Journal of Multivariate Analysis **77**, 187–228 (2001)
doi:10.1006/jmva.2000.1926, available online at http://www.idealibrary.com on IDEAL®

Second Order Hadamard Differentiability in Statistical Applications

Jian-Jian Ren[1]

Tulane University

and

Pranab Kumar Sen

University of North Carolina at Chapel Hill

Received November 17, 1998; published online February 5, 2001

A formulation of the second-order Hadamard differentiability of (extended) statistical functionals and some related theoretical results are established. These results are applied to derive the limiting distributions of a class of generalized Cramér–von Mises type test statistics, which include some proposed new ones for the tests of goodness of fit in the 3-sample problems, the tests in linear regression models, and the tests of bivariate independence, as special cases. © 2001 Academic Press

AMS 1991 subject classifications: 62E20; 62G10; 62G99; 62J05.

Key words and phrases: Cramér–von Mises statistic; goodness of fit test; Hadamard differentiability; limiting distribution; linear models; M-estimator; statistical functionals; uniform asymptotic linearity; weighted empirical processes.

1. INTRODUCTION

Let $X_1, ..., X_n$ be independent random variables each having distribution function (d.f.) $F_1, ..., F_n$, respectively, and let $c_1, ..., c_n$ be a sequence of real numbers. Consider a weighted empirical process

$$S_n^*(x) = C_n^{-1} \sum_{i=1}^{n} c_{ni} I\{X_i \leqslant x\},$$

where $c_{ni} = (c_i - \bar{c}_n)/C_n$, $\bar{c}_n = n^{-1} \sum_{i=1}^n c_i$, and $C_n^2 = \sum_{i=1}^n (c_i - \bar{c}_n)^2$, and consider

$$S_n(x) = (n+1)^{-1} \sum_{i=1}^{n} I\{X_i \leqslant x\}.$$

[1] The author's research was partially supported by NSF grants DMS-9510367 and DMS-9626532.

0047-259X/01 $35.00
Copyright © 2001 by Academic Press
All rights of reproduction in any form reserved.

Motivated by the pioneering work of Hájek (1963), we are interested in the following functional of (S_n^*, S_n),

$$\Psi(S_n^*, S_n) = \int_{-\infty}^{\infty} [S_n^*(x)]^2 \, \psi(S_n(x)) \, dS_n(x), \tag{1.1}$$

where ψ is a non-negative weight function. The following example shows one application of this functional in statistics problems.

EXAMPLE 1 (Tests in Simple Linear Regression Model). Let X_i's be the observations of the following simple linear regression model,

$$X_i = \alpha + c_i \beta + \varepsilon_i, \qquad i = 1, ..., n, \tag{1.2}$$

where c_i are known regression constants, (α, β) is the vector of unknown (regression) parameters to be estimated, and ε_i are independent and identically distributed random variables (i.i.d.r.v.). For testing the null hypothesis

$$H_0; \beta = 0 \qquad \text{vs} \qquad H_1: \beta \neq 0, \tag{1.3}$$

we have that with $\psi \equiv 1$, $C_n^2 \Psi(S_n^*, S_n)$ is equivalent to the Cramér–von Mises type test statistic (Hájek and Šidák, 1967, p. 103). We refer to Hájek and Šidák (1967) for some general motivation and a useful survey of the related asymptotic theory. For a general weight function ψ, $\Psi(S_n^*, S_n)$ is termed as *generalized Cramér–von Mises* (GCvM) *type statistic* and its asymptotic properties have not been studied in the literature. The motivation of such a use of the weight function in certain situations is given in Example 3 of Section 2, which shows that the use of $\psi(t) = [t(1-t)]^{-1}$ is of special importance.

Several additional examples on the applications of functional $\Psi(S_n^*, S_n)$ in statistical problems are presented in Section 2, where some new test statistics are proposed for the tests of goodness of fit in the 3-sample problems, the tests in linear regression models and the tests of bivariate independence.

For all these examples, asymptotic distribution theory of $\Psi(S_n^*, S_n)$ is of focal importance. Thus we intend to develop a general approach for such studies in this paper. Let F be a continuous d.f., and observe that Ψ induces a function defined on the space $D[0, 1] \times D[0, 1]$,

$$\tau(U_n^*, U_n) = \int_0^1 [U_n^*(t)]^2 \, \psi(U_n(t)) \, dU_n(t), \tag{1.4}$$

where $U_n^* = S_n^* \circ F^{-1}$, $U_n = S_n \circ F^{-1}$ and $D[0, 1]$ is the space of right continuous real valued functions with left hand limits endowed with the supremum norm $\|\cdot\|$. In Example 1, F is the common d.f. of X_i's under H_0. We will see later on that this functional τ is Hadamard differentiable at $(0, U)$, where U is the uniform d.f. on $[0, 1]$. Hence, we have a form of the Taylor expansion,

$$\tau(U_n^*, U_n) = \tau(0, U) + \tau'_{(0, U)}(U_n^*, U_n - U) + \text{Rem}(U_n^*, U_n - U; \tau), \tag{1.5}$$

where $\tau'_{(0, U)}$ is a linear functional and is the Hadamard derivative at $(0, U)$, and $\text{Rem}(U_n^*, U_n - U; \tau)$ is the remainder term of this first-order expansion. Usually, we would expect to derive the asymptotic normality of $\tau(U_n^*, U_n)$ through verifying

$$C_n \,\text{Rem}(U_n^*, U_n - U; \tau) \xrightarrow{P} 0, \qquad \text{as} \quad n \to \infty. \tag{1.6}$$

References on this approach can be found in Ren (1994) and Ren and Sen (1995) or a simpler case in Fernholz (1983). However, we have $\tau'_{(0, U)} \equiv 0$ in (1.5) for our examples. Hence, the first-order expansion (1.5) cannot help us obtain the limiting distribution of $\tau(U_n^*, U_n) = \Psi(S_n^*, S_n)$. As noted in von Mises (1947), this leads us to consider a higher order expansion of (1.5), i.e.,

$$\begin{aligned}\tau(U_n^*, U_n) = \tau(0, U) + \tau'_{(0, U)}(U_n^*, U_n - U) + \tfrac{1}{2}\tau''_{(0, U)}(U_n^*, U_n - U)\\ + \text{Rem}_2(U_n^*, U_n - U; \tau),\end{aligned} \tag{1.7}$$

where $\tau''_{(0, U)}$ is the second-order Hadamard derivative at $(0, U)$ and $\text{Rem}_2(U_n^*, U_n - U; \tau)$ is the remainder term of this second-order expansion.

It appears that this second-order expansion (1.7) also provides some deeper asymptotic results on regression M-estimators in linear models. Such an application is discussed in Example 5 of Section 2. Moreover, the bivariate version of (1.7), i.e., for bivariate random vectors $\mathbf{X}_i = (V_i, W_i)$ (see Ren and Sen, 1995, for the first-order expansion with multivariate random vectors), gives a convenient tool to study the limiting distributions of the test statistics for the tests of independence of V_i and W_i, which is briefly discussed in Example 6 of Sections 2 and 5.

Motivated by the studies described above and by broader potential applications in other studies of the asymptotic properties of the statistics in different model settings (for instance, the L- and R-estimators in linear model, etc.), the concept of second-order Hadamard differentiability (SOHD) and some related theoretical results are established in Section 3 with proofs deferred to Section 6. While there might be methods other than the SOHD method to study the asymptotic distributions of the statistics considered in this paper, the general motivation of our work here is that the SOHD property possessed by a statistic provides more information than the first-order Hadamard differentiability property about the statistic's asymptotic behavior. Thereby, deeper asymptotic properties of the statistics can be obtained more easily. In some situations (e.g., Examples 2–4 and 6 of Section 2), we may hope that the testing or estimating procedure can be constructed and studied conveniently by the SOHD method.

The applications of the concept of SOHD are considered in Section 4 in deriving the limiting distributions of $\tau(U_n^*, U_n)$ given by (1.4) with proofs deferred to Section 6. These results are used in Section 5 to study the specific limiting distributions of the test statistics given in Examples 1–4 and 6 of Sections 1 and 2. The use of the special weight function $\psi(S_n) = [S_n(1-S_n)]^{-1}$ is included as a special case in our studies.

2. EXAMPLES

In the following examples, the GCvM type statistics $\Psi(S_n^*, S_n)$ are used as test statistics, some of which are new or have not been studied in the literature.

EXAMPLE 2 (Tests in General Linear Regression Model). We generalize the Cramér–von Mises type test statistic by Hájek and Šidák (1967) to a general linear model. Let X_i's be the observations of the following linear regression model,

$$X_i = \alpha + c_i\beta + d_i\gamma + \varepsilon_i, \qquad i = 1, ..., n, \tag{2.1}$$

where c_i, d_i are known regression constants, (α, β, γ) is the vector of unknown (regression) parameters to be estimated, and ε_i are i.i.d.r.v.'s. Let $d_{ni} = (d_i - \bar{d}_n)/D_n$, $\bar{d}_n = n^{-1}\sum_{i=1}^n d_i$, and $D_n^2 = \sum_{i=1}^n (d_i - \bar{d}_n)^2$, and let

$$T_n^*(x) = D_n^{-1} \sum_{i=1}^n d_{ni} I\{X_i \leqslant x\}. \tag{2.2}$$

For testing the null hypothesis

$$H_0: (\beta, \gamma) = \underset{\sim}{0} \qquad \text{vs} \qquad H_i: (\beta, \gamma) \neq \underset{\sim}{0}, \tag{2.3}$$

proceeding as in Hájek and Šidák (1967), we propose the following functionals of (S_n^*, T_n^*, S_n):

$$\int_{-\infty}^{\infty} (C_n S_n^*(x), D_n T_n^*(x))(C_n S_n^*(x), D_n T_n^*(x))^{\mathrm{T}} \, \psi(S_n(x)) \, dS_n(x)$$
$$= C_n^2 \Psi(S_n^*, S_n) + D_n^2 \Psi(T_n^*, S_n) \tag{2.4}$$

or

$$\max\left\{C_n^2\int_{-\infty}^{\infty}[S_n^*(x)]^2\,\psi(S_n(x))\,dS_n(x),\ D_n^2\int_{-\infty}^{\infty}[T_n^*(x)]^2\,\psi(S_n(x))\,dS_n(x)\right\}$$

$$=\max\{C_n^2\Psi(S_n^*,S_n),\ D_n^2\Psi(T_n^*,S_n)\} \tag{2.5}$$

to be used as GCvM type test statistics. One may note that through rank statistics $C_n^2\Psi(S_n^*, S_n)$ and $D_n^2\Psi(T_n^*, S_n)$ measure the effects that c_i's and d_i's have on X_i's, respectively. Clearly, these statistics can be easily further extended to the linear regression model with p unknown regression parameters, which will be briefly discussed in Section 5.

Example 3 (Goodness of Fit Tests in 3-Sample Problem). The test statistic given in (2.4) may be used for the 3-sample problem.

(a) *Case with equal sample size.* For $n=3m$, let X_i, $i=1, ..., m$, have d.f. F, X_i, $i=m+1, ..., 2m$, have d.f. G, and X_i, $i=2m+1, ..., n$, have d.f. H. We are interested in the following 3-sample problem

$$H_0\colon F=G=H \qquad \text{vs} \qquad H_1\colon H_0 \text{ not true.} \tag{2.6}$$

Suppose that

$$\mathfrak{Q}=\begin{bmatrix}\sqrt{\frac{1}{3}} & \sqrt{\frac{1}{3}} & \sqrt{\frac{1}{3}}\\ a_1 & a_2 & a_3\\ b_1 & b_2 & b_3\end{bmatrix}$$

is an orthogonal matrix. In model (2.1), if we let

$$(c_i, b_i)=\begin{cases}(a_1, b_1), & i=1, ..., m\\ (a_2, b_2), & i=m+1, ..., 2m\\ (a_3, b_3), & i=2m+1, ..., n,\end{cases} \tag{2.7}$$

then model (2.1) becomes

$$X_i=\begin{cases}\alpha+a_1\beta+b_1\gamma+\varepsilon_i, & i=1, ..., m\\ \alpha+a_2\beta+b_2\gamma+\varepsilon_i, & i=m+1, ..., 2m\\ \alpha+a_3\beta+b_3\gamma+\varepsilon_i, & i=2m+1, ..., n\end{cases}$$

and H_0 of the 3-sample problem (2.6) implies H_0: $a_1\beta+b_1\gamma=a_2\beta+b_2\gamma=a_3\beta+b_3\gamma$, which is equivalent to test (2.3) because of $a_1^2+a_2^2+a_3^2=b_1^2+b_2^2+b_3^2=1$ and $a_1+a_2+a_3=b_1+b_2+b_3=a_1b_1+a_2b_2+a_3b_3=0$. From (2.7), we have $C_n^2=D_n^2=m$,

$$(c_{ni}, d_{ni})=\begin{cases}(a_1/\sqrt{m}, b_1/\sqrt{m}), & i=1, ..., m\\ (a_2/\sqrt{m}, b_2/\sqrt{m}), & i=m+1, ..., 2m\\ (a_3/\sqrt{m}, b_3/\sqrt{m}), & i=2m+1, ..., n\end{cases}$$

and

$$\begin{aligned}C_nS_n^*(x)&=a_1\sqrt{m}\,F_m(x)+a_2\sqrt{m}\,G_m(x)+a_3\sqrt{m}\,H_m(x)\\ D_nT_n^*(x)&=b_1\sqrt{m}\,F_m(x)+b_2\sqrt{m}\,G_m(x)+b_3\sqrt{m}\,H_m(x),\end{aligned}$$

where F_m, G_m and H_m are the empirical d.f.'s for $(X_1, ..., X_m)$, $(X_{m+1}, ..., X_{2m})$ and $(X_{2m+1}, ..., X_{3m})$, respectively. Thus, we have

$$\left(\frac{n+1}{\sqrt{n}}S_n, C_nS_n^*, D_nT_n^*\right)^{\mathrm{T}}=\mathfrak{Q}(\sqrt{m}\,F_m, \sqrt{m}\,G_m, \sqrt{m}\,H_m)^{\mathrm{T}}$$

and

$$\begin{aligned}C_n^2S_n^{*2}+D_n^2T_n^{*2}&=m(F_m^2+G_m^2+H_m^2)-n\left(\frac{n+1}{n}S_n\right)^2\\ &=m\{(F_m^2+G_m^2+H_m^2)-3[(F_m+G_m+H_m)/3]^2\}\\ &=m\left\{\left(F_m-\frac{n+1}{n}S_n\right)^2+\left(G_m-\frac{n+1}{n}S_n\right)^2\right.\\ &\quad\left.+\left(H_m-\frac{n+1}{n}S_n\right)^2\right\}.\end{aligned}$$

Hence, from (2.4), the test statistic for the 3-sample problem (2.6) is given by

$$\begin{aligned}&C_n^2\Psi(S_n^*, S_n)+D_n^2\Psi(T_n^*, S_n)\\ &\quad=m\int_{-\infty}^{\infty}\left\{\left(F_m-\frac{n+1}{n}S_n\right)^2+\left(G_m-\frac{n+1}{n}S_n\right)^2\right.\\ &\qquad\left.+\left(H_m-\frac{n+1}{n}S_n\right)^2\right\}\psi(S_n)\,dS_n. \qquad (2.8)\end{aligned}$$

With $\psi\equiv 1$ and $\psi(t)=[t(1-t)]^{-1}$, the test statistics (2.8) are equivalent to those by Kiefer (1959) and Scholz and Stephens (1987), respectively. We note that under H_0, the asymptotic variance of $\sqrt{m}[F_m(x) - \frac{n+1}{n} S_n(x)]$ is given by $\frac{2}{3}F(x)(1-F(x))$. In the one-sample goodness of fit tests, Anderson and Darling (1952) used the reciprocal of the asymptotic variance of $\sqrt{m}[F_m(x) - F(x)]$ under the null hypothesis as the weight function to put equal weight (in a certain sense) to each point of the distribution. However, in the 3-sample goodness of fit test (2.6), F is not known even under H_0. To generalize this idea of using the weight function from one-sample tests to 3-sample tests, we may use the reciprocal of the estimator for the asymptotic variance of $\sqrt{m}[F_m(x) - \frac{n+1}{n} S_n(x)]$ as the weight function, i.e., we may use $\psi(S_n(x)) = [S_n(x)(1 - S_n(x))]^{-1}$ as a weight function. Note that for this special weight function, $\Psi(S_n^*, S_n)$ in (1.1) is well defined because of our choice of S_n. For a general weight function ψ, (2.8) gives GCvM type test statistic for the 3-sample goodness of fit tests.

(b) *Case with non-equal sample size.* Suppose that for $n = n_1 + n_2 + n_3$, the random samples $(X_1, ..., X_{n_1})$, $(X_{n_1+1}, ..., X_{n_1+n_2})$ and $(X_{n_1+n_2+1}, ..., X_{n_1+n_2+n_3})$, are drawn from F, G and H with the empirical d.f.'s F_{n_1}, G_{n_2} and H_{n_3}, respectively. Let

$$\mathfrak{Q}_n = \begin{bmatrix} \sqrt{\frac{n_1}{n}} & \sqrt{\frac{n_2}{n}} & \sqrt{\frac{n_3}{n}} \\ a_{n1} & a_{n2} & a_{n3} \\ b_{n1} & b_{n2} & b_{n3} \end{bmatrix}$$

be an orthogonal matrix, and let $C_n^2 = D_n^2 = n$ with

$$(c_{ni}, d_{ni}) = \begin{cases} (a_{n1}/\sqrt{n_1}, b_{n1}/\sqrt{n_1}), & i = 1, ..., n_1 \\ (a_{n2}/\sqrt{n_2}, b_{n2}/\sqrt{n_2}), & i = n_1 + 1, ..., n_1 + n_2 \\ (a_{n3}/\sqrt{n_3}, b_{n3}/\sqrt{n_3}), & i = n_1 + n_2 + 1, ..., n. \end{cases}$$

Then, we have

$$\left(\frac{n+1}{\sqrt{n}} S_n, C_n S_n^*, D_n T_n^*\right)^{\mathrm{T}} = \mathfrak{Q}_n(\sqrt{n_1}\, F_{n_1}, \sqrt{n_2}\, G_{n_2}, \sqrt{n_3}\, H_{n_3})^{\mathrm{T}},$$

$$C_n^2 S_n^{*2} + D_n^2 T_n^{*2} = n_1 \left(F_{n_1} - \frac{n+1}{n} S_n\right)^2 + n_2 \left(G_{n_2} - \frac{n+1}{n} S_n\right)^2 + n_3 \left(H_{n_3} - \frac{n+1}{n} S_n\right)^2,$$

and from (2.4), we have the following test statistic for (2.6) with non-equal sample size

$$C_n^2 \Psi(S_n^*, S_n) + D_n^2 \Psi(T_n^*, S_n)$$
$$= \int_{-\infty}^{\infty} \left\{ n_1 \left(F_{n_1} - \frac{n+1}{n} S_n \right)^2 + n_2 \left(G_{n_2} - \frac{n+1}{n} S_n \right)^2 + n_3 \left(H_{n_3} - \frac{n+1}{n} S_n \right)^2 \right\} \psi(S_n)\, dS_n. \tag{2.9}$$

EXAMPLE 4 (Alternative Tests in 3-Sample Problem). Functionals (2.4) and (2.5) can be used to construct alternative tests for the 3-sample problem (2.6). Consider a more general case of model (2.1),

$$X_i = \alpha + c_i \beta + d_i \gamma + e_i \eta + \varepsilon_i, \qquad i = 1, ..., n, \tag{2.10}$$

and denote

$$R_n^*(x) = E_n^{-1} \sum_{i=1}^{n} e_{ni} I\{X_i \leqslant x\},$$

where $\varepsilon_{ni} = (e_i - \bar{e}_n)/E_n$, $\bar{e}_n = n^{-1} \sum_{i=1}^n e_i$, and $E_n^2 = \sum_{i=1}^n (e_i - \bar{e}_n)^2$. Several alternative test statistics for (2.6) are given as follows.

(a) In (2.10), let

$$(c_i, d_i, e_i) = \begin{cases} (1, 0, 0), & i = 1, ..., m \\ (0, 1, 0), & i = m+1, ..., 2m \\ (0, 0, 1), & i = 2m+1, ..., n; \end{cases} \tag{2.11}$$

then model (2.10) becomes a special case of the 3-sample problem with equal sample size, and H_0 of the goodness of fit test (2.6) implies H_0: $\beta = \gamma = \eta$. Since under this H_0, X_i's are i.i.d., from Hájek and Šidák (1967, see discussion on p. 90), we know that the test statistics (2.4) or (2.5) may be used here. For (2.11), we have

$$C_n^2 = D_n^2 = E_n^2 = \frac{2}{3} m$$

$$(C_n S_n^*, D_n T_n^*, E_n R_n^*)$$
$$= \sqrt{3m/2} \left(\left(F_m - \frac{n+1}{n} S_n \right), \left(G_m - \frac{n+1}{n} S_n \right), \left(H_m - \frac{n+1}{n} S_n \right) \right),$$

and from (2.5), we have the following test statistic for (2.6):

$$\max\{C_n^2\Psi(S_n^*, S_n), D_n^2\Psi(T_n^*, S_n), E_n^2\Psi(R_n^*, S_n)\}$$
$$=\frac{3}{2}\max\left\{m\int_{-\infty}^{\infty}\left(F_m-\frac{n+1}{n}S_n\right)^2\psi(S_n)\,dS_n,\right.$$
$$m\int_{-\infty}^{\infty}\left(G_m-\frac{n+1}{n}S_n\right)^2\psi(S_n)\,dS_n,$$
$$\left.m\int_{-\infty}^{\infty}\left(H_m-\frac{n+1}{n}S_n\right)^2\psi(S_n)\,dS_n\right\}. \tag{2.12}$$

For a general weight function ψ, this statistic has not been studied; for the case of $\psi\equiv 1$, it is equivalent to that given by Kiefer (1959), where the limiting distribution of this test statistic was not derived.

(b) For the case of the 3-sample problem with non-equal sample size, let in (2.10)

$$(c_{ni}, d_{ni}, e_{ni})=\begin{cases} C_n^{-1}\left(1-\frac{n_1}{n}, -\frac{n_2}{n}, -\frac{n_3}{n}\right), & i=1, ..., n_1\\ D_n^{-1}\left(-\frac{n_1}{n}, 1-\frac{n_2}{n}, -\frac{n_3}{n}\right), & i=n_1+1, ..., n_1+n_2\\ E_n^{-1}\left(-\frac{n_1}{n}, -\frac{n_2}{n}, 1-\frac{n_3}{n}\right) & i=n_1+n_2+1, ..., n \end{cases} \tag{2.13}$$

and

$$C_n^2=n_1\left(1-\frac{n_1}{n}\right), \qquad D_n^2=n_2\left(1-\frac{n_2}{n}\right), \qquad E_n^2=n_3\left(1-\frac{n_3}{n}\right).$$

Then, we have

$$C_nS_n^*=\sqrt{\frac{n_1n}{n-n_1}}\left(F_{n_1}-\frac{n+1}{n}S_n\right)$$
$$D_nT_n^*=\sqrt{\frac{n_2n}{n-n_2}}\left(G_{n_2}-\frac{n+1}{n}S_n\right),$$
$$E_nR_n^*=\sqrt{\frac{n_3n}{n-n_3}}\left(H_{n_3}-\frac{n+1}{n}S_n\right),$$

and from (2.5), we have the following test statistic for the 3-sample problem (2.6) with non-equal sample size

$$\max\{C_n^2\Psi(S_n^*, S_n), D_n^2\Psi(T_n^*, S_n), E_n^2\Psi(R_n^*, S_n)\}$$
$$=\max\left\{\frac{n_1 n}{n-n_1}\int_{-\infty}^{\infty}\left(F_{n_1}-\frac{n+1}{n}S_n\right)^2\psi(S_n)\,dS_n,\right.$$
$$\frac{n_2 n}{n-n_2}\int_{-\infty}^{\infty}\left(G_{n_2}-\frac{n+1}{n}S_n\right)^2\psi(S_n)\,dS_n,$$
$$\left.\frac{n_3 n}{n-n_3}\int_{-\infty}^{\infty}\left(H_{n_3}-\frac{n+1}{n}S_n\right)^2\psi(S_n)\,dS_n\right\}. \tag{2.14}$$

(c) In (2.10), let

$$(c_i, d_i, e_i)=\begin{cases}(1, 0, -1), & i=1, ..., m\\ (-1, 1, 0), & i=m+1, ..., 2m\\ (0, -1, 1) & i=2m+1, ..., n;\end{cases} \tag{2.15}$$

then

$$C_n^2=D_n^2=E_n^2=2m$$
$$(C_nS_n^*, D_nT_n^*, E_nR_n^*)=\sqrt{m/2}\,((F_m-G_m), (G_m-H_m), (H_m-F_m)),$$

model (2.10) becomes a special case of the 3-sample problem with equal sample size, and H_0 of the goodness of fit test (2.6) implies H_0: $\beta=\gamma=\eta$. From (2.5) and (2.4), we have the following test statistics for (2.6)

$$\max\{C_n^2\Psi(S_n^*, S_n), D_n^2\Psi(T_n^*, S_n), E_n^2\Psi(R_n^*, S_n)\}$$
$$=\tfrac{1}{2}\max\left\{m\int_{-\infty}^{\infty}(F_m-G_m)^2\,\psi(S_n)\,dS_n,\right.$$
$$m\int_{-\infty}^{\infty}(G_m-H_m)^2\,\psi(S_n)\,dS_n,$$
$$\left.m\int_{-\infty}^{\infty}(H_m-F_m)^2\,\psi(S_n)\,dS_n\right\}, \tag{2.16}$$

and

$$C_n^2\Psi(S_n^*, S_n)+D_n^2\Psi(T_n^*, S_n)+E_n^2\Psi(R_n^*, S_n)$$
$$=\frac{m}{2}\int_{-\infty}^{\infty}\{(F_m-G_m)^2+(G_m-H_m)^2+(H_m-F_m)^2\}\,\psi(S_n)\,dS_n, \tag{2.16a}$$

respectively. For a general weight function ψ, these statistics have not been studied; for the case of $\psi \equiv 1$, David (1958) constructed and studied a test statistic which is the Kolmogorov–Smirnov version of (2.16).

(d) For the case of the 3-sample problem with non-equal sample size, let in (2.10)

$$(c_{ni}, d_{ni}, e_{ni}) = \begin{cases} C_n^{-1}(\sqrt{n_2/n_1}, 0, -\sqrt{n_3/n_1}), & i=1, ..., n_1 \\ D_n^{-1}(-\sqrt{n_1/n_2}, \sqrt{n_3/n_2}, 0), & i=n_1+1, ..., n_1+n_2 \\ E_n^{-1}(0, -\sqrt{n_2/n_3}, \sqrt{n_1/n_3}), & i=n_1+n_2+1, ..., n, \end{cases} \tag{2.17}$$

and

$$C_n^2 = n_1 + n_2, \qquad D_n^2 = n_2 + n_3, \qquad E_n^2 = n_1 + n_3;$$

then we have

$$C_n S_n^* = \sqrt{\frac{n_1 n_2}{n_1 + n_2}} (F_{n_1} - G_{n_2}),$$

$$D_n T_n^* = \sqrt{\frac{n_2 n_3}{n_2 + n_3}} (G_{n_2} - H_{n_3}), \qquad E_n R_n^* = \sqrt{\frac{n_1 n_3}{n_1 + n_3}} (H_{n_3} - F_{n_1}),$$

and from (2.5) and (2.4), we have the following test statistics for (2.6) with non-equal sample size

$$\begin{aligned} &\max\{C_n^2 \Psi(S_n^*, S_n), D_n^2 \Psi(T_n^*, S_n), E_n^2 \Psi(R_n^*, S_n)\} \\ &\quad = \max \left\{ \frac{n_1 n_2}{n_1 + n_2} \int_{-\infty}^{\infty} (F_{n_1} - G_{n_2})^2 \psi(S_n)\, dS_n, \right. \\ &\quad \frac{n_2 n_3}{n_2 + n_3} \int_{-\infty}^{\infty} (G_{n_2} - H_{n_3})^2 \psi(S_n)\, dS_n, \\ &\quad \left. \frac{n_1 n_3}{n_1 + n_3} \int_{-\infty}^{\infty} (H_{n_3} - F_{n_1})^2 \psi(S_n)\, dS_n \right\}, \end{aligned} \tag{2.18}$$

and

$$\begin{aligned} &C_n^2 \Psi(S_n^*, S_n) + D_n^2 \Psi(T_n^*, S_n) + E_n^2 \Psi(R_n^*, S_n) \\ &\quad = \int_{-\infty}^{\infty} \left\{ \frac{n_1 n_2}{n_1 + n_2} (F_{n_1} - G_{n_2})^2 + \frac{n_2 n_3}{n_2 + n_3} (G_{n_2} - H_{n_3})^2 \right. \\ &\quad \left. + \frac{n_1 n_3}{n_1 + n_3} (H_{n_3} - F_{n_1})^2 \right\} \psi(S_n)\, dS_n, \end{aligned} \tag{2.18a}$$

respectively.

Clearly, our test statistics in Examples 3–4 here can be easily extended to treat the k-sample problem ($k \geqslant 2$).

The next example describes the application of the second-order expansion (1.7) for statistical functionals to the studies of deeper asymptotic results on regression M-estimators in linear models.

EXAMPLE 5 (Regression M-estimators). For simplicity of presentation, we consider the simple linear regression model given by (1.2) with error distribution F. The robust M-estimator of (α, β), denoted by $(\hat{\alpha}_n, \hat{\beta}_n)$, is given by a solution (with respect to (θ_1, θ_2)) of the estimating equations

$$\begin{aligned} &\sum_{i=1}^{n} \psi(X_i - \theta_1 - c_i\theta_2) = 0 \\ &\sum_{i=1}^{n} c_i\psi(X_i - \theta_1 - c_i\theta_2) = 0, \end{aligned} \tag{2.19}$$

where $\psi: \mathbb{R} \to \mathbb{R}$ is a suitable score function. Setting $Y_i = X_i - \alpha - c_i\beta$ (i.i.d.r.v.'s with d.f. F), $1 \leqslant i \leqslant n$, with $\sum_{i=1}^{n} c_i = 0$, and $\mathbf{u} = (u_1, u_2)^{\mathrm{T}} \in \mathbb{R}^2$ for $u_1 = \sqrt{n}(\theta_1 - \alpha)$, $u_2 = C_n(\theta_2 - \beta)$, we have that (2.19) is equivalent to

$$\begin{aligned} M_{1n}(\mathbf{u}) &= \sum_{i=1}^{n} n^{-1/2}\psi(Y_i - \mathbf{c}_{ni}^{\mathrm{T}}\mathbf{u}) = 0 \\ M_{2n}(\mathbf{u}) &= \sum_{i=1}^{n} c_{ni}\psi(Y_i - \mathbf{c}_{ni}^{\mathrm{T}}\mathbf{u}) = 0, \end{aligned} \tag{2.20}$$

where $\mathbf{c}_{ni} = (n^{-1/2}, c_{ni})^{\mathrm{T}}$. Letting $\mathbf{M}_n(\mathbf{u}) = (M_{1n}, M_{2n})^{\mathrm{T}}$, the asymptotic normality and related properties of $(\hat{\alpha}_n, \hat{\beta}_n)$ have been studied most conveniently by incorporating the following uniform asymptotic linearity for the M-estimators:

$$\sup_{|\mathbf{u}| \leqslant K} |\mathbf{M}_n(\mathbf{u}) - \mathbf{M}_n(\mathbf{0}) + \gamma \mathbf{Q}_n \mathbf{u}| \xrightarrow{P} 0, \qquad \text{as} \quad n \to \infty, \tag{2.21}$$

where K is any finite positive real number, $|\cdot|$ stands for supremum norm on $\mathbb{R}^2$, $\gamma = \int \psi' \, dF > 0$, and $\mathbf{Q}_n = \sum_{i=1}^{n} \mathbf{c}_{ni}\mathbf{c}_{ni}^{\mathrm{T}}$. Note that $\mathbf{M}_n(\mathbf{u})$ is a linear functional of the empirical functions

$$\begin{aligned} V_n(t, \mathbf{u}) &= \sum_{i=1}^{n} n^{-1/2} I\{Y_i \leqslant F^{-1}(t) + \mathbf{c}_{ni}^{\mathrm{T}}\mathbf{u}\} \\ V_n^*(t, \mathbf{u}) &= \sum_{i=1}^{n} c_{ni} I\{Y_i \leqslant F^{-1}(t) + \mathbf{c}_{ni}^{\mathrm{T}}\mathbf{u}\}, \end{aligned}$$

where $t \in [0, 1]$ and $\mathbf{u} \in \mathbb{R}^2$, viz.,

$$\mathbf{M}_n(\mathbf{u}) = \tau_L(V_n(\cdot, \mathbf{u}), V_n^*(\cdot, \mathbf{u}))$$
$$= \left(\int \psi \circ F^{-1} \, dV_n(\cdot, \mathbf{u}), \int \psi \circ F^{-1} \, dV_n^*(\circ, \mathbf{u}) \right)^{\mathrm{T}},$$

and that $\mathbf{M}_n(\mathbf{u})$ is a functional defined on $D[0, 1] \times D[0, 1]$, because $V_n(\cdot, \mathbf{u})$ and $V_n^*(\cdot, \mathbf{u})$ are elements on $D[0, 1]$ for any fixed $\mathbf{u} \in \mathbb{R}^2$. Since the Hadamard derivative is a linear functional, $\mathbf{M}_n(\mathbf{u})$ could be the Hadamard derivative of some appropriate functional τ. The choice of τ may not be unique, and general motivations for this are given in Ren and Sen (1991). Ren and Sen (1991) showed that if a functional τ is Hadamard differentiable with first-order derivative $\mathbf{M}_n(\mathbf{u})$, then as $n \to \infty$

$$\sup_{|\mathbf{u}| \leqslant K} \left| \begin{bmatrix} \sqrt{n} & 0 \\ 0 & a_n \end{bmatrix} \{\tau(V_n(\cdot \mathbf{u}), V_n^*(\cdot, \mathbf{u}))\} - [\mathbf{M}_n(\mathbf{u}) - \mathbf{M}_n(\mathbf{0})] \right| \xrightarrow{P} 0, \tag{2.22}$$

where $a_n = \sum_{i=1}^n c_{ni}^+ = \sum_{i=1}^n c_{ni}^-$ with $c_{ni}^+ = \max\{0, c_{ni}\}$ and $c_{ni}^- = -\min\{0, c_{ni}\}$, and for some functional τ_1, τ_2: $D[0, 1] \to \mathbb{R}$,

$$\tau(V_n(\cdot, \mathbf{u}), V_n^*(\cdot, \mathbf{u})) = (\tau_1(V_n(\cdot, \mathbf{u})/\sqrt{n}), \tau_2(V_n^*(\cdot, \mathbf{u})/a_n))^{\mathrm{T}}. \tag{2.23}$$

Thereby, using a convenient functional τ, Ren and Sen (1994) established (2.21) from showing

$$\sup_{|\mathbf{u}| \leqslant K} \left| \begin{bmatrix} \sqrt{n} & 0 \\ 0 & a_n \end{bmatrix} \{\tau(V_n(\cdot, \mathbf{u}), V_n^*(\cdot, \mathbf{u}))\} + \gamma \mathbf{Q}_n \mathbf{u} \right| \xrightarrow{P} 0, \qquad \text{as} \quad n \to \infty.$$

This (first-order) Hadamard differentiability approach for establishing (2.21) compares quite favorably with the alternative ones in the literature (Ren and Sen, 1994). Note that (2.22) is given by the asymptotic behavior of $\mathrm{Rem}(V_n(\cdot, \mathbf{u}) - U, V_n^*(\cdot, \mathbf{u}) - U; \tau)$. Hence, we naturally expect to obtain more detailed asymptotic properties of (2.21): the convergence rate in probability, using the second-order expansion (1.7) for the functional τ in (2.22) and $(V_n(\cdot, \mathbf{u}), V_n^*(\cdot, \mathbf{u}))$. In this context, some asymptotic results on $\mathrm{Rem}_2(V_n(\cdot, \mathbf{u}) - U, V_n^*(\cdot, \mathbf{u}) - U; \tau)$ will be given in Section 3. These results have been used to establish an asymptotic representation of $(\hat{\alpha}_n, \hat{\beta}_n)$ under weaker conditions than those available in the literature (Ren and Sen, 1993).

EXAMPLE 6 (Test of Independence). Let $(V_1, W_1), ..., (V_m, W_m)$ be a random sample from a bivariate d.f. $F(v, w)$. If one wishes to test if V_i and W_i are independent, the test statistic for the following hypothesis,

$$\begin{aligned} &H_0\colon F(v, w) = F(v, \infty)\, F(\infty, w) \qquad \text{vs} \\ &H_1\colon F(v, w) \neq F(v, \infty)\, F(\infty, w), \end{aligned} \tag{2.24}$$

may be given by the bivariate version of (1.1). To see this, let $n = m + m^2$ and for $1 \leqslant i \leqslant n$, denote

$$\mathbf{X}_i = \begin{cases} (V_i, W_i) & \text{if} \quad 1 \leqslant i \leqslant m \\ (V_{i-jm}, W_j) & \text{if} \quad jm+1 \leqslant i \leqslant (j+1)\, m, \quad j = 1, ..., m \end{cases} \tag{2.25}$$

$$\mathbf{x} = (v, w), \qquad I\{\mathbf{X}_i \leqslant \mathbf{x}\} = I\{V_i \leqslant v, W_i \leqslant w\}.$$

Thus, in (1.1) for $c_i = 2$, $1 \leqslant i \leqslant m$; $c_i = 1$, $m+1 \leqslant i \leqslant n$, we have $C_n^2 = m^2/(m+1)$,

$$\begin{aligned} S_n^*(\mathbf{x}) &= C_n^{-1} \sum_{i=1}^{n} c_{ni} I\{\mathbf{X}_i \leqslant \mathbf{x}\} = [F_m(v, w) - F_m(v, \infty)\, F_m(\infty, w)] \\ S_n(\mathbf{x}) &= \frac{n}{n+1} \left[\frac{1}{m+1} F_m(v, w) + \frac{m}{m+1} F_m(v, \infty)\, F_m(\infty, w) \right], \end{aligned} \tag{2.26}$$

where F_m is the bivariate empirical d.f. of $(V_1, W_1), ..., (V_m, W_m)$. Hence, the bivariate version of (1.1) gives a test statistic for (2.24):

$$\begin{aligned} C_n^2 \Psi(S_n^*, S_n) &= C_n^2 \iint [S_n^*(\mathbf{x})]^2\, \psi(S_n(\mathbf{x}))\, dS_n(\mathbf{x}) \\ &= \frac{m^2}{m+1} \iint [F_m(v, w) - F_m(v, \infty)\, F_m(\infty, w)]^2 \\ &\quad \times \psi(S_n(\mathbf{x}))\, dS_n(\mathbf{x}). \end{aligned} \tag{2.27}$$

We note that although the random vectors $\mathbf{X}_i$ in (2.25), $1 \leqslant i \leqslant n$, are not all independent from one and other, and $S_n^*(\mathbf{x})$, $S_n(\mathbf{x})$ are bivariate processes, a more general multivariate form of (1.5) for the first-order Hadamard derivative is studied by Ren and Sen (1995), and the multivariate analogue of (1.7) for the second order Hadamard derivative applies to the statistic given in (2.27), because under H_0, $\sqrt{m}\, S_n^*(\mathbf{x})$ weakly converges to a centered Gaussian process and $S_n(\mathbf{x})$ uniformly converges to $F(v, \infty)\, F(\infty, w)$ with probability 1. This will be briefly discussed in Section 5. Hoeffding (1948) studied the independence test (2.24) using

U-statistics (its limiting distribution under H_0 was not specifically given), while our formulation in this paper does not require U-statistics representation and directly connects the degenerated limiting distribution of the test statistic under H_0 with the second-order Hadamard derivative.

3. SECOND-ORDER HADAMARD DIFFERENTIABILITY

First-order Hadamard differentiability is usually defined as follows. Let V and W be the topological vector spaces, $\mathfrak{L}_1(V, W)$ be the set of continuous linear transformations from V to W, and $\mathscr{A}$ be an open set of V.

DEFINITION I. A functional $T: \mathscr{A} \to W$ is *Hadamard differentiable* (or *compact differentiable*) at $F \in \mathscr{A}$ if there exists $T'_F \in \mathfrak{L}_1(V, W)$ such that for any compact set Γ of V,

$$\lim_{t \to 0} \frac{T(F+tH) - T(F) - T'_F(tH)}{t} = 0 \tag{3.1}$$

uniformly for any $H \in \Gamma$. The linear function T'_F is called the *Hadamard derivative* of T at F.

For the sake of convenience, in (3.1) we usually denote

$$\mathrm{Rem}_1(tH) = T(F+tH) - T(F) - T'_F(tH) \tag{3.2}$$

as the remainder term of the first-order expansion. This definition is related to the original one given in Reeds (1976) (see Fernholz, 1983). We should note that in normed vector spaces, (3.1) is equivalent to the following form (viz., Gill, 1989)

$$\lim_{t \to 0} \frac{\mathrm{Rem}_1(F + tH_n)}{t} = 0, \tag{3.1a}$$

for any sequences H_n with $H_n \to H \in V$.

Let $\mathscr{C}(V, W)$ be the set of continuous transformations from V to W, and let

$$\mathfrak{L}_2(V, W) = \{f;\ f \in \mathscr{C}(V, W),\ f(tH) = t^2 f(H) \text{ for any } H \in V,\ t \in \mathbb{R}\}. \tag{3.3}$$

We define second-order Hadamard differentiability as follows.

DEFINITION II. A functional $T: \mathscr{A} \to W$ is *second-order Hadamard differentiable* at $F \in \mathscr{A}$ if there exists $T'_F \in \mathfrak{L}_1(V, W)$ and $T''_F \in \mathfrak{L}_2(V, W)$ such that for any compact set Γ of V

$$\lim_{t \to 0} \frac{T(F+tH) - T(F) - T'_F(tH) - \frac{1}{2} T''_F(tH)}{t^2} = 0 \tag{3.4}$$

uniformly for any $H \in \Gamma$. T'_F and T''_F are called the first- and second-order Hadamard derivatives of T at F, respectively.

We denote the remainder term of the second-order expansion as below:

$$\text{Rem}_2(tH) = T(F+tH) - T(F) - T'_F(tH) - \tfrac{1}{2} T''_F(tH). \tag{3.5}$$

In normed vector spaces, (3.4) may be presented in an equivalent form,

$$\lim_{t \to 0} \frac{\text{Rem}_2(F + tH_n)}{t^2} = 0, \tag{3.4a}$$

for any sequences H_n with $H_n \to H \in V$.

Remark 1. In the literature, various types of higher order derivatives have been considered by some authors, such as von Mises (1947), Averbukh and Smolyanov (1967), Keller (1974), Reeds (1976), Sen (1988), among others. Averbukh and Smolyanov (1967) defined the higher order derivative inductively; that is the second-order derivative is defined if the map $F \in \mathscr{A} \to T'_F \in \mathfrak{L}_1(V, W)$ is differentiable. In Keller (1974) and Reeds (1976), the second-order derivative is required to be a "2-linear map". One may note that all these definitions require the functional T to be at least continuously differentiable at F, while our definition of the second-order Hadamard derivative in Definition II does not require this, thus is a weaker differentiability condition. Examples show that in some situations, continuous differentiability condition fails to hold (see Gill, 1989). One may also note that based on our Definition II, the computation of the second-order Hadamard derivative is more straightforward. In (2.2) of Sen (1988), if we let

$$T'_F(H) = \int T_1(F; x)\, dH(x) \qquad \text{and}$$

$$T''_F(H) = \iint T_2(F; x, y)\, dH(x)\, dH(y),$$

then (2.2) of Sen (1988) coincides with our (3.5) for continuous and bounded T_1 and T_2. Later on one will see that our concept of second-order Hadamard differentiability in Definition II suffices in our studies here.

Remark 2. From our definition of second-order Hadamard differentiability, it is obvious that the existence of the second-order Hadamard derivative implies the existence of the first-order Hadamard derivative, and we have

$$T'_F(\delta_x - F) = \mathrm{IC}(x; F, T) = \frac{d}{dt} T(F + t(\delta_x - F))|_{t=0} \tag{3.6}$$

and

$$T''_F(\delta_x - F) = \frac{d^2}{dt^2} T(F + t(\delta_x - F))|_{t=0}, \tag{3.7}$$

where δ_x is the d.f. of the point mass one at x.

It is known that the chain rule holds for first-order Hadamard differentiability (Fernholz, 1983), which makes it useful. We will show that the chain rule also holds for second-order Hadamard differentiability. The proof is given in Section 6.

PROPOSITION 3.1. *Let V, W and Z be the topological vector spaces with T: $V \to W$ and Q: $W \to Z$. If T is second-order Hadamard differentiable at $F \in V$ and if Q is second-order Hadamard differentiable at $T(F) \in W$, then $\tau = Q \circ T$ is second-order Hadamard differentiable at F with derivatives*

$$\tau'_F = Q'_{T(F)} \circ T'_F \tag{3.8}$$

$$\tau''_F = (Q \circ T)''_F = Q''_{T(F)} \circ T'_F + Q'_{T(F)} \circ T''_F. \tag{3.9}$$

In our current study, we will be primarily interested in the functionals defined on Banach space $(D[0, 1] \times D[0, 1], \|\cdot\|)$, where $\|\cdot\|$ stands for the supremum norm and the σ-field generated by all open balls is equipped (see Shorack and Wellner, 1986, for references). In the next few propositions, we give some sufficient conditions for a second-order Hadamard differentiable functional defined on the space $D[0, 1]$ or $D[0, 1] \times D[0, 1]$ with the proofs deferred to Section 6.

The next proposition is a generalization of Proposition 6.1.2 of Fernholz (1983) for second-order Hadamard differentiability.

PROPOSITION 3.2. *Let L: $\mathbb{R} \to \mathbb{R}$ be differentiable and L' be continuous, bounded and piecewise differentiable with a bounded derivative. Let γ: $D[0, 1] \to L^p[0, 1]$, $p \geqslant 1$, be defined by*

$$\gamma(H) = L \circ H, \qquad H \in D[0, 1]$$

and let $\mathbb{A}$ be the set of points in $\mathbb{R}$ where L' is not differentiable. If γ is defined in a neighborhood of $Q \in D[0, 1]$ and if $\mu\{x; Q(x) \in \mathbb{A}\} = 0$, where μ is Lebesgue measure, then γ is second-order Hadamard differentiable at Q with derivatives

$$\gamma'_Q(H) = (L' \circ Q)\, H \qquad \text{and} \qquad \gamma''_Q(H) = (L'' \circ Q)\, H^2, \qquad H \in D[0, 1].$$

PROPOSITION 3.3. *Let $\gamma: D[0, 1] \times D[0, 1] \to D[0, 1]$ be defined by*

$$\gamma(G, H) = G\phi(H), \qquad G, H \in D[0, 1]$$

where ϕ is a real valued function with continuous second derivative ϕ''. Then for $G_0, H_0 \in D[0, 1]$, γ is second-order Hadamard differentiable at (G_0, H_0) with derivatives

$$\gamma'_{(G_0, H_0)}(G, H) = G_0 \phi'(H_0)\, H + \phi(H_0)\, G, \qquad G, H \in D[0, 1]$$

and

$$\gamma''_{(G_0, H_0)}(G, H) = 2\phi'(H_0)\, GH + G_0 \phi''(H_0)\, H^2, \qquad G, H \in D[0, 1].$$

We should notice that a special case of the above proposition with $G_0 \equiv 0$ requires weaker conditions on ϕ. We state this case as a corollary without proof.

COROLLARY 3.4. *Let $\gamma: D[0, 1] \times D[0, 1] \to D[0, 1]$ be defined by*

$$\gamma(G, H) = G\phi(H), \qquad G, H \in D[0, 1],$$

where ϕ is a real valued function with continuous derivative ϕ'. Then for $H_0 \in D[0, 1]$, γ is second-order Hadamard differentiable at $(0, H_0)$ with derivatives

$$\gamma'_{(0, H_0)}(G, H) = \phi(H_0)\, G \qquad \text{and}$$
$$\gamma''_{(0, H_0)}(G, H) = 2\phi'(H_0)\, GH, \qquad G, H \in D[0, 1].$$

The proof of the following proposition, given in Section 6, is similar to that of Lemma 3 by Gill (1989), where a class of elements in $D[0, 1] \times D[0, 1]$ is considered,

$$E = \left\{ (G, H);\, G, H \in D[0, 1] \text{ with } \int_0^1 |dH| \leqslant C \right\}, \tag{3.10}$$

for a positive constant C. For detailed discussions on E, see Gill (1989).

PROPOSITION 3.5. *Let* γ: $D[0,1] \times D[0,1] \to \mathbb{R}$ *be defined by*

$$\gamma(G, H) = \int_0^1 G^2(t)\, dH(t), \qquad G, H \in D[0,1],$$

and let $(G_0, H_0) \in E$ *with* $\int_0^1 |dG_0| < \infty$. *Then* γ *is second-order Hadamard differentiable at* (G_0, H_0) *with derivatives*

$$\gamma'_{(G_0, H_0)}(G, H) = 2 \int_0^1 G(t)\, G_0(t)\, dH_0(t) + \int_0^1 G_0^2(t)\, dH(t),$$
$$G, H \in D[0,1],$$

and

$$\gamma''_{(G_0, H_0)}(G, H) = 2 \int_0^1 G^2(t)\, dH_0(t) + 4 \int_0^1 G(t)\, G_0(t)\, dH(t),$$
$$G, H \in D[0,1].$$

For U_n^* and U_n given in Section 1, the second-order remainder term Rem_2 in (1.7) has the following results which are similar to those for the first-order remainder term. The proofs are given in Section 6.

Let $C[0,1]$ be the space of real valued continuous functions endowed with the supremum norm $\|\cdot\|$. We impose the following assumptions on U_n^* and U_n throughout this paper.

Assumption A. (A1) For some U^* and $\bar{U} \in C[0,1]$, we have that as $n \to \infty$, $C_n[E\{U_n^*\} - U^*] \xrightarrow{\|\cdot\|} 0$ and $\sqrt{n}\,[E\{U_n\} - \bar{U}] \xrightarrow{\|\cdot\|} 0$;

(A2) $n^{-1} C_n^2 \leqslant M$, for all $n \geqslant 1$ and some $0 < M < \infty$;

(A3) $\lim_{n \to \infty} \max_{1 \leqslant i \leqslant n} \{c_{ni}^2\} = 0$.

THEOREM 3.6. *Suppose* τ: $D[0,1] \times D[0,1] \to \mathbb{R}$ *is a functional. Then, under Assumption* A,

(i) *if* τ *is Hadamard differentiable at* $(U^*, \bar{U})$, *we have*

$$C_n\, \mathrm{Rem}_1(U_n^* - U^*, U_n - \bar{U}; \tau) \xrightarrow{P} 0, \qquad \text{as} \quad n \to \infty, \tag{3.11}$$

and

$$C_n[\tau(U_n^*, U_n) - \tau(U^*, \bar{U})] = C_n \tau'_{(U^*, \bar{U})}(U_n^* - U^*, U_n - \bar{U}) + o_p(1); \tag{3.12}$$

(ii) *if* τ *is second-order Hadamard differentiable at* $(U^*, \bar{U})$, *we have*

$$C_n^2\, \mathrm{Rem}_2(U_n^* - U^*, U_n - \bar{U}; \tau) \xrightarrow{P} 0, \qquad \text{as} \quad n \to \infty, \tag{3.13}$$

and

$$C_n^2[\tau(U_n^*, U_n) - \tau(U^*, \bar{U})] = C_n^2 \tau'_{(U^*, \bar{U})}(U_n^* - U^*, U_n - \bar{U}) + \tfrac{1}{2} C_n^2 \tau''_{(U^*, \bar{U})}(U_n^* - U^*, U_n - \bar{U}) + o_p(1). \tag{3.14}$$

Referring to Example 5 in Section 2, we have the following theorem on the second-order remainder term $\mathrm{Rem}_2(V_n(\cdot, \mathbf{u}) - U, V_n^*(\cdot, \mathbf{u}) - U; \tau)$ for the functional τ in (2.22). Let

$$V_n^{*+}(t, \mathbf{u}) = \sum_{i=1}^{n} c_{ni}^{+} I\{Y_i \leqslant F^{-1}(t) + \mathbf{c}_{ni}^{\mathrm{T}} \mathbf{u}\} \tag{3.15}$$

$$V_n^{*-}(t, \mathbf{u}) = \sum_{i=1}^{n} c_{ni}^{-} I\{Y_i \leqslant F^{-1}(t) + \mathbf{c}_{ni}^{\mathrm{T}} \mathbf{u}\}; \tag{3.16}$$

then for this example, we usually have

$$\tau_2(V_n^*(\cdot, \mathbf{u})/a_n) = \tau_1(V_n^{*+}(\cdot, \mathbf{u})/a_n) - \tau_1(V_n^{*-}(\cdot, \mathbf{u})/a_n) \tag{3.17}$$

in (2.23) (see Ren and Sen, 1993).

THEOREM 3.7. *Suppose* $\tau_1 : D[0, 1] \to \mathbb{R}$ *is a functional and is second-order Hadamard differentiable at* U. *Assume that* F *is absolutely continuous with a positive and uniformly continuous derivative. Then, for any* $K > 0$,

(i) *we have*

$$\sup_{|\mathbf{u}| \leqslant K} n \, |\mathrm{Rem}_2(n^{-1/2} V_n(\cdot, \mathbf{u}) - U; \tau_1)| \xrightarrow{P} 0, \qquad \text{as} \quad n \to \infty; \tag{3.18}$$

(ii) *under* (A3), *we have*

$$\sup_{|\mathbf{u}| \leqslant K} a_n^2 \, |\mathrm{Rem}_2(a_n^{-1} V_n^{*+}(\cdot, \mathbf{u}) - U; \tau_1)| \xrightarrow{P} 0, \qquad \text{as} \quad n \to \infty$$

and

$$\sup_{|\mathbf{u}| \leqslant K} a_n^2 \, |\mathrm{Rem}_2(a_n^{-1} V_n^{*-}(\cdot, \mathbf{u}) - U; \tau_1)| \xrightarrow{P} 0, \qquad \text{as} \quad n \to \infty.$$

The proof of Theorem 3.7 is similar to that of Theorem 3.1 in Ren and Sen (1991), where we only need to replace $t^{-1} \mathrm{Rem}(tH; \tau_1)$ by $t^{-2} \mathrm{Rem}_2(tH; \tau_1)$. One may note that the Skorohod topology was used in

Ren and Sen (1991), but their results apply here (and in Section 6) because the limiting Gaussian process for the weak convergence here and in Ren and Sen (1991) has continuous sample path (see discussions in Gill, 1989) and the use of the σ-field generated by all open balls in this paper ensures the measurability of the weighted empirical processes.

Remark 3. Ren and Sen (1993) have used Theorem 3.7 along with Proposition 3.2 to study the convergence rate in probability of (2.21). Thereby, an asymptotic representation of M-estimators in linear models is given under weaker conditions on the score function ψ, the error distribution F and the design matrix than those in Jurěcková and Sen (1984).

4. LIMITING DISTRIBUTIONS OF $\tau(U_n^*, U_n)$

In this section, we will consider the functional τ given by (1.4). Let τ: $D[0,1] \times D[0,1] \to \mathbb{R}$ be a functional, given by

$$\tau(G, H) = \int_0^1 [G(x)]^2 \psi(H(x))\, dH(x), \qquad G, H \in D[0,1]. \tag{4.1}$$

The following conditions may be required in our theorems.

Assumption B. (B1) ψ is positive on $[0,1]$ with continuous derivative ψ';

(B2) For any $\delta > 0$, ψ is positive on $[\delta, 1-\delta]$ with continuous derivative ψ';

(B3) There exists $\delta_0 > 0$ and $0 < M_1$, $M_2 < \infty$ such that $\psi(t) \leqslant M_1/t$, $t \in (0, \delta_0]$ and $\psi(t) \leqslant M_2/(1-t)$, $t \in [1-\delta_0, 1)$.

In the next lemma, we show that τ given by (4.1) is second-order Hadamard differentiable with the proof deferred to Section 6.

LEMMA 4.1. *Let $\phi = \sqrt{\psi}$ and let $(G_0, H_0) \in E$ with $\int_0^1 |dG_0| < \infty$. Then,*

(i) *Under Assumption* (B1), *we have that the functional τ given by* (4.1) *is first-order Hadamard differentiable at (G_0, H_0) with the first-order Hadamard derivative*

$$\tau'_{(G_0, H_0)}(G, H) = 2 \int_0^1 [G_0 \phi'(H_0) H + \phi(H_0) G]\, G_0 \phi(H_0)\, dH_0 + \int_0^1 G_0^2 \phi^2(H_0)\, dH; \tag{4.2}$$

(ii) *If in addition to Assumption* (B1), *we assume that ψ has continuous second derivative ψ'', then the functional τ given by* (4.1) *is second-order Hadamard differentiable at* (G_0, H_0) *with the first-order Hadamard derivative given by* (4.2), *and the second-order Hadamard derivative*

$$\tau''_{(G_0, H_0)}(G, H) = 2 \int_0^1 [G_0 \phi'(H_0) H + \phi(H_0) G]^2 \, dH_0$$
$$+ 4 \int_0^1 [G_0 \phi'(H_0) H + \phi(H_0) G] \, G_0 \phi(H_0) \, dH$$
$$+ 2 \int_0^1 [2\phi'(H_0) GH + G_0 \phi''(H_0) H^2] \, G_0 \phi(H_0) \, dH_0, \tag{4.3}$$

where $G, H \in D[0, 1]$.

In a special case of Lemma 4.1 with $G_0 \equiv 0$, weaker conditions are required on ψ. We state this case as a corollary. From Corollary 3.4 and Proposition 3.5, the proof is similar to that of Lemma 4.1.

Corollary 4.2. *Under Assumption* (B1), *for* $H_0 \in D[0, 1]$ *with* $\int_0^1 |dH_0| < \infty$, *the functional τ given by* (4.1) *is second-order Hadamard differentiable at* $(0, H_0)$ *with the first-order Hadamard derivative*

$$\tau_{(0, H_0)}(G, H) \equiv 0, \qquad G, H \in D[0, 1] \tag{4.4}$$

and the second-order Hadamard derivative

$$\tau''_{(0, H_0)}(G, H) = 2 \int_0^1 \psi(H_0) \, G^2 \, dH_0, \qquad G, H \in D[0, 1]. \tag{4.5}$$

To study the asymptotic distribution of $\tau(U_n^*, U_n)$ given by (1.4), we are particularly interested in the following special case for (U_n^*, U_n):

Assumption C. $X_1, X_2, ..., X_n$ are i.i.d. with a continuous d.f. F.

In this case, we have $U^* = 0$ and $\bar{U} = U$ in Assumption (A1), because we always have

$$\sum_{i=1}^{n} c_{ni} = 0, \qquad n \geqslant 1. \tag{4.6}$$

THEOREM 4.3. *Under Assumption* (A), (B1) *and* (C), *the functional* τ *given by* (4.1) *is second-order Hadamard differentiable at* $(0, U)$ *with derivatives*

$$\tau'_{(0,\,U)}(G, H) \equiv 0 \tag{4.7}$$

and

$$\tau''_{(0,\,U)}(G, H) = 2 \int_0^1 G^2 \psi(U)\, dU, \tag{4.8}$$

where $G, H \in D[0, 1]$. *Therefore,*

$$C_n^2 \tau(U_n^*, U_n) = C_n^2 \int_0^1 U_n^{*2} \psi(U)\, dU + o_p(1), \qquad as \quad n \to \infty \tag{4.9}$$

and

$$C_n^2 \tau(U_n^*, U_n) \xrightarrow{D} \int_0^1 W^2(t)\, \psi(t)\, dt = \sum_{j=1}^{\infty} \lambda_j Z_j^2, \qquad as \quad n \to \infty, \tag{4.10}$$

where W *is a Gaussian process on* $[0, 1]$ *with mean* 0 *and covariance*

$$\gamma(s, t) = \min\{s, t\} - st, \tag{4.11}$$

Z_j *are independent standard normal random variables, and* λ_j *are the eigenvalues for the following eigenvalue problem*:

$$\int_0^1 \gamma(s, t) \sqrt{\psi(s)} \sqrt{\psi(t)}\, \chi(t)\, dt = \lambda \chi(s). \tag{4.12}$$

The proof of Theorem 4.3 is given in Section 6. It is clear that in Theorem 4.3, we require ψ to be bounded. To handle more general weight functions, we first establish the following lemma, then extend Theorem 4.3 to unbounded weight functions. The proofs are deferred to Section 6. Under assumption (C), we denote

$$W_n(t) = C_n U_n^*(t) = \sum_{i=1}^{n} c_{ni}(I\{Y_i \leqslant t\} - t), \qquad t \in [0, 1], \tag{4.13}$$

where $Y_i = F(X_i)$ are i.i.d.r.v.'s with d.f. U.

Lemma 4.4. *Under Assumption* (A), (B2)–(B3) *and* (C), *we have that for any* $\varepsilon > 0$, *there exists* $\delta > 0$ *and* N *such that for* $n \geqslant N$

$$P\left\{\left|\int_0^\delta W_n^2 \psi(U_n)\, dU_n\right| \leqslant \varepsilon\right\} \geqslant 1-\varepsilon \tag{4.14}$$

$$P\left\{\left|\int_{1-\delta}^1 W_n^2 \psi(U_n)\, dU_n\right| \leqslant \varepsilon\right\} \geqslant 1-\varepsilon \tag{4.15}$$

$$P\left\{\left|\int_0^\delta W_n^2 \psi(U)\, dU\right| \leqslant \varepsilon\right\} \geqslant 1-\varepsilon \tag{4.16}$$

$$P\left\{\left|\int_{1-\delta}^1 W_n^2 \psi(U)\, dU\right| \leqslant \varepsilon\right\} \geqslant 1-\varepsilon. \tag{4.17}$$

Theorem 4.5. *Under Assumption* (A), (B2)–(B3) *and* (C), *the functional* τ *given by* (4.1) *satisfies*

$$C_n^2 \tau(U_n^*, U_n) = C_n^2 \int_0^1 U_n^{*2} \psi(U)\, dU + o_p(1), \qquad as \quad n \to \infty, \tag{4.18}$$

$$C_n^2 \tau(U_n^*, U_n) \xrightarrow{D} \int_0^1 W^2(t)\, \psi(t)\, dt = \sum_{j=1}^{\infty} \lambda_j Z_j^2, \qquad as \quad n \to \infty, \tag{4.19}$$

where W, Z_j *and* λ_j *are as those in Theorem* 4.3.

Remark 4. Note that our condition (B2) is required for having the decomposition of $W(t)\sqrt{\psi(t)}$ in (4.19) (see Anderson and Darling, 1952).

5. APPLICATIONS

In this section, we give the specific limiting distributions of those test statistics in Examples 1–4 and 6 of Sections 1 and 2.

Example 1. In model (1.2), we assume that the error variables ε_i's are continuous. Then, we know that under H_0, $X_1, ..., X_n$ are i.i.d.r.v.'s with some continuous d.f. F. Assuming (A2)–(A3), from Theorem 4.5, we know that the test statistics $\Psi(S_n^*, S_n)$ has the following limiting distribution under H_0,

$$\begin{aligned} C_n^2 \Psi(S_n^*, S_n) = C_n^2 \tau(U_n^*, U_n) &\xrightarrow{D} \int_0^1 W^2(t)\, \psi(t)\, dt \\ &= \sum_{j=1}^{\infty} \lambda_j Z_j^2, \qquad \text{as} \quad n \to \infty, \end{aligned} \tag{5.1}$$

where W, Z_j and λ_j are as those in Theorem 4.3, ψ satisfies Assumption (B2)–(B3), and τ is given by (4.1). Particularly, for some special weight function ψ, the values of λ_j in (5.1) have been given. If $\psi \equiv 1$, then

$$\lambda_j = 1/(j\pi)^2, \qquad j = 1, 2, ... \tag{5.2}$$

(Shorack and Wellner, 1986, p. 214). If $\psi(t) = [t(1-t)]^{-1}$, then

$$\lambda_j = 1/[j(j+1)], \qquad j = 1, 2, ... \tag{5.3}$$

(Anderson and Darling, 1952).

EXAMPLE 2. If the error distribution is continuous in model (2.1), we have that $X_1, ..., X_n$ are i.i.d.r.v.'s with a continuous d.f. F under H_0. If for $\{c_{ni}\}$ and $\{d_{ni}\}$, (A2)–(A3) hold with $\rho_n = \sum_{i=1}^n c_{ni} d_{ni} \to \rho$, as $n \to \infty$, then by Theorem 4.5, we have that under H_0,

$$C_n^2 \tau(U_n^*, U_n) + D_n^2 \tau(V_n^*, U_n)$$
$$= C_n^2 \int_0^1 U_n^{*2} \psi(U)\, dU + D_n^2 \int_0^1 V_n^{*2} \psi(U)\, dU + o_p(1),$$
$$\text{as} \quad n \to \infty,$$

where $V_n^* = T_n^* \circ F^{-1}$. Furthermore, from a generalization of Theorem 3.1.1 of Shorack and Wellner (1986, p. 93) and from the proof of our Theorem 4.3 given in Section 6, we have

$$C_n^2 \int_0^1 U_n^{*2} \psi(U)\, dU = \int_0^1 W_1^2(t)\, \psi(t)\, dt + o_p(1), \qquad \text{as} \quad n \to \infty,$$

and

$$D_n^2 \int_0^1 V_n^{*2} \psi(U)\, dU = \int_0^1 W_2^2(t)\, \psi(t)\, dt + o_p(1), \qquad \text{as} \quad n \to \infty,$$

where W_1 and W_2 are Brownian bridges with $\text{Cov}\{W_1(s), W_2(t)\} = \rho[\min\{s, t\} - st]$. Hence, we have that under H_0,

$$\int_{-\infty}^{\infty} (C_n S_n^*(x), D_n T_n^*(x))(C_n S_n^*(x), D_n T_n^*(x))^{\mathrm{T}}\, \psi(S_n(x))\, dS_n(x)$$
$$= C_n^2 \Psi(S_n^*, S_n) + D_n^2 \Psi(T_n^*, S_n) = C_n^2 \tau(U_n^*, U_n) + D_n^2 \tau(V_n^*, U_n)$$
$$\xrightarrow{D} \int_0^1 [W_1^2(t) + W_2^2(t)]\, \psi(t)\, dt$$
$$= \sum_{j=1}^{\infty} \lambda_j (Z_j^2 + \hat{Z}_j^2), \qquad \text{as} \quad n \to \infty, \tag{5.4}$$

and

$$\max\left\{C_n^2\int_{-\infty}^{\infty}[S_n^*(x)]^2\,\psi(S_n(x))\,dS_n(x),\right.$$

$$\left.D_n^2\int_{-\infty}^{\infty}[T_n^*(x)]^2\,\psi(S_n(x))\,dS_n(x)\right\}$$

$$=\max\{C_n^2\Psi(S_n^*,S_n),D_n^2\Psi(T_n^*,S_n)\}$$

$$\xrightarrow{D}\max\left\{\sum_{j=1}^{\infty}\lambda_jZ_j^2,\ \sum_{j=1}^{\infty}\lambda_j\hat{Z}_j^2\right\},\qquad \text{as}\quad n\to\infty, \tag{5.5}$$

where λ_j are given by (4.12), $(Z_j, \hat{Z}_j)$ are i.i.d. bivariate normal r.v.'s with $E\{Z_j\}=E\{\hat{Z}_j\}=0$, $\mathrm{Var}\{Z_j\}=\mathrm{Var}\{\hat{Z}_j\}=1$ and $\mathrm{Cov}\{Z_j,\hat{Z}_j\}=\rho$, ψ satisfies Assumption (B2)–(B3), and τ is given by (4.1). For the special weight function $\psi\equiv 1$ and $\psi(t)=[t(1-t)]^{-1}$, the values of λ_j are given by (5.2) and (5.3), respectively.

To find the critical values of the tests using (5.4) or (5.5), one may generate i.i.d. standard normal r.v.'s: $Y_1, \hat{Y}_1, ..., Y_N, \hat{Y}_N$ for some large N. Then, set $Z_j=Y_j$ and $\hat{Z}_j=\rho_nY_j-\sqrt{1-\rho_n^2}\,\hat{Y}_j$, $j=1, ..., N$. Since $\sum_{j=1}^{\infty}\lambda_j(Z_j^2+\hat{Z}_j^2)\approx\sum_{j=1}^{N}\lambda_j(Z_j^2+\hat{Z}_j^2)$ in (5.4) and $\max\{\sum_{j=1}^{\infty}\lambda_jZ_j^2,\sum_{j=1}^{\infty}\lambda_j\hat{Z}_j^2\}\approx\max\{\sum_{j=1}^{N}\lambda_jZ_j^2,\sum_{j=1}^{N}\lambda_j\hat{Z}_j^2\}$ in (5.5), the critical values of the test statistics (2.4) and (2.5) can be determined by the quantiles of $\sum_{j=1}^{N}\lambda_j(Z_j^2+\hat{Z}_j^2)$ and $\max\{\sum_{j=1}^{N}\lambda_jZ_j^2,\sum_{j=1}^{N}\lambda_j\hat{Z}_j^2\}$, respectively.

If a general linear model is considered,

$$X_i=\alpha+\mathbf{c}_i^{\mathrm{T}}\boldsymbol{\beta}+\varepsilon_i,\qquad i=1, ..., n, \tag{5.6}$$

where $\mathbf{c}_i$ are known p-vectors of regression constants, $\boldsymbol{\beta}$ is the p-vector unknown regression parameters, and ε_i are i.i.d.r.v.'s with a continuous d.f., we let

$$\mathbf{D}_n=(\mathbf{c}_1, ..., \mathbf{c}_n)^{\mathrm{T}},\qquad \bar{\mathbf{D}}_n=n^{-1}\mathbf{1}_n\mathbf{1}_n^{\mathrm{T}}\mathbf{D}_n$$

$$\mathbf{C}_n=\mathrm{Diag}(\|\mathbf{d}_1\|_E, ..., \|\mathbf{d}_n\|_E),\qquad \mathbf{c}_{ni}=\mathbf{C}_n^{-1}(\mathbf{c}_i-\bar{\mathbf{c}}_n)$$

$$\boldsymbol{\Sigma}_n=\sum_{i=1}^{n}\mathbf{c}_{ni}\mathbf{c}_{ni}^{\mathrm{T}},$$

where $\mathbf{d}_i$ is the column vector of $\mathbf{D}_n-\bar{\mathbf{D}}_n$, $\bar{\mathbf{c}}_n$ is a column of $\bar{\mathbf{D}}_n$, and $\|\cdot\|_E$ stands for Euclidean norm. Also, let

$$\mathbf{S}_n^*(x)=\mathbf{C}_n^{-1}\sum_{i=1}^{n}\mathbf{c}_{ni}I\{X_i\leqslant x\} \tag{5.7}$$

with $\lim_{n\to\infty} \max_{1\leqslant i\leqslant n} \|\mathbf{c}_{ni}\|_E^2 = 0$, $\lim_{n\to\infty} \boldsymbol{\Sigma}_n = \boldsymbol{\Sigma}$ and $n^{-1}\|1^{\mathrm{T}}\mathbf{C}_n\|_E^2 \leqslant M < \infty$. Then, under H_0: $\boldsymbol{\beta} = \mathbf{0}$ (vs. H_1: $\boldsymbol{\beta} \neq \mathbf{0}$), we have that the test statistics

$$\int_{-\infty}^{\infty} (\mathbf{C}_n, \mathbf{S}_n^*)^{\mathrm{T}} (\mathbf{C}_n \mathbf{S}_n^*)\, \psi(S_n)\, dS_n \xrightarrow{D} \sum_{j=1}^{\infty} \lambda_j \mathbf{Z}_j^{\mathrm{T}} \mathbf{Z}_j, \qquad \text{as} \quad n\to\infty, \tag{5.8}$$

and

$$\max_{1\leqslant k\leqslant p} \left\{ \int_{-\infty}^{\infty} (\mathbf{e}_k^{\mathrm{T}} \mathbf{C}_n \mathbf{S}_n^*)^2\, \psi(S_n)\, dS_n \right\} \xrightarrow{D} \max_{1\leqslant k\leqslant p} \left\{ \sum_{j=1}^{\infty} \lambda_j (\mathbf{e}_k^{\mathrm{T}} \mathbf{Z}_j)^2 \right\}, \qquad \text{as} \quad n\to\infty, \tag{5.9}$$

where $\mathbf{e}_k$ is a $p\times 1$ vector with the kth component 1 and others 0, λ_j are given by (4.12), $\mathbf{Z}_j$ are i.i.d. $N_p(\mathbf{0}, \boldsymbol{\Sigma})$ and ψ satisfies Assumption (B2)–(B3). The critical values of the tests can be obtained similarly as that for the case of $p = 2$ outlined above.

Example 3. Consider the test statistic given in (2.9) for the 3-sample problem (2.6) with non-equal sample size. For $\{c_{ni}\}$ and $\{d_{ni}\}$, we have

$$\rho_n = \sum_{i=1}^{n} c_{ni} d_{ni} = a_{n1} b_{n1} + a_{n2} b_{n2} + a_{n3} b_{n3} = 0$$

and (A2)–(A3) hold if $n_j \to \infty$, as $n\to\infty$, $j = 1, 2, 3$. Hence, if (B2)–(B3) hold for the weight function ψ, from (5.4) the test statistic (2.9) has the following limiting distribution under H_0,

$$\int_{-\infty}^{\infty} \left\{ n_1 \left(F_{n_1} - \frac{n+1}{n} S_n \right)^2 + n_2 \left(G_{n_2} - \frac{n+1}{n} S_n \right)^2 + n_3 \left(H_{n_3} - \frac{n+1}{n} S_n \right)^2 \right\} \psi(S_n)\, dS_n \xrightarrow{D} \sum_{j=1}^{\infty} \lambda_j \chi_j^2, \qquad \text{as} \quad n\to\infty, \tag{5.10}$$

where χ_j^2 are i.i.d. Chi-square random variables with degrees of freedom 2, and for the special weight function $\psi \equiv 1$ and $\psi(t) = [t(1-t)]^{-1}$, the values of λ_j are given by (5.2) and (5.3), respectively.

EXAMPLE 4. *Case* (b): Consider the test statistic given in (2.14) for the 3-sample problem (2.6). For $\{c_{ni}\}$, $\{d_{ni}\}$ and $\{e_{ni}\}$ in (2.13), we have

$$\rho_n(c, d) = \sum_{k=1}^{n} c_{nk} d_{nk} = -\frac{\sqrt{n_1 n_2}}{\sqrt{(n-n_1)(n-n_2)}}$$
$$\rho_n(c, e) = -\frac{\sqrt{n_1 n_3}}{\sqrt{(n-n_1)(n-n_3)}}, \qquad \rho_n(d, e) = -\frac{\sqrt{n_2 n_3}}{\sqrt{(n-n_2)(n-n_3)}}. \tag{5.11}$$

Suppose that $n_j(1-\frac{n_j}{n}) \to \infty$, as $n \to \infty$, $j = 1, 2, 3$, and $(\rho_n(c, d), \rho_n(c, e), \rho_n(d, e))$ converges to $(\rho_{cd}, \rho_{ce}, \rho_{de})$, as $n \to \infty$, and that (B2)–(B3) hold for the weight function ψ. Then, (A2)–(A3) hold, and from (5.9), the test statistic (2.14) has the following limiting distribution under H_0,

$$\max \left\{ \frac{n_1 n}{n-n_1} \int_{-\infty}^{\infty} \left(F_{n_1} - \frac{n+1}{n} S_n \right)^2 \psi(S_n)\, dS_n, \right.$$
$$\frac{n_2 n}{n-n_2} \int_{-\infty}^{\infty} \left(G_{n_2} - \frac{n+1}{n} S_n \right)^2 \psi(S_n)\, dS_n,$$
$$\left. \frac{n_3 n}{n-n_3} \int_{-\infty}^{\infty} \left(H_{n_3} - \frac{n+1}{n} S_n \right)^2 \psi(S_n)\, dS_n \right\}$$
$$\xrightarrow{D} \max_{1 \leqslant k \leqslant 3} \left\{ \sum_{j=1}^{\infty} \lambda_j (\mathbf{e}_k^{\mathrm{T}} \mathbf{Z}_j)^2 \right\}, \qquad \text{as} \quad n \to \infty, \tag{5.12}$$

where $\mathbf{Z}_j$ are i.i.d. $N_3(\mathbf{0}, \boldsymbol{\Sigma})$ with

$$\boldsymbol{\Sigma} = \begin{bmatrix} 1 & \rho_{cd} & \rho_{ce} \\ \rho_{cd} & 1 & \rho_{de} \\ \rho_{ce} & \rho_{de} & 1 \end{bmatrix}, \tag{5.13}$$

and for the special weight function $\psi \equiv 1$ and $\psi(t) = [t(1-t)]^{-1}$, the values of λ_j are given by (5.2) and (5.3), respectively. In particular, for *Case* (a) with equal sample size as described in Section 2, we have $\rho_{cd} = \rho_{ce} = \rho_{de} = -1/2$ in (5.13).

Case (d): Consider the test statistics given by (2.18) and (2.18a) for the 3-sample problem (2.6). For $\{c_{ni}\}$, $\{d_{ni}\}$ and $\{e_{ni}\}$ in (2.17), we have

$$\rho'_n(c, d) = \rho_n(c, e), \qquad \rho'_n(c, e) = \rho_n(d, e), \qquad \rho'_n(d, e) = \rho_n(c, d), \tag{5.14}$$

where $\rho_n(c, d)$, $\rho_n(c, e)$ and $\rho_n(d, e)$ are given in (5.11). Suppose that $n_j \to \infty$, as $n \to \infty$, $j = 1, 2, 3$, and $(\rho'_n(c, d), \rho'_n(c, e), \rho'_n(d, e))$ converges to $(\rho'_{cd}, \rho'_{ce}, \rho'_{de})$, as $n \to \infty$, and that (B2)–(B3) hold for the weight function ψ. Then, (A2)–(A3) hold, and from (5.9) and (5.8), the test statistics given by (2.18) and (2.18a) have the following limiting distributions under H_0,

$$\max\left\{\frac{n_1n_2}{n_1+n_2}\int_{-\infty}^{\infty}(F_{n_1}-G_{n_2})^2\,\psi(S_n)\,dS_n,\right.$$
$$\frac{n_2n_3}{n_2+n_3}\int_{-\infty}^{\infty}(G_{n_2}-H_{n_3})^2\,\psi(S_n)\,dS_n,$$
$$\left.\frac{n_1n_3}{n_1+n_3}\int_{-\infty}^{\infty}(H_{n_3}-F_{n_1})^2\,\psi(S_n)\,dS_n\right\}$$
$$\xrightarrow{D}\max_{1\leqslant k\leqslant 3}\left\{\sum_{j=1}^{\infty}\lambda_j(\mathbf{e}_k^{\mathrm{T}}\mathbf{Z}_j)^2\right\},\qquad \text{as}\quad n\to\infty \tag{5.15}$$

and

$$\int_{-\infty}^{\infty}\left\{\frac{n_1n_2}{n_1+n_2}(F_{n_1}-G_{n_2})^2+\frac{n_2n_2}{n_2+n_3}(G_{n_2}-H_{n_2})^2\right.$$
$$\left.+\frac{n_1n_3}{n_1+n_3}(H_{n_3}-F_{n_1})^2\right\}\psi(S_n)\,dS_n$$
$$\xrightarrow{D}\sum_{j=1}^{\infty}\lambda_j\mathbf{Z}_j^{\mathrm{T}}\mathbf{Z}_j,\qquad \text{as}\quad n\to\infty, \tag{5.16}$$

respectively, where for $\mathbf{\Sigma}'$ given by (5.13) with ρ_{cd}, ρ_{ce}, ρ_{de} replaced by ρ'_{cd}, ρ'_{ce}, ρ'_{de}, respectively, $\mathbf{Z}_j$ are i.i.d. $N_3(\mathbf{0},\mathbf{\Sigma}')$, and for the special weight function $\psi\equiv 1$ and $\psi(t)=[t(1-t)]^{-1}$, the values of λ_j are given by (5.2) and (5.3), respectively. In particular, for *Case* (c) with equal sample size as described in Section 2, we have $\rho'_{cd}=\rho'_{ce}=\rho'_{de}=-1/2$, thus $\mathbf{\Sigma}=\mathbf{\Sigma}'$.

Remark 5. It is worth mentioning that with equal sample size in the 3-sample problem (2.6), the limiting distribution of the test statistic given by (2.18a) is $\frac{3}{2}\sum_{j=1}^{\infty}\lambda_j\chi_j^2$, where χ_j^2 are i.i.d. Chi-square random variables with degrees of freedom 2, because the eigenvalues of $\mathbf{\Sigma}'$ in (5.16) are 0, 3/2 and 3/2. Also, one may note that by using Theorem 3.6 we can easily show that under a fixed alternative hypothesis, the limiting distribution of the test statistic given by (2.18a) is normal.

Remark 6. One may note that in Examples 1–2, it is shown that the SOHD method allows us to generalize the Cramér–von Mises type of test statistics (Hájek and Šidák, 1967, p. 103) from model (1.2) to (2.1) in a rather straightforward way and allows us to derive their asymptotic distributions conveniently. The investigation on the power of the proposed tests here in comparison with the alternative tests, say, (normal) rank tests, is technical, which will not be studied in this current paper. In Hájek and Šidák (1967, pp. 229–232), a comparison of the asymptotic local power of the one-sided Kolmogorov–Smirnov test with that of the normal rank test was given. We may expect similar results for our generalized Cramér–von Mises tests. One may also note that for the 3-sample problem (2.6), the null limiting distributions of the test statistic studied by Scholz and Stephens (1987) and the alternative test statistics constructed in Example 4 are just the special cases of that in Example 2 for linear regression model.

EXAMPLE 6. Consider the test statistic given by (2.27) for the independence test (2.24), and let $\mathbf{x} = (u, v)$, $\mathbf{y} = (s, t)$, $F_1(u) = F(u, \infty)$ and $F_2(v) = F(\infty, v)$. It can be shown that under H_0, $\sqrt{m}\, S_n^*(F_1^{-1}(u), F_2^{-1}(v)) = \sqrt{m}\, U_n^*(u, v)$ weakly converges to a centered Gaussian process $G(\mathbf{x})$ with covariance function

$$\gamma(\mathbf{x}, \mathbf{y}) = (\min\{u, s\} - us)(\min\{v, t\} - vt),$$

and $S_n(F_1^{-1}(u), F_2^{-1}(v)) = U_n(u, v)$ uniformly converges to uv with probability 1. Following the concepts in Ren and Sen (1995) for the first-order Hadamard derivative with bivariate random vectors, it is easy to generalize Theorems 3.6 and 4.3 to their bivariate versions under suitable conditions. Thus, we have that under H_0,

$$\begin{aligned} C_n^2 \Psi(S_n^*, S_n) &= C_n^2 \int_0^1 \int_0^1 [U_n^*(u, v)]^2\, \psi(U_n(u, v))\, dU_n(u, v) \\ &\xrightarrow{D} \int_0^1 \int_0^2 G^2(u, v)\, \psi(uv)\, du\, dv, \qquad \text{as} \quad m \to \infty. \end{aligned}$$

6. PROOFS

Proof of Proposition 3.1. Since T and Q are second-order Hadamard differentiable at F and $T(F)$, respectively, for any compact set Γ_V of V and compact set Γ_W of W, we have

$$\lim_{t \to 0} \frac{T(F + tH) - T(F) - T'_F(tH) - \frac{1}{2} T''_F(tH)}{t^2} = 0 \tag{6.1}$$

uniformly for any $H \in \Gamma_V$, and

$$\lim_{t \to 0} \frac{Q(T(F) + tG) - Q(T(F)) - Q'_{T(F)}(tG) - \frac{1}{2} Q''_{T(F)}(tG)}{t^2} = 0. \tag{6.2}$$

uniformly for any $G \in \Gamma_W$. By Proposition 3.1.2 of Fernholz (1983), we know

$$\tau'_F = Q'_{T(F)} \circ T'_F,$$

and obviously $\tau'_F \in \mathfrak{L}_1(V, Z)$. It is also obvious that τ''_F given by (3.9) is an element of $\mathfrak{L}_2(V, Z)$. From (6.1) we have

$$T(F + tH) = T(F) + T'_F(tH) + \tfrac{1}{2} T''_F(tH) + o(1)\, t^2,$$

where $o(1)$ converges to 0 uniformly for any $H \in \Gamma_V$, as $t \to 0$. Hence, by the linearity of $Q'_{T(F)}$, we have

$$\begin{aligned}
&\tau(F + tH) - \tau(F) - \tau'_F(tH) \\
&\quad = Q(T(F + tH)) - Q(T(F)) - Q'_{T(F)}(T'_F(tH)) \\
&\quad = Q(T(F) + T'_F(tH) + \tfrac{1}{2} T''_F(tH) + o(1)\, t^2) \\
&\qquad - Q(T(F)) - Q'_{T(F)}(T'_F(tH)) \\
&\quad = Q(T(F) + t\{ T'_F(H) + \tfrac{1}{2} t T''_F(H) + o(1)\, t\}) - Q(T(F)) \\
&\qquad - Q'_{T(F)}(t\{ T'_F(H) + \tfrac{1}{2} t T''_F(H) + o(1)\, t\}) \\
&\qquad + \tfrac{1}{2} t^2 Q'_{T(F)}(T''_F(H)) + t^2 Q'_{T(F)}(o(1)).
\end{aligned} \tag{6.3}$$

It is easy to show that

$$\begin{aligned}
\Gamma'_W &= \{ T'_F(H) + \tfrac{1}{2} t T''_F(H) + o(1)\, t;\ H \in \Gamma_V,\ t \in [-1, 1]\} \\
&= \{ T'_F(H) + \kappa(H, t);\ H \in \Gamma_V,\ t \in [-1, 1]\}
\end{aligned}$$

is compact, where $\kappa(H, t) = [T(F + tH) - T(F) - T'_F(tH)]/t$ converges to 0 uniformly for any $H \in \Gamma_V$, as $t \to 0$. Hence, by (6.2), we have

$$\begin{aligned}
&\lim_{t \to 0} \frac{1}{t^2} \{\{ Q(T(F) + t\{ T'_F(H) + \tfrac{1}{2} t T''_F(H) + o(1)\, t\}) - Q(T(F)) \\
&\qquad - Q'_{T(F)}(t\{ T'_F(H) + \tfrac{1}{2} t T''_F(H) + o(1)\, t\})\} - \tfrac{1}{2} Q''_{T(F)}(T'_F(tH))\} \\
&\quad = \lim_{t \to 0} \frac{Q''_{T(F)}(t\{ T'_F(H) + \tfrac{1}{2} t T''_F(H) + o(1)\, t\}) - Q''_{T(F)}(T'_F(tH))}{2t^2} \\
&\quad = \lim_{t \to 0} \tfrac{1}{2}\{ Q''_{T(F)}(T'_F(H) + \tfrac{1}{2} t T''_F(H) + o(1)\, t) - Q''_{T(F)}(T'_F(H))\} \\
&\quad = \lim_{t \to 0} \tfrac{1}{2}\{ Q''_{T(F)}(T'_F(H) + \kappa(H, t)) - Q''_{T(F)}(T'_F(H))\}
\end{aligned} \tag{6.4}$$

uniformly for any $H \in \Gamma_V$. Let

$$\phi(H, t) = Q''_{T(F)}(T'_F(H) + \kappa(H, t)) - Q''_{T(F)}(T'_F(H));$$

then ϕ: $V \times [-1, 1] \to W$ is continuous and $\phi(H, t)$ converges to 0 for any fixed $H \in \Gamma_V$, as $t \to 0$. Hence, for any open set O of W such that $0 \in O$ and for any $H_0 \in \Gamma_V$, there exists an open set $O_{(H_0, 0)} = O_{H_0} \times I_{(H_0, 0)}$ of $V \times [-1, 1]$ such that O_{H_0} is an open set of V and $H_0 \in O_{H_0}$, $I_{(H_0, 0)}$ is an open interval and $0 \in I_{(H_0, 0)}$, and $[\phi(H, t) - \phi(H_0, 0)] \in O$ if $(H, t) \in O_{(H_0, 0)}$. Since $\{O_H; H \in \Gamma_V\}$ is an open covering of Γ_V and since Γ_V is compact, there exists a finite open covering of Γ_V, say, $\Gamma_V \subset \bigcup_{i=1}^{N} O_{H_i}$. Therefore, for any $H \in \Gamma_V$ and any small enough t, there exists i such that $(H, t) \in O_{(H_i, 0)}$. Hence, $\phi(H, t) = [\phi(H, t) - \phi(H_i, 0)] \in O$. Therefore, we have

$$\lim_{t \to 0} \tfrac{1}{2}\{Q''_{T(F)}(T'_F(H) + \kappa(H, t)) - Q''_{T(F)}(T'_F(H))\} = 0 \tag{6.5}$$

uniformly for any $H \in \Gamma_V$. Since $Q'_{T(F)}$ is continuous, then

$$\lim_{t \to 0} \frac{t^2 Q'_{T(F)}(o(1))}{t^2} = \lim_{t \to 0} Q'_{T(F)}(o(1)) = 0 \tag{6.6}$$

uniformly for any $H \in \Gamma_V$. Therefore, (6.3) through (6.6) imply that

$$\lim_{t \to 0} \frac{\tau(F + tH) - \tau(F) - \tau'_F(tH) - \frac{1}{2}\{Q''_{T(F)}(T'_F(tH)) + Q'_{T(F)}(T''_F(tH))\}}{t^2} = 0$$

uniformly for any $H \in \Gamma_V$. ∎

Proof of Proposition 3.2. For any compact set Γ of $D[0, 1]$, we need to show that

$$\left\| \frac{\mathrm{Rem}_2(tH)}{t^2} \right\|_{L^p} \to 0 \tag{6.7}$$

uniformly for $H \in \Gamma$, as $t \to 0$, where

$$\mathrm{Rem}_2(tH) = L \circ (Q + tH) - L \circ Q - (L' \circ Q)\, tH - \tfrac{1}{2}(L'' \circ Q)\, t^2 H^2.$$

Since Γ is a compact set, for arbitrary $\varepsilon > 0$, we can choose $H_1, ..., H_n \in \Gamma$ such that for any $H \in \Gamma$,

$$\inf_{1 \leqslant i \leqslant n} \|H - H_i\| < \varepsilon. \tag{6.8}$$

Since for a given H_i,

$$\frac{\text{Rem}_2(tH_i)(x)}{t^2} = \frac{L(Q(x) + tH_i(x)) - L(Q(x)) - L'(Q(x))\, tH_i(x) - \frac{1}{2}L''(Q(x))\, t^2 H_i^2(x)}{t^2}$$

$$= \frac{L'(\xi) - L'(Q(x))}{t} H_i(x) - \tfrac{1}{2}L''(Q(x))\, H_i^2(x),$$

where ξ is between $Q(x)$ and $Q(x) + tH_i(x)$, by Lemma 5.4.3 of Fernholz (1983), we have that

$$\left|\frac{L'(\xi) - L'(Q(x))}{t}\right| \leqslant M\, |H_i(x)|,$$

where M is a bound for L''. Therefore, for each i,

$$\left|\frac{\text{Rem}_2(tH_i)(x)}{t^2}\right| \leqslant M_1\, |H_i(x)|^2,$$

where M_1 is a constant. Moreover, for x such that $Q(x) \notin \mathbb{A}$,

$$\frac{\text{Rem}_2(tH_i)(x)}{t^2} \to 0, \qquad \text{as} \quad t \to 0.$$

So, by the Dominated Convergence Theorem, we have

$$\left\|\frac{\text{Rem}_2(tH_i)}{t^2}\right\|_{L^p} \to 0, \qquad \text{as} \quad t \to 0. \tag{6.9}$$

For any $H \in \Gamma$,

$$\left\|\frac{\text{Rem}_2(tH)}{t^2}\right\|_{L^p} \leqslant \left\|\frac{\text{Rem}_2(tH_i)}{t^2}\right\|_{L^p} + \left\|\frac{\text{Rem}_2(tH)}{t^2} - \frac{\text{Rem}_2(tH_i)}{t^2}\right\|_{L^p} \tag{6.10}$$

and

$$\frac{\operatorname{Rem}_2(tH)-\operatorname{Rem}_2(tH_i)}{t^2}$$

$$=\frac{L(Q(x)+tH(x))-L(Q(x)+tH_i(x))}{t^2}$$

$$-\frac{L'(Q(x))[H(x)-H_i(x)]}{t}-\frac{1}{2}L''(Q(x))[H^2(x)-H_i^2(x)]$$

$$=\frac{L'(\eta)-L'(Q(x))}{t}[H(x)-H_i(x)]-\frac{1}{2}L''(Q(x))[H^2(x)-H_i^2(x)],$$

where η is between $Q(x)+tH(x)$ and $Q(x)+tH_i(x)$. As above, by Lemma 5.4.3 of Fernholz (1983), we have

$$\left|\frac{L'(\eta)-L'(Q(x))}{t}\right|\leqslant M\,|H(x)-H_i(x)|.$$

Hence, by (6.8), we have

$$\inf_{1\leqslant i\leqslant n}\left\|\frac{\operatorname{Rem}_2(tH)}{t^2}-\frac{\operatorname{Rem}_2(tH_i)}{t^2}\right\|_{L^p}\leqslant M_\Gamma\varepsilon, \tag{6.11}$$

where M_Γ is a constant which depends on Γ. Therefore, (6.7) follows from (6.9) through (6.11). ∎

Proof of Proposition 3.3. It suffices to show that

$$\lim_{t\to 0}\frac{\gamma(G_0+tG_n,H_0+tH_n)-\gamma(G_0,H_0)-\gamma'_{(G_0,H_0)}(tG_n,tH_n)-\frac{1}{2}\gamma''_{(G_0,H_0)}(tG_n,tH_n)}{t^2}=0,$$

for any $G_n\to G$, $H_n\to H$, as $n\to\infty$, where $G_n, G, H_n, H\in D[0,1]$. Note that

$$\frac{1}{t^2}\{\gamma(G_0+tG_n,H_0+tH_n)-\gamma(G_0,H_0)$$

$$-\gamma'_{(G_0,H_0)}(tG_n,tH_n)-\tfrac{1}{2}\gamma''_{(G_0,H_0)}(tG_n,tH_n)\}$$

$$=\frac{G_n[\phi(H_0+tH_n)-\phi(H_0)]}{t}-\phi'(H_0)\,G_nH_n$$

$$+G_0\frac{\phi(H_0+tH_n)-\phi(H_0)-\phi'(H_0)\,tH_n-\frac{1}{2}\phi''(H_0)\,t^2H_n^2}{t^2}$$

$$=[\phi'(\xi)-\phi(H_0)]\,G_nH_n+\tfrac{1}{2}[\phi''(\eta)-\phi''(H_0)]\,G_0H_n^2, \tag{6.12}$$

where ξ and η are between H_0 and H_0+tH_n. Since ϕ'' is continuous, the proof follows from (6.12). ∎

Proof of Proposition 3.5. It suffices to show that

$$\lim_{n\to\infty}\frac{\mathrm{Rem}_2(t_nG_n,\, t_nH_n;\gamma)}{t_n^2}=0,$$

for any $G_n\to G$, $H_n\to H$, $t_n\to 0$, as $n\to\infty$, where G_n, G, H_n, $H\in D[0,1]$ such that $(G_0+t_nG_n,\, H_0+t_nH_n)\in E$. Note that

$$\gamma(G_0+t_nG_n,\, H_0+t_nH_n)-\gamma(G_0,H_0)$$
$$-t_n\gamma'_{(G_0,H_0)}(G_n,H_n)-\tfrac{1}{2}t_n^2\gamma''_{(G_0,H_0)}(G_n,H_n)=t_n^3\int_0^1 G_n^2(x)\,dH_n(t).$$

Hence,

$$\frac{\mathrm{Rem}_2(t_nG_n,\, t_nH_n;\gamma)}{t_n^2}=t_n\int_0^1 G_n^2(x)\,dH_n(t).$$

The proof follows from the one of Lemma 3 by Gill (1989). ∎

Proof of Theorem 3.6. The proof of (3.11) is similar to that of Theorem 3.1 by Ren and Sen (1991). We sketch the idea of the proof as below.

Let $\tilde{U}_n^*$ and $\tilde{U}_n$ be the continuous version of U_n^* and U_n, respectively, with

$$\|\tilde{U}_n^*-U_n^*\|\leqslant C_n^{-1}\max_{1\leqslant i\leqslant n}|c_{ni}|,\qquad \text{a.s.}\tag{6.13}$$

and

$$\|\tilde{U}_n-U_n\|\leqslant (n+1)^{-1},\qquad \text{a.s.}\tag{6.14}$$

Since $C_n[U_n^*-E\{U_n^*\}]$ and $\sqrt{n}[U_n-E\{U_n\}]$ weakly converge on $(D[0,1],\|\cdot\|)$ (see Shorack and Wellner, 1986, p. 109), by (A1), we have that $W_n=C_n[U_n^*-U^*]$ and $V_n=\sqrt{n}[U_n-\bar{U}]$ also weakly converge on $(D[0,1],\|\cdot\|)$. If we denote $Z_n=C_n(\tilde{U}_n^*-U^*,\, \tilde{U}_n-\bar{U})$, we easily see that $Z_n\in C[0,1]\times C[0,1]$. Note that

$$Z_n=C_n(\tilde{U}_n^*-U_n^*,\, \tilde{U}_n-U_n)+C_n(U_n^*-U^*,\, U_n-\bar{U}).\tag{6.15}$$

Hence, from (A2) and (6.13)–(6.15), we know that $C_n[\tilde{U}_n^* - U^*]$ and $C_n[\tilde{U}_n - \bar{U}]$ are relatively compact on $C[0, 1]$. Since, $C[0, 1]$ is complete and separable, by Prohorov's Theorem (Billingsley, 1968), we have that for every $\varepsilon > 0$, there exist compact sets K_1 and K_2 in $C(0, 1]$ such that

$$P\{Z_n \in K\} > 1 - \varepsilon, \qquad \text{all} \quad n \geqslant 1, \tag{6.16}$$

where $K = K_1 \times K_2$ is a compact set in $C[0, 1] \times C[0, 1]$. For $Q(G, H, t) = \text{Rem}(tG, tH; \tau)/t$ with $G, H \in D[0, 1]$, the rest of the proof follows along the lines of the proof of Theorem 3.1 by Ren and Sen (1991), and hence is omitted.

The proof of (3.13) follows similarly by using $Q(G, H, t) = \text{Rem}_2(tG, tH; \tau)/t^2$. ∎

Proof of Lemma 4.1. We will only prove (ii) since the proof of (i) is quite similar. Note that the functional τ can be expressed as a composition of the following second-order Hadamard differentiable transformations:

$\gamma_1 \colon D[0, 1] \times D[0, 1] \to D[0, 1] \times D[0, 1]$ defined by $\gamma_1(G, H) = (G\phi(H), H)$, is, by Proposition 3.3, second-order Hadamard differentiable at (G_0, H_0) with derivatives

$$\gamma'_{1(G_0, H_0)}(G, H) = (G_0 \phi'(H_0)\, H + \phi(H_0)\, G, H),$$

and

$$\gamma''_{1(G_0, H_0)}(G, H) = (2\phi'(H_0)\, GH + G_0 \phi''(H_0)\, H^2, 0),$$

where $G, H \in D[0, 1]$;

$\gamma_2 \colon D[0, 1] \times D[0, 1] \to \mathbb{R}$ defined by $\gamma_2(G, H) = \int_0^1 G^2\, dH$, is, by Proposition 3.5, second-order Hadamard differentiable at $(G_0 \phi(H_0), H_0)$ with derivatives

$$\gamma'_{2(G_0\phi(H_0), H_0)}(G, H) = 2 \int_0^1 G_0 \phi(H_0)\, G\, dH_0 + \int_0^1 G_0^2 \phi^2(H_0)\, dH,$$

and

$$\gamma''_{2(G_0\phi(H_0), H_0)}(G, H) = 2 \int_0^1 G^2\, dH_0 + 4 \int_0^1 G_0 \phi(H_0)\, G\, dH,$$

where $G, H \in D[0, 1]$.

We have $\tau = \gamma_2 \circ \gamma_1$. Hence, by Proposition 3.1. τ is second-order Hadamard differentiable at (G_0, H_0) with derivatives given by (4.2) and (4.3). ∎

Proof of Theorem 4.3. From Corollary 4.2 and Theorem 3.6, we immediately have (4.7)–(4.9). For W_n given by (4.13), we have

$$E\{W_n(t)\} = 0,$$

$$E\{W_n(s)\, W_n(t)\} = \min\{s, t\} - st,$$

because we always have (4.6) and

$$\sum_{i=1}^{n} c_{ni}^2 = 1, \qquad n \geqslant 1. \tag{6.17}$$

For the special construction (Shorack and Wellner, 1986, p. 93), by Theorem 3.7.1 of Shorack and Wellner (1986, p. 140), we have

$$\begin{aligned}
&\left| \int_0^1 W_n^2 \psi(U)\, dU - \int_0^1 W^2 \psi(U)\, dU \right| \\
&\quad \leqslant \int_0^1 |W_n^2 - W^2|\, \psi(U)\, dU \\
&\quad \leqslant \|(W_n^2 - W^2)/(U(1-U))^{1/4}\| \int_0^1 \psi(U)(U(1-U))^{1/4}\, dU \\
&\quad \leqslant \|(W_n - W)/(U(1-U))^{1/4}\|\, \{\|W_n - W\| + 2\, \|W\|\} \\
&\qquad \times \int_0^1 \psi(U)(U(1-U))^{1/4}\, dU \\
&\quad = o_p(1)(o_p(1) + O_p(1)) \int_0^1 \psi(U)(U(1-U))^{1/4}\, dU = o_p(1),
\end{aligned} \tag{6.18}$$

where W is a Brownian bridge. Since for the special construction and W_n given in (4.13), the distributions of $\int_0^1 W_n^2 \psi(U)\, dU$ are the same, we have that

$$C_n^2 \tau(U_n^*, D_n) \xrightarrow{D} \int_0^1 W^2(t)\, \psi(t)\, dt, \qquad \text{as} \quad n \to \infty. \tag{6.19}$$

For the decomposition of $\int_0^1 W^2(t)\,\psi(t)\,dt$ in (4.10), one may see Anderson and Darling (1952) or Shorack and Wellner (1986, Chap. 5). ▮

Proof of Lemma 4.4. Recall that we have

$$U_n(t)=(n+1)^{-1}\sum_{i=1}^{n} I\{Y_i\leqslant t\},$$

where Y_i's are as those in (4.13). We notice that for the special construction (Shorack and Wellner, 1986, p. 93) and for our W_n, U_n, the distributions of $\int_0^\delta W_n^2\psi(U_n)\,dU_n$, $\int_{1-\delta}^1 W_n^2\psi(U_n)\,dU_n$, $\int_0^\delta W_n^2\psi(U)\,dU$ and $\int_{1-\delta}^1 W_n^2\psi(U)\,dU$ are the same, respectively. Hence, it suffices to establish (4.14)–(4.17) for the special construction, which will be also denoted by W_n and U_n.

Let $Y_{(1)}, ..., Y_{(n)}$ be the order statistics of $Y_1, ..., Y_n$. Note that for any $\delta>0$,

$$\begin{aligned}\int_0^\delta W_n^2\psi(U_n)\,dU_n &\leqslant M_1\int_0^\delta w_n^2/U_n\,dU_n\\ &=M_1(n+1)^{-1}\sum_{i=1}^{n}\frac{W_n^2(Y_{(i)})}{U_n(Y_{(i)})}I\{Y_{(i)}\leqslant\delta\}\\ &=M_1(n+1)^{-1}\sum_{i=1}^{n}\frac{W_n^2(Y_{(i)})}{i/(n+1)}I\{Y_{(i)}\leqslant\delta\},\end{aligned}\tag{6.20}$$

and

$$\begin{aligned}&(n+1)^{-1}\sum_{i=1}^{n}\frac{W_n^2(Y_{(i)})}{i/(n+1)}I\{Y_{(i)}\leqslant\delta\}\\ &\quad=(n+1)^{-1}\sum_{i=1}^{n}\frac{W_n^2(Y_{(i)})-W^2(Y_{(i)})}{i/(n+1)}I\{Y_{(i)}\leqslant\delta\}\\ &\qquad+(n+1)^{-1}\sum_{i=1}^{n}\frac{W^2(Y_{(i)})}{i/(n+1)}I\{Y_{(i)}\leqslant\delta\},\end{aligned}\tag{6.21}$$

where W is a Brownian bridge. From Theorem 3.7.1 of Shorack and Wellner (1986, p. 140), we know that

$$\begin{aligned}&\|(W_n^2-W^2)/(U(1-U))^{1/4}\|\\ &\quad\leqslant\|(W_n-W)/(U(1-U))^{1/4}\|\,\{\|W_n-W\|+2\,\|W\|\}\\ &\quad=o_p(1)(o_p(1)+O_p(1))=o_p(1).\end{aligned}\tag{6.22}$$

Hence, by $\sqrt{n}\,\|U_n - U\| = O_p(1)$, we have

$$(n+1)^{-1}\sum_{i=1}^{n}\frac{|W_n^2(Y_{(i)}) - W^2(Y_{(i)})|}{i/(n+1)} I\{Y_{(i)} \leqslant \delta\}$$

$$\leqslant (n+1)^{-1}\sum_{i=1}^{n}\left|\frac{W_n^2(Y_{(i)}) - W^2(Y_{(i)})}{[Y_{(i)}(1-Y_{(i)})]^{1/4}}\right|\frac{(Y_{(i)})^{1/4}}{i/(n+1)} I\{Y_{(i)} \leqslant \delta\}$$

$$\leqslant \|(W_n^2 - W^2)/[U(1-U)]^{1/4}\|$$

$$\times\left\{(n+1)^{-1}\sum_{i=1}^{n}\frac{|Y_{(i)} - i/(n+1)|^{1/4} + [i/(n+1)]^{1/4}}{i/(n+1)}\right\}$$

$$\leqslant o_p(1)\left\{(n+1)^{-1}\sum_{i=1}^{n}\frac{\|U_n - U\|^{1/4} + [i/(n+1)]^{1/4}}{i/(n+1)}\right\}$$

$$\leqslant o_p(1)\left\{n^{-1/8}O_p(1)\sum_{i=1}^{n} 1/i + (n+1)^{-1}\sum_{i=1}^{n}[i/(n+1)]^{-3/4}\right\}$$

$$\leqslant o_p(1)\{n^{-1/8}O_p(1)\{\log n + C_0 + o(1)\} + O(1)\} = o_p(1), \qquad (6.23)$$

where C_0 is a constant. From Anderson and Darling (1952), we know that with probability 1,

$$W^2(t) \leqslant 2t(1-t)\log\log\frac{1-t}{t}, \qquad \text{for} \quad 0 < t < t_0, \qquad (6.24)$$

where $0 < t_0 < 1$. Since there exists $0 < B < \infty$ such that $\sqrt{x}\log\log\frac{1-x}{x} \leqslant B$ for $x \in (0, 1/2]$, we have

$$(n+1)^{-1}\sum_{i=1}^{n}\frac{W^2(Y_{(i)})}{i/(n+1)} I\{Y_{(i)} \leqslant \delta\}$$

$$\leqslant \frac{2}{(n+1)}\sum_{i=1}^{n}\frac{Y_{(i)}\log\log((1-Y_{(i)})/Y_{(i)})}{i/(n+1)} I\{Y_{(i)} \leqslant \delta\}$$

$$\leqslant \frac{2B}{(n+1)}\sum_{i=1}^{n}\frac{(Y_{(i)})^{1/2}}{i/(n+1)} I\{Y_{(i)} \leqslant \delta\}$$

$$\leqslant \frac{2B}{(n+1)}\sum_{i=1}^{n}\frac{|Y_{(i)} - i/(n+1)|^{1/2} + [i/(n+1)]^{1/2}}{i/(n+1)} I\{Y_{(i)} \leqslant \delta\}$$

$$= \frac{2B}{(n+1)}\sum_{i=1}^{n}\frac{n^{-1/2}O_p(1) + [i/(n+1)]^{1/2}}{i/(n+1)} I\{Y_{(i)} \leqslant \delta\}$$

$$\leqslant 2BO_p(1)\,n^{-1/2}\sum_{i=1}^{n}1/i+\frac{2B}{n+1}\sum_{i=1}^{n}\frac{I\{Y_{(i)}\leqslant\delta\}}{[i/(n+1)]^{1/2}}$$

$$=2BO_p(1)\frac{\log n+C_0+o(1)}{\sqrt{n}}+2B\sum_{i=1}^{n}I\{Y_{(i)}\leqslant\delta\}$$

$$\times\int_{(i-1)/(n+1)}^{i/(n+1)}\left(\frac{i}{n+1}\right)^{-1/2}dt$$

$$\leqslant o_p(1)+2B\sum_{i=1}^{n}I\{Y_{(i)}\leqslant\delta\}\int_{(i-1)/(n+1)}^{i/(n+1)}\frac{dt}{\sqrt{t}}$$

$$=o_p(1)+2B\int_0^{j/(n+1)}\frac{dt}{\sqrt{t}}, \tag{6.25}$$

where

$$j=\max\{i;\ Y_{(i)}\leqslant\delta\}$$

Since

$$Y_{(j)}\leqslant\delta<Y_{(j+1)},$$

and with probability 1,

$$|U_n(Y_{(j)})-Y_{(j)}|=|j/(n+1)-Y_{(j)}|\to 0, \qquad \text{as} \quad n\to\infty, \tag{6.26}$$

we have that for sufficiently large n,

$$\frac{1}{(n+1)}\sum_{i=1}^{n}\frac{W^2(Y_{(i)})}{i/(n+1)}I\{Y_{(i)}\leqslant\delta\}\leqslant o_p(1)+2B\int_0^{2\delta}\frac{dt}{\sqrt{t}}$$

$$=o_p(1)+4B\sqrt{2\delta}, \tag{6.27}$$

with probability 1. Therefore, (4.14) follows from (6.20), (6.21), (6.23) and (6.27).

From (6.22) and (6.24), we have

$$\int_0^{\delta}W_n^2\psi(U)\,dU\leqslant M_1\int_0^{\delta}\frac{W_n^2(t)}{t}\,dt$$

$$=M_1\int_0^{\delta}\frac{W_n^2(t)-W^2(t)}{t}\,dt+M_1\int_0^{\delta}\frac{W^2(t)}{t}\,dt$$

$$\leqslant o_p(1)\,M_1\int_0^{\delta}\frac{dt}{t^{3/4}}+2M_1\int_0^{\delta}\log\log\frac{1-t}{t}\,dt. \tag{6.28}$$

Therefore, (4.16) follows from (6.28) and the fact:

$$\int_0^{1/2} \log\log\frac{1-t}{t}\,dt < \infty. \tag{6.29}$$

Similarly, we can establish (4.15) and (4.17). ∎

Proof of Theorem 4.5. For any fixed $\delta > 0$, consider a functional τ_δ given by

$$\tau_\delta(G, H) = \int_\delta^{1-\delta} (G(x))^2\,\psi(H(x))\,dH(x), \qquad G, H \in D[0, 1]. \tag{6.30}$$

From Theorem 4.3, we have

$$C_n^2\tau_\delta(U_n^*, U_n) = C_n^2 \int_\delta^{1-\delta} U_n^{*2}\psi(U)\,dU + o_p(1), \qquad \text{as} \quad n \to \infty. \tag{6.31}$$

Hence, (4.18) follows from Lemma 4.4 and (6.31).

The proof of (4.19) follows from (6.18). For the decomposition of $\int_0^1 W^2(t)\,\psi(t)\,dt$ in (4.19), one may see Anderson and Darling (1952). ∎

REFERENCES

T. W. Anderson and D. A. Darling, Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes, *Ann. Math. Statist.* **23** (1952), 193–212.

V. I. Averbukh and O. G. Smolyanov, The theory of differentiation in linear topological spaces, *Russian Math. Survey* **22** (1967), 201–258.

P. Billingsley, "Convergence of Probability Measures," Wiley, New York, 1968.

H. T. David, A three-sample Kolmogorov–Smirnov test, *Ann. Math. Statist.* **29** (1958), 842–851.

L. T. Fernholz, "Von Mises Calculus for Statistical Functionals," Lecture Notes in Statistics, Vol. 19, Springer-Verlag, New York, [1.3f, 2.1b, 2.3a] 1983.

R. D. Gill, Non- and semi-parametric maximum likelihood estimators and the von Mises method-part 1, *Scand. J. Statist.* **16** (1989), 97–128.

J. Hájek, Extensions of the Kolmogorov–Smirnov tests to regression alternatives, *in* "Proc. Bernoulli-Bayes-Laplace Seminar" (L. LeCam, Ed.), pp. 45–60, University of California Press, Los Angeles, 1963.

J. Hájek and Z. Šidák, "Theory of Rank Tests," Academic Press, New York, 1967.

W. Hoeffding, A non-parametric test of independence, *Ann. Math. Statist.* **19** (1948), 546–557.

J. Jurečková and P. K. Sen, On adaptive scale-equivariant M-estimators in linear models, *Statist. Decision* (*Suppl.*) **1** (1984), 31–46.

H. H. Keller, "Differential Calculus in Local Convex Spaces," Lecture Notes in Mathematics, Vol. 417, Springer-Verlag, New York, 1974.

J. Kiefer, K-sample analogues of the Kolmogorov–Smirnov and Cramér–von Mises tests, *Ann. Math. Statist.* **30** (1959), 420–447.

J. A. Reeds, "On the Definition of Von Mises Functional," Ph.D. dissertation Harvard University, Cambridge, MA, 1976.

J. Ren, On Hadamard differentiability and its application to R-estimation in linear models, *Statist. Decisions* **12** (1994), 1–22.

J. Ren and P. K. Sen, On Hadamard differentiability of extended statistical functional, *J. Multivariate Anal.* **39** (1991), 30–43.

J. Ren and P. K. Sen, An asymptotic representation of regression M-estimators: Hadamard differentiability approaches, preprint (1993).

J. Ren and P. K. Sen, Asymptotic normality of regression M-estimators: Hadamard differentiability approaches, *in* "Asymptotic Statistics (Prague, 1993)," pp. 131–147, Contri. Statist. Physica, Heidelberg, 1994.

J. Ren and P. K. Sen, Hadamard differentiability on $D[0, 1]^p$, *J. Multivariate Anal.* **55** (1995), 14–28.

M. Rosenblatt, Limit theorems associated with variants of the von Mises statistics, *Ann. Math. Statist.* **23** (1952), 617–623.

F. W. Scholz and M. A. Stephens, K-sample Anderson-Darling tests, *J. Amer. Statist. Assoc.* **82**, 399 (1987), 918–924.

P. K. Sen, Functional jackknifing: relationality and general asymptotics, *Ann. Statist.* **16** (1988), 405–469.

G. R. Shorack and J. A. Wellner, "Empirical Processes with Applications to Statistics," Wiley, New York, 1986.

R. Von Mises, On the asymptotic distribution of differentiable statistical functions, *Ann. Math. Statist.* **18** (1947), 309–348.