemma 6.3.1. Hence $\mathbf{Z}_n \Rightarrow \mathbf{Z}$ in $L_2^p(\mathbb{R}, G)$. In particular $\int \|Z_n\|^2 dG \underset{d}{\rightarrow} \int \|Z\|^2 dG$. This and (5) proves (16). The proof of (15) is similar. □

Remark 6.3.3. The r.v. $\overline{G}_1$ can be rewritten as

$$\overline{G}_1 = \int B^2(F_0)dG - \frac{\{\int B(F_0) f_0 dG\}^2}{\int f_0^2 dG}$$

Recall that $\overline{G}_1$ is the same as the limiting r.v. obtained in the one sample location model. Its distribution for various G and F_0 has been theoretically studied by Martynov (1975). Boos (1981) has tabulated some critical values of $\overline{G}_1$ when $dG = \{F_0(1 - F_0)\}^{-1} dF_0$ and F_0 = Logistic. From Anderson–Darling or Boos one obtains that in this case

$$\overline{G}_1 = \int_0^1 B^2(t)(t(1-t))^{-1}dt - 6\Big(\int_0^1 B(t)dt\Big)^2 = \sum_{j\geq 2} N_j^2/j(j+1)$$

where $\{N_j\}$ are i.i.d. $N(0, 1)$ r.v.'s. From Boos (Table 3), one obtains the following

Table II

α	.005	.01	.025	.05
$t\alpha$	1.710	1.505	1.240	1.046

In Table II, $t\alpha$ is such that $P(\overline{G}_1 > t\alpha) = \alpha$. For some other tables see Stephens (1979).

The r.v. $\overline{G}_2$ can be rewritten as

$$\begin{aligned}\overline{G}_2 &:= \int \|B(F_0)\|^2 dG - \frac{\|\int B(F_0) f_0 dG\|^2}{\int f_0^2 dG} \\ &= \sum_{j=1}^{p} \Big[\int B_j^2(F_0)dG - \frac{(\int B_j(F_0) f_0 dG)^2}{\int f_0^2 dG}\Big],\end{aligned}$$

which is a sum of p independent r.v.'s identically distributed as $\overline{G}_1$. The distribution of such r.v.'s does not seem to have been studied yet. Until the distribution of $\overline{G}_2$ is tabulated one could use the independence of the summands in $\overline{G}_2$ and the bounds between the sum and the maximum to obtain a crude approximation to the significance level.

For p = 1, the a.n.d. of K_1^* and K_2^* is the same but the conditions under which the results for K_1^* hold are stronger than those for K_2^*. □

The next result gives an approximation for $\hat{K}_j$, j = 1, 2. It also follows from Theorem 5.5.1 in a fashion similar to the previous theorem, and hence no details are given.

Theorem 6.3.2. *Assume that* (1.1.1), H_0, (NX), (5.5.68) – (5.5.70) *with* $F \equiv F_0$ *and* (6.2a.3) *hold. Then,*

$$\hat{K}_1 = \int [W_1^o(y) + n^{1/2}\, \bar{x}A\cdot A^{-1}(\hat{\beta}-\beta)f_0(y)]^2 dG(y) + o_p(1). \tag{17}$$

$$\hat{K}_2 = \int \|W^o(y) + A^{-1}(\hat{\beta}-\beta)f_0(y)\|^2 dG(y) + o_p(1). \qquad \square$$

From this we can obtain the asymptotic null distribution of these statistics when $\hat{\beta}$ is estimated efficiently for the large samples as follows. Recall the definition of $\{s_i\}$ from (6.2a.13) and let

$$\gamma_i(y) := I(e_i \le y) - F_0(y) + n\bar{x}'(X'X)^{-1}x_i\, s_i I_0^{-1} f_0(y),$$

$$\alpha_i(y) := I(e_i \le y) - F_0(y) + s_i I_0^{-1} f_0(y), \qquad 1 \le i \le n, \qquad y \in \mathbb{R},$$

$$\alpha = (\alpha_1, \ldots, \alpha_n)', \qquad \gamma = (\gamma_1, \ldots, \gamma_n)'.$$

Also, define

$$Z_{n1}(y) := W_1^o(y) + n^{1/2}\bar{x}'AAX'sI_0^{-1}f_0(y) = n^{-1/2}\sum_{i=1}^{n}\gamma_i(y) \tag{19}$$

$$Z_{n2}(y) := W^o(y) + AX'sI_0^{-1}f_0(y) = AX'\alpha(y), \qquad y \in \mathbb{R}.$$

From Theorem 6.3.2 we readily obtain the

Corollary 6.3.2. *Assume that* (1.1.1), H_0, (NX), (5.5.68) – (5.5.70) *with* $F \equiv F_0$, (6.2a.12) *and* (6.2a.14) *hold. Then,*

$$\hat{K}_1 = \int Z_{n1}^2 dG + o_p(1). \tag{20}$$

$$\hat{K}_2 = \int \|Z_{n2}\|^2 dG + o_p(1). \tag{21}$$ □

Next, observe that for $y \le z$,

$$
\begin{aligned}
K_{n1}(y, z) &:= \mathrm{Cov}(Z_{n1}(y), Z_{n1}(z)) \\
&= F_0(y)(1-F_0(z)) - n\bar{x}'(X'X)^{-1}\bar{x}\,\frac{f_0(y)f_0(z)}{I_0} =: \ell_{n1}(y, z), \\
K_{n2}(y, z) &:= EZ_{n2}(y)Z_{n2}'(z) \\
&= \{F_0(y)(1 - F_0(z)) - \frac{f_0(y)f_0(z)}{I_0}\}I_{p\times p} =: r_0(y, z), \quad \text{say.}
\end{aligned}
$$

Now apply Lemma 6.3.1 and argue just as in the proof of Corollary 6.3.1 to conclude

Corollary 6.3.3. (a). *In addition to the conditions of Corollary* 6.3.2, *assume that*

$$(22) \qquad n\bar{x}'(X'X)^{-1}\bar{x} \longrightarrow c, \quad |c| < \infty.$$

Then,

$$(23) \qquad \hat{K}_1 \xrightarrow[d]{} \int Z_1^2(y)dG(y)$$

where Z_1 *is a Gaussian process in* $L_2(\mathscr{R}, G)$ *with the covariance function*

$$(24) \qquad K_1(x, y) := F_0(x)(1 - F_0(y)) - cf_0(x)f_0(y)\, I_0^{-1}, \qquad x \leq y.$$

(b) *Under the conditions of Corollary* 6.3.2,

$$(25) \qquad \hat{K}_2 \xrightarrow[d]{} \int \| Y_0\|^2 dG$$

where Y_0 *is a vector of* p *independent Gaussian processes in* $L_2^p(\mathbb{R}, G)$ *with the covariance matrix* $r_0 \cdot I_{p\times p}$. □

Remark 6.3.4. Again, observe that the test statistic $\hat{K}_1$ based on the ordinary empirical of the residuals has an a.n.d. which is design dependent whereas the a.n.d. of the test based on the weighted empiricals $\hat{K}_2$ is design free. In fact, for $p = 1$, the limiting r.v. in (25) is the same as the one that appears in the one sample location model. For $G = F_0 = N(0, 1)$ d.f., Martynov (1976) has tabulated the distribution of this r.v.. Stephens (1976) has also tabulated the distribution of this r.v. for $G = F_0$, $dG = dG_0 = \{F_0(1 - F_0)\}^{-1}dF_0$, and for $F_0 = N(0, 1)$. For $G = F_0$, $F_0 = N(0, 1)$ d.f., Stephens and Martynov's tables generally agree up to the two decimal places, though occasionally there is an agreement up to three decimal places. In any case, for $p = 1$, one could use these tables to implement the test based on $\hat{K}_2$, at least asymptotically, whereas the test based on $\hat{K}_1$, being design dependent, can not be readily implemented. For the sake of convenience we reproduce some of the Stephens (1976, 1979) tables below.

Table III

$F_0 = N(0, 1)$

$\hat{K}_2 \backslash \alpha$	0.10	.025	.05	.10
$\hat{K}_2(F_0)$	.237	.196	.165	.135
$\hat{K}_2(G_0)$	1.541	1.281	1.088	.897

In Table III, $\hat{K}_2(G)$ stands for the $\hat{K}_2$ with G being the integrating measure. $\hat{K}_2(G_0)$ is the $\hat{K}_2$ with the Anderson–Darling weights. Table III is, of course, useful only when $p = 1$. □

As far as the *asymptotic power* of the above L_2-tests is concerned, it is apparent that Theorems 5.5.1, 5.5.3 and Lemma 6.3.1 can be used to deduce the asymptotic power of these tests against fairly general alternatives. Here we shall discuss the asymptotic behavior of only K_j^*, $j = 1, 2$ under the heteroscedastic gross errors alternatives. More precisely, suppose that

$$F_{ni} = (1 - \delta_{ni})F_0 + \delta_{ni}F_1, \quad 0 \le \delta_{ni} \le 1, \quad \max_i \delta_{ni} \longrightarrow 0, \tag{26}$$

F_1 a fixed d.f. Let

$$m_1 := n^{-1/2} \Sigma_i \, \delta_{ni}(F_1 - F_0), \qquad \mathbf{m}_2 := \Sigma_i \, \mathbf{A}\mathbf{x}_{ni}\delta_{ni}(F_1 - F_0).$$

Lemma 6.3.2. *Let* (1.1.1) *hold with* e_{ni} *having the* d.f. F_{ni} *given by* (26), $1 \le i \le n$. *Suppose that* $\mathbf{X}$ *satisfies* (NX); (F_0, G) *and* (F_1, G) *satisfy* (5.5.68) – (5.5.70) *and that*

$$\int |F_1 - F_0| \, dG < \infty, \tag{27}$$

(a) *If, in addition,* (5.6a.10) *and* (5.6a.11) *hold, then*

$$K_1^* = \int \Big\{ W_1^o + m_1 - f_0 \frac{\int (W_1^o + m_1) f_0 \, dG}{\int f_0^2 \, dG} \Big\}^2 dG + o_p(1) \tag{28}$$

provided

$$n^{-1/2} \Sigma_i \, \delta_{ni} = O(1). \tag{29}$$

(b) *Without any additional conditions,*

$$(30) \qquad K_2^* = \int \|\mathbf{W}^o + \mathbf{m}_2 - f_0 \frac{\int (\mathbf{W}^o + \mathbf{m}_2) f_0 dG}{\int f_0^2 dG}\|^2 dG + o_p(1),$$

provided

$$(31) \qquad \Sigma_i \mathbf{A}\mathbf{x}_{ni}\delta_{ni} = O(1).$$

Proof. Apply Theorem 5.5.1 and (5.5.49) to $D = n^{-1/2}[1, 0, ..., 0]$, $Y_{ni} \equiv e_{ni}$, $H_{ni} \equiv F_0$, $\{F_{ni}\}$ given by (26) to conclude (a). Apply the same results to $D = \mathbf{A}\mathbf{X}$ and the rest of the entities as in the proof of (a) to conclude (b). □

Now apply Lemma 6.3.1 to

$$(30) \qquad Z_n := W^o + m_1 - f_0 \frac{\int (W^o + m_1) f_0 dG}{\int f_0^2 dG},$$

$$Z := \mathcal{B}(F_0) + a_1(F_1 - F_0) - f_0 \frac{\int \{\mathcal{B}(F_0) + a_1(F_1 - F_0)\} f_0 dG}{\int f_0^2 dG},$$

where $a_1 := \limsup_n n^{-1/2} \Sigma_i \delta_{ni}$, to obtain

Corollary 6.3.4. *Under the conditions of Lemma* 6.3.2(a),

$$K_1^* \underset{d}{\longrightarrow} \int Z^2 dG, \quad \textit{where } Z \textit{ is as in } (30). \qquad □$$

Similarly, apply Lemma 6.3.1 to

$$(31) \qquad \mathbf{Z}_n := \mathbf{W}^o + \mathbf{m}_2 - f_0 \frac{\int (\mathbf{W}^o + \mathbf{m}_2) f_0 dG}{\int f_0^2 dG},$$

$$\mathbf{Z} := \mathcal{B}(F_0) + \mathbf{a}_2(F_1 - F_0) - f_0 \frac{\int \{\mathcal{B}(F_0) + \mathbf{a}_2(F_1 - F_0)\} f_0 dG}{\int f_0^2 dG},$$

where $\mathbf{a}_2 = \limsup_n \Sigma_i \mathbf{A}\mathbf{x}_{ni}\delta_{ni}$, to obtain

Corollary 6.3.5. *Under the conditions of Lemma* 6.3.2(b),

$$K_2^* \underset{d}{\longrightarrow} \int \|\mathbf{Z}\|^2 dG, \quad \textit{where } \mathbf{Z} \textit{ is as in } (31). \qquad □$$

An interesting choice of $\delta_{ni} = p^{-1/2}\|\mathbf{A}\mathbf{x}_{ni}\|$. Another choice is $\delta_{ni} \equiv n^{-1/2}$. Both a priori satisfy (26), (29) and (31). □

6.4. TESTING WITH UNKNOWN SCALE

Now consider (1.1.1) and the problem of testing H_1 of (6.1.4). Here we shall discuss the modifications of $\hat{D}_j$, $\hat{K}_j$, $j = 1, 2$, of Sections 6.2, 6.3 that will be suitable for H_1. With W_1^o, $\mathbf{W}^o$ as before, define

$$
\begin{aligned}
(1)\qquad & D_1(a, \mathbf{u}) := \sup_y |W_1^o(ay, \mathbf{u})|, \\
& D_2(a, \mathbf{u}) := \sup_y |\mathbf{W}^o(ay, \mathbf{u})|, \\
& K_1(a, \mathbf{u}) := \int \{W_1^o(ay, \mathbf{u})\}^2 dG(y), \\
& K_2(a, \mathbf{u}) := \int \|\mathbf{W}^o(ay, \mathbf{u})\|^2 dG(y), \qquad a > 0,\ \mathbf{u} \in \mathbb{R}^p.
\end{aligned}
$$

Let $(\tilde{\sigma}, \tilde{\boldsymbol{\beta}})$ be estimators of $(\sigma, \boldsymbol{\beta})$, $\tilde{D}_j$ and $\tilde{K}_j$ stand for $D_j(\tilde{\sigma}, \tilde{\boldsymbol{\beta}})$ and $K_j(\tilde{\sigma}, \tilde{\boldsymbol{\beta}})$, respectively, $j = 1, 2$. The following two theorems give the a.n.d.'s of these statistics. Theorem 6.4.1 follows from Corollary 2.3.4 in a similar fashion as does Theorem 6.2.1 from Corollaries 2.3.3 and 2.3.5. Theorem 6.4.2 follows from Theorems 5.5.8 in a similar fashion as does Theorem 6.3.2 from Theorem 5.5.1. Recall the conditions ($\mathbf{F_0 1}$) and ($\mathbf{F_0 3}$) from Section 2.3.

Theorem 6.4.1. *In addition to* (1.1.1) *and* H_1, *assume that* (**NX**), ($\mathbf{F_0 1}$), ($\mathbf{F_0 3}$) *and the following hold.*

$$(2)\quad \text{(a)}\ |n^{1/2}(\tilde{\sigma} - \sigma)\sigma^{-1}| = O_p(1). \quad \text{(b)}\ \|\mathbf{A}^{-1}(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta})\| = O_p(1).$$

Then,

$$\tilde{D}_1 = \sup |W_1(t) + q_0(t)\{n^{1/2}\bar{\mathbf{x}}_n'(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta}) + n^{1/2}(\tilde{\sigma}-\sigma)F_0^{-1}(t)\}\sigma^{-1}| + o_p(1),$$

and

$$\tilde{D}_2 = \sup \|\mathbf{W}(t) + q_0(t)\{\mathbf{A}^{-1}(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta}) + n^{1/2}\mathbf{A}\bar{\mathbf{x}}_n \cdot n^{1/2}(\tilde{\sigma}-\sigma)F_0^{-1}(t)\}\sigma^{-1}\| + o_p(1),$$

where now $W_1(\cdot) := W_1^o(\sigma F_0^{-1}(\cdot), \boldsymbol{\beta})$ *and* $\mathbf{W}(\cdot) := \mathbf{W}^o(\sigma F_0^{-1}(\cdot), \boldsymbol{\beta})$.

Theorem 6.4.2. *In addition to* (1.1.1) *and* H_1, *assume that* (**NX**), (2), (5.5.69) *with* $F = F_0$, *and the following hold.*

(3) F_0 *has a continuous density* f_0 *such that*

(a) $0 < \int |y|^j f_0^k(y) dG(y) < \infty$, $j = 0, k = 1, 2$; $j = 2, k = 2$.

(b) $\lim_{s\to 0} \limsup_n \int f_0^k(y+\tau n^{-1/2}+s)dG(y) = \int f_0^k\, dG(y)$, $k = 1, 2$, $\tau \in \mathbb{R}$.

(c) $\lim_{s\to 0} \int |y|\, f_0(y(1+s))dG(y) = \int |y|\, f_0(y)dG(y)$.

Then,

$$\tilde{K}_1 = \int [W_1^o(\sigma y, \beta) + f_0(y)\{n^{1/2}\bar{x}_n'(\tilde{\beta}-\beta) + n^{1/2}(\tilde{\sigma}-\sigma)y\}\sigma^{-1}]^2 dG(y) + o_p(1),$$

$$\tilde{K}_2 = \int \|W^o(\sigma y, \beta) + f_0(y)\{A^{-1}(\tilde{\beta}-\beta) + n^{1/2}A\bar{x}_n \cdot n^{1/2}(\tilde{\sigma}-\sigma)y\}\sigma^{-1}\|^2 dG(y) + o_p(1).$$

Clearly, from these theorems one can obtain an analogue of Corollary 6.3.2 when $(\tilde{\sigma}, \tilde{\beta})$ are chosen to be asymptotically efficient estimators.

As is the case in the classical least square theory or in the M-estimation methodology, neither of the two dispersions $K_1(a, \mathbf{u})$ and $K_2(a, \mathbf{u})$ can be used to satisfactorily estimate (σ, β) by the simultaneous minimization process. The analogues of the m.d. goodness-of-fit tests that should be used are $\inf\{K_j(\tilde{\sigma}, \mathbf{u});\ \mathbf{u} \in \mathbb{R}^p\}$, $j = 1, 2$. The methodology of Section 5 may be used to obtain the asymptotic distributions of these statistics in a fashion similar to the above. □

6.5. TESTING FOR SYMMETRY OF THE ERRORS

Consider the model (1.1.1) and the hypothesis H_s of symmetry of the errors specified at (6.1.5). The proposed tests are to be based on $\hat{D}_{js}$, $j = 1, 2, 3$, of (6.1.6), (6.1.7), $K_j^+(\hat{\beta})$, and $\inf\{K_j^+(\mathbf{t});\ \mathbf{t} \in \mathbb{R}^p\}$, $j = 1, 2$, where

(1) $$K_1^+(\mathbf{t}) := \int \{W_1^+(y, \mathbf{t})\}^2 dG(y), \quad K_2^+(\mathbf{t}) := \int \|\mathbf{W}^+(y, \mathbf{t})\|^2 dG(y), \quad \mathbf{t} \in \mathbb{R}^p,$$

with W_1^+ and $\mathbf{W}^+$ as in (6.1.7) and (6.1.9). Large values of these statistics are considered to be significant for H_s.

Although the results of Chapters 2 and 5 can be used to obtain their asymptotic behavior under fairly general alternatives, here we shall focus only on the a.n.d.'s of these tests. To state these, we need some more notation. For a d.f. F, define

(2) $$F_+(y) := F(y) - F(-y), \qquad y \geq 0.$$

Then, with F^{-1} denoting the usual inverse of a d.f. F, we have

$$(3) \qquad F_+^{-1}(t) = F^{-1}((1+t)/2), \qquad -F_+^{-1}(t) = F^{-1}((1-t)/2), \qquad 0 \le t \le 1,$$

for all F that are continuous and symmetric around 0. Finally, let

$$(4) \qquad W_1^*(t) := W_1^+(F_+^{-1}(t), \beta), \quad W^*(t) := W^+(F_+^{-1}(t), \beta),$$
$$q^+(t) := f(F_+^{-1}(t)), \qquad 0 \le t \le 1.$$

We are now ready to state and prove

Theorem 6.5.1. *In addition to* (1.1.1), H_s *and* (**NX**), *assume that* F *in* H_s *and the estimator* $\hat{\beta}$ *satisfy* (**F1**) *and*

$$(5) \qquad \|A^{-1}(\hat{\beta} - \beta)\| = O_p(1), \qquad \textit{under } H_s.$$

Then,

$$(6) \qquad \hat{D}_{1s} = \sup_{0 \le t \le 1} |W_1^*(t) + 2q^+(t)\, n^{1/2}\bar{x}_n' A\, A^{-1}(\hat{\beta} - \beta)| + o_p(1),$$

$$(7) \qquad \hat{D}_{2s} = \sup_{0 \le t \le 1} |W^*(t) + 2q^+(t)\, A^{-1}(\hat{\beta} - \beta)| + o_p(1).$$

and

$$(8) \qquad \hat{D}_{3s} = \sup_{0 \le t \le 1} \|W^*(t) + 2q^+(t)\, A^{-1}(\hat{\beta} - \beta)\| + o_p(1).$$

Proof. The proof follows from Theorem 2.3.1 in the following fashion. The details will be given only for (8), as they are the same for (7) and quite similar for (6). Because F is continuous and symmetric around 0 and because $W^+(\cdot, \cdot) \equiv W^+(-\cdot, \cdot)$, $\hat{D}_{3s} =_d \sup_{0 \le t \le 1} W^+(F_+^{-1}(t), \hat{\beta})$. But, from the definition (6.1.8) and (3), it follows that for a $v \in \mathbb{R}^p$,

$$W^+(F_+^{-1}(t), v)$$
$$= \Sigma_i\, A x_{ni}\{I(e_{ni} \le F^{-1}(\tfrac{1+t}{2}) + c_{ni}' u) + I(e_{ni} \le F^{-1}(\tfrac{1-t}{2}) + c_{ni}' u) - 1\}$$
$$(9) \qquad = S(\tfrac{1+t}{2}, u) + S(\tfrac{1-t}{2}, u) - \Sigma_i\, A x_{ni}, \qquad 0 \le t \le 1,$$

where

$$S(t, u) := \Sigma_i\, A x_{ni}\, I(e_{ni} \le F^{-1}(t) + c_{ni}' u), \qquad 0 \le t \le 1,$$

is a p–vector of S_d–processes of (2.3.1) with $X_{ni} \equiv e_{ni}$, $F_{ni} \equiv F \equiv H$, $c_{ni} \equiv$

$\mathbf{Ax}_{ni}$, $\mathbf{u} = \mathbf{A}^{-1}(\mathbf{v} - \boldsymbol{\beta})$ and where the j^{th} process has the weights $\{d_{ni}\}$ given by the j^{th} column of $\mathbf{AX}$. The assumptions about F and $\mathbf{X}$ imply all the assumptions of Theorem 2.3.1. Hence (8) follows from (3.2.6), (5) and (9) in an obvious fashion. □

Next, we state an analogous result for the L_2-distances.

Theorem 6.5.2. *In addition to* (1.1.1), H_s, (NX) *and* (5), *assume that* F *in* H_s *and the integrating measure* G *satisfy* (5.3.8), (5.5.68), (5.5.70) *and* (5.6a.13). *Then,*

$$(10) \qquad K_1^+(\hat{\beta}) = \int [W_1^+(y) + 2f(y)\, n^{1/2}\bar{\mathbf{x}}_n'(\hat{\beta} - \beta)]^2 dG(y) + o_p(1),$$

$$(11) \qquad K_2^+(\hat{\beta}) = \int \|\mathbf{W}^+(y) + 2f(y)\, \mathbf{A}^{-1}(\hat{\beta} - \beta)\|^2 dG(y) + o_p(1),$$

where $W_1^+(\cdot)$, $\mathbf{W}^+(\cdot)$ *now stand for* $W_1^+(\cdot, \beta)$, $\mathbf{W}^+(\cdot, \beta)$.

Proof. The proof follows from two applications of Theorem 5.5.2, once with $\mathbf{D} = n^{-1/2}[1, 0, \ldots, 0]$ and once with $\mathbf{D} = \mathbf{XA}$. In both cases, take Y_{ni} and F_{ni} of that theorem to be equal to e_{ni} and F, $1 \le i \le n$, respectively. The Claim 5.5.2 justifies the applicability of that theorem under the present assumptions. □

The next result is useful in obtaining the a.n.d.'s of the m.d. test statistics. Its proof uses Theorem 5.5.2 and 5.5.4 in a similar fashion as Theorems 5.5.1 and 5.5.3 are used in the proof of Theorem 6.3.1, and hence no details are given. Let

$$K_j^s := \inf\{K_j^+(\mathbf{t});\ \mathbf{t} \in \mathbb{R}^p\}, \quad j = 1, 2.$$

Theorem 6.5.3. *Assume that* (1.1.1), H_s, (NX), (5.3.8), (5.5.68), (5.5.70) *and* (5.6a.13) *hold.*

(a) *If, in addition,* (5.6a.10) *and* (5.6a.11) *hold, then*

$$(12) \qquad K_1^s = 2\int_0^\infty \Big\{W_1^+(y) - f(y)\int_0^\infty W_1^+ f dG \,\big(\int_0^\infty f^2 dG\big)^{-1}\Big\}^2 dG + o_p(1).$$

(b) *Under no additional assumptions,*

$$(13) \qquad K_2^s = 2\int_0^\infty \|\mathbf{W}^+(y) - f(y)\int_0^\infty \mathbf{W}^+ f dG \,\big(\int_0^\infty f^2 dG\big)^{-1}\|^2 dG + o_p(1). \qquad □$$

To obtain the a.n.d.'s of the given statistics from the above theorem we now apply Lemma 6.3.1 to the approximating processes. The details will be given for K_2^s only as they are similar for K_1^s. Accordingly, let

$$\mathbf{Z}_n(y) := \mathbf{W}^+(y) - f(y)\int_0^\infty \mathbf{W}^+ f dG \left(\int_0^\infty f^2 dG\right)^{-1}, \quad n \geq 1, \ y \geq 0. \tag{14}$$

To determine the approximating r.v. for K_2^s we shall first obtain the covariance matrix function for this $\mathbf{Z}_n$, the computation of which is made easy by rewriting $\mathbf{Z}_n$ as follows.

Recall the definition of ψ from (5.6a.2) and define

$$\alpha_i(y) := I(e_i \leq y) + I(e_i \leq -y) - 1, \ y \in \mathbb{R}, \qquad \bar{\alpha}_i := \int_0^\infty \alpha_i \, d\psi, \ 1 \leq i \leq n;$$

$$\boldsymbol{\alpha}' := (\alpha_1, \ldots, \alpha_n); \qquad \bar{\boldsymbol{\alpha}}' := (\bar{\alpha}_1, \ldots., \bar{\alpha}_n); \qquad a := \int_0^\infty f^2 dG.$$

Then

$$\mathbf{Z}_n(y) = \mathbf{A}\mathbf{X}'[\boldsymbol{\alpha}(y) - f(y)\bar{\boldsymbol{\alpha}}\, a^{-1}], \qquad y \geq 0. \tag{15}$$

Now observe that under H_s, $E\boldsymbol{\alpha} = \mathbf{0}$, $E\alpha_1(x)\alpha_1(y) = 2(1-F(y))$, $0 \leq x \leq y$, and, because of the independence of the errors,

$$E\boldsymbol{\alpha}(x)\boldsymbol{\alpha}'(y) = 2(1-F(y))\, \mathbf{I}_{p \times p}, \qquad 0 \leq x \leq y. \tag{16}$$

Again, because of the symmetry and the continuity of F and Fubini, for $y \geq 0$,

$$\begin{aligned}
E\alpha_1(y)\bar{\alpha}_1 &= \int_0^\infty E[I(e_1 \leq y)+I(e_1 \leq -y)-1][I(e_1 \leq x)+I(e_1 \leq -x)-1]\, d\psi(x) \\
&= \int_0^\infty [F(x\wedge y)+F(-x\wedge y)-F(y)+F(x\wedge -y)+F(-x\wedge -y)-F(-y)]\, d\psi(x) \\
&= 2(1-F(y))\{\psi(y) - \psi(0)\} + \int_y^\infty 2(1-F(x))\, d\psi(x) \\
&= 2\int_y^\infty [\psi(x) - \psi(0)]\, dF(x) =: k(y), \quad \text{say.}
\end{aligned}$$

The last equality is obtained by integrating the second expression in the previous one by parts. From this and the independence of the errors, we obtain

$$E\boldsymbol{\alpha}(y)\bar{\boldsymbol{\alpha}}' = k(y)\, \mathbf{I}_{p \times p}, \qquad y \geq 0.$$

Similarly,

$$E\bar{\alpha}\bar{\alpha}' = I_{pxp}\, 4\int_0^\infty \int_x^\infty (1-F(y))\, d\psi(x)d\psi(y) =: I_{pxp}\, r(F,G), \text{ say.}$$

From these calculations one readily obtains that under H_s, for $0 \le x \le y$,

$$(17) \qquad K_n(x, y) := EZ_n(x)Z_n'(y)$$
$$= [2(1-F(y)) - k(y)f(x)a^{-1} - k(x)f(y)a^{-1} + r(F,G)]I_{pxp}.$$

We also need the weak convergence of W^+ to a continuous Gaussian process in uniform topology. One way to prove this is as follows. By (16),

$$(18) \qquad EW^+(x)W^+(y)' = 2(1-F(y))\, I_{pxp}, \qquad 0 \le x \le y,$$

From the definition (6.1.9) and the symmetry of F,

$$(19) \quad W^+(y) = \Sigma_i\, Ax_{ni}\{I(e_{ni} \le y) - I(-e_{ni} < y)\}$$
$$= \Sigma_i\, Ax_{ni}\{I(e_{ni} \le y) - F(y)\} - \Sigma_i\, Ax_{ni}\{I(-e_{ni} \le y) - F(y)\}$$
$$+ \Sigma_i\, Ax_{ni}I(-e_{ni} = y)$$

$$(20) \qquad = \mathcal{W}_1(y) - \mathcal{W}_2(y) + \Sigma_i\, Ax_{ni}I(-e_{ni} = y)\ , \quad \text{say,} \qquad y \ge 0.$$

Now, let $\mathcal{W}' := (\mathcal{W}_1, \mathcal{W}_2, ..., \mathcal{W}_p)$ be a vector of independent Wiener processes on [0, 1] such that $\mathcal{W}(0) = 0$, $E\mathcal{W} \equiv 0$, and $E\mathcal{W}_j(s)\mathcal{W}_j(t) = s\wedge t$, $1 \le j \le p$. Note that

$$E\mathcal{W}(2(1-F(x)))\mathcal{W}(2(1-F(y)))' = 2(1-F(y))\, I_{pxp}, \qquad 0 \le x \le y.$$

From (18) and (19), it hence follows, with the aid of the L–F CLT and the Cramer–Wold device, that under (NX), all finite dimensional distributions of W^+ converge to those of $\mathcal{W}(2(1-F))$.

To prove the tightness in the uniform metric, proceed as follows. From (20) and the triangle inequality, because of (NX), it suffices to show that $\mathcal{W}_1$ and $\mathcal{W}_2$ are tight. But by the symmetry and the continuity of F,

$$\{\mathcal{W}_1(y), y\in\mathbb{R}\} \underset{d}{=} \{\mathcal{W}_2(y), y\in\mathbb{R}\} \underset{d}{=} \{\mathcal{W}_1(F^{-1}(t)), 0\le t\le 1\}.$$

But, $\mathcal{W}_1(F^{-1})$ is obviously a p–vector of w.e.p.'s of the type W_d^* specified at (2.2a.33). Thus the tightness follows from (2.2a.35) of Corollary 2.2a.1. We summarize this weak convergence result as

Lemma 6.5.1. *Let* F *be a continuous d.f. that is symmetric around* 0 *and* $\{e_{ni}, 1\le i\le n\}$ *be i.i.d.* F *r.v.'s. Assume that* (NX) *holds. Then,*

$$\mathbf{W}^{+}(\cdot) \Rightarrow \mathcal{W}(2(1-F(\cdot))) \text{ in } (\mathbb{D}[0, \infty], \mathcal{d}). \qquad \square$$

The above discussion suggests the approximating process for the $\mathbf{Z}_n$ of (16) to be

$$\text{(21)} \qquad \mathbf{Z}(y) := \mathcal{W}(2(1-F(y)))-f(y)\int_0^\infty \mathcal{W}(2(1-F))fdG \left(\int_0^\infty f^2dG\right)^{-1}, \quad y\geq 0.$$

Straightforward calculations show that $K_n(x, y) = E\mathbf{Z}(x)\mathbf{Z}'(y)$, $0\leq x\leq y$, $n\geq 1$. This then verifies (i), (ii) and (iv) of Lemma 6.3.1 in the present case. Condition (iii) is verified as in the proof of Corollary 6.3.1(b) with the help of Lemma 6.5.1. To summarize, we have

Corollary 6.5.1. (a) *Under the conditions of Theorem* 6.5.3(a),

$$\text{(22)} \quad K_1^s \xrightarrow[d]{} 2\int_0^\infty [\mathcal{W}_1(2(1-F(y)))-f(y)\int_0^\infty \mathcal{W}_1(2(1-F))fdG \left(\int_0^\infty f^2dG\right)^{-1}]^2dG(y).$$

(b) *Under the conditions of Theorem* 6.5.3(b),

$$\text{(23)} \qquad K_2^s \xrightarrow[d]{} 2\int_0^\infty \|\mathbf{Z}\|^2dG(y), \quad \textit{with } \mathbf{Z} \textit{ given at } (21). \qquad \square$$

Remark 6.5.1. The distributions of the limiting r.v.'s in (22) and (23) have been studied by Martynov (1975, 1976) and Boos (1982) for some F and G. An interesting G in the present case is $G = \lambda$. But the corresponding tests are not a.d.f.. Also because the F in H_s is unknown, one can not use $G = F$ or the Anderson–Darling integrating measures $dG = dF/\{F(1-F)\}$ in these test statistics.

One way to overcome this problem would be to use the signed rank analogues of the above tests which is equivalent to replacing the F in the integrating measure by an appropriate empirical of the residuals $\{Y_{nj}-\mathbf{x}_{nj}'\mathbf{u};\ 1 \leq j \leq n\}$. Let R_{iu}^{+} denote the rank of $|Y_{ni}-\mathbf{x}_{ni}'\mathbf{u}|$ among $\{|Y_{nj}-\mathbf{x}_{nj}'\mathbf{u}|;\ 1 \leq j \leq n\}$, $1 \leq i \leq n$, and define

$$\mathcal{Z}_1^{+}(t, \mathbf{u}) := n^{-1/2}\, \Sigma_i\, I(R_{iu}^{+} \leq nt)\, \mathrm{sgn}(Y_{ni}-\mathbf{x}_{ni}'\mathbf{u}),$$

$$\mathcal{Z}_2^{+}(t, \mathbf{u}) := \mathbf{A}\, \Sigma_i\, \mathbf{x}_{ni}\, I(R_{iu}^{+} \leq nt)\, \mathrm{sgn}(Y_{ni}-\mathbf{x}_{ni}'\mathbf{u}), \qquad 0 \leq t \leq 1,\ \mathbf{u} \in \mathbb{R}^p.$$

The signed rank analogues of K_1^s, K_2^s statistics, respectively, are $\mathcal{K}_1^s := \inf\{\mathcal{K}_1(\mathbf{u});\ \mathbf{u}\in\mathbb{R}^p\}$, $\mathcal{K}_2^s := \inf\{\mathcal{K}_2(\mathbf{u});\ \mathbf{u}\in\mathbb{R}^p\}$, where

$$\mathcal{K}_1(\mathbf{u}) := \int_0^1 [\mathcal{Z}_1^{+}(t, \mathbf{u})]^2dL(t), \qquad \mathcal{K}_2(\mathbf{u}) := \int_0^1 \|\mathcal{Z}_2^{+}(t, \mathbf{u})\|^2dL(t), \qquad \mathbf{u} \in \mathbb{R}^p,$$

with $L \in \mathbb{D}\mathcal{I}[0,1]$. If $L(t) \equiv t$ then $\mathcal{K}_j^s$, $j=1, 2$, are analogues of the Cramer–Von Mises statistics. If L is specified by the relation $dL(t) = \{1/t(1-t)\}dt$, then the corresponding tests would be the Anderson–Darling type test of symmetry.

Note that if in (3.3.1) we put $d_{ni} \equiv n^{-1/2}$, $X_{ni} \equiv e_{ni}$, $F_{ni} \equiv F$, then Z_d^+ of (3.3.1) reduces to $\mathcal{Z}_1^+$. Similarly, $\mathcal{Z}_2^+$ corresponds to a p-vector of Z_d^+-processes of (3.3.1) whose j^{th} component has $d_{ni} \equiv (j^{th}$ column of $\mathbf{A})'\mathbf{x}_{ni}$ and the rest of the entities the same as above. Consequently, from (3.3.17) and arguments like those used for Theorem 6.5.3, we can deduce the following

Theorem 6.5.4. *Assume that* (1.1.1), H_s *and* (**NX**) *hold;* L *is a d.f. on* [0, 1], *and* F *of* H_s *satisfies* (**F1**), (**F2**).

(a) *If, in addition,* (5.6a.10) *and* (5.6a.11) *hold, then*

$$\mathcal{K}_1^s \xrightarrow[d]{} \int_0^1 [\mathcal{W}_1(t) - q^+(t)\int_0^1 \mathcal{W}_1\, q^+ dL\, (\int_0^1 (q^+)^2 dL)^{-1}]^2 dL(t). \tag{24}$$

(b) *Under no additional assumptions,*

$$\mathcal{K}_2^s \xrightarrow[d]{} \int_0^1 \|\mathcal{W}(t) - q^*(t)\int_0^1 \mathcal{W}\, q^* dL\, (\int_0^1 (q^*)^2 dL)^{-1}\|^2 dL(t), \tag{25}$$

where $q^*(t) := 2[f(F^{-1}((t+1)/2) - f(0)]$, $0 \le t \le 1$. □

Clearly this theorem covers $L(t) \equiv t$ case but not the case where $dL(t) = \{1/t(1-t)\}dt$. The *problem of proving an analogue of the above theorem for a general* L is *unsolved* at the time of this writing. □□