# SECOND-ORDER ASYMPTOTIC RELATIONS AND GOODNESS-OF-FIT TESTS

by

**JANA JUREČKOVÁ**
**and**
**PRANAB K. SEN**

Department of Biostatistics
University of North Carolina

Institute of Statistics
Mimeo Series No. 2189

May 1998

# SECOND-ORDER ASYMPTOTIC RELATIONS AND GOODNESS-OF-FIT TESTS*†

JANA JUREČKOVÁ AND PRANAB KUMAR SEN‡
*Charles University, Prague, and University of North Carolina at Chapel Hill*

For goodness-of-fit tests for a model (distribution) admitting nuisance location or scale parameters, second-order asymptotic distributional representations are exploited for some robust estimators that are asymptotically first-order equivalent; their difference then provides a goodness-of-fit test criterion, whose asymptotic properties are studied under the null hypothesis as well as under local alternatives.

**1. Introduction.** Let $\{X_i;\ i \geq 1\}$ be a sequence of independent, identically distributed *(i.i.d.)* random variables with an absolutely continuous distribution function *(d.f.)* $F((x-\theta_1)/\theta_2)$, $x \in \mathbb{R}$ where $\theta_1 \in \mathbb{R}$, and $\theta_2 \in \mathbb{R}^+$ are unknown location and scale parameters; we let $\boldsymbol{\theta} = (\theta_1, \theta_2)$ and denote the parameter space by $\boldsymbol{\Theta}$. Let $F_0$ be a hypothetical *d.f.*, absolutely continuous with density $f_0$, symmetric around 0. We intend to test the null hypothesis of goodness-of-fit (GOF):

$$\mathbf{H}_0:\ F \equiv F_0 \quad \text{vs.} \quad \mathbf{K}: F \not\equiv F_0,\ \boldsymbol{\theta} \text{ nuisance.} \tag{1.1}$$

In the classical GOF testing problem with known $\boldsymbol{\theta}$, we may reduce the testing problem to the case of uniform $R(0,1)$ *d.f.* by means of the probability integral transformation; hence, the classical Kolmogorov-Smirnov, Cramér-von Mises and allied GOF tests, that are exact distribution-free (EDF), are applicable. The situation is more complex when $\boldsymbol{\theta}$ is nuisance. In some specific cases, such as the normal $F_0$, we could adapt a transformation on the $X_i$ eliminating $\boldsymbol{\theta}$ and then apply the classical GOF tests. For example, Durbin (1961) formed the residuals $(X_i - \bar{X}_n)/s_n$, $i = 1, \ldots, n$ by using the sample mean $\bar{X}_n$ and sample standard deviation $s_n$, characterized the spherical-uniformity property for normal $F_0$ and used these residuals for GOF testing. However, the distribution theory is more complex in such a case. Also for normal $F_0$, Shapiro and Wilk (1965) exploited the best linear unbiased estimators (BLUE) and, contrasting with $s_n^2$, they formulated a highly intuitive GOF test for normality.

Both of these tests seem to suffer from two drawbacks: (i) spherical uniformity or the specific distributional properties of the BLUE are are tied-down to the assumed normality of $F_0$; for non-normal $F_0$ we may need some other approach, and (ii) because these tests involve the highly nonrobust $s_n^2$, they are likely to be sensitive to the outliers and error contaminations (even to an infinitesimal extent).

Bhattacharyya and Sen (1977) considered an alternative approach to GOF testing, applicable not only for testing the normality but any $F_0$ admitting a minimal sufficient statistic for the associated

*Key words and phrases: Hadamard differentiability; logistic model; M-, L- and R-estimator; nuisance location and scale; one-step version; SOADR; spherical uniformity.

†AMS 1991 subject classification: 60G99, 62G10, 62G20

‡This work was supported by the US-Czech Collaborative Research Grant (NSF INT-9600518 and ES 046/1996); the first author's research was also supported by the Czech Republic Grant 201/96/0230

parameters; their test criterion is based on the difference $F_n - F_n^*$ of the empirical *d.f.* $F_n$ and its Rao-Blackwell version $F_n^*$. Though such test is exact (under $F_0$), its sampling distribution is usually quite complicated. In asymptotic considerations we may use the dominance of the distribution of a functional $\varphi(F_n - F_n^*)$ by $\varphi(F_n - F_0)$, but it may result in considerable conservativeness. Moreover, robustness aspects of such GOF tests merit careful examination.

There has been another line of attack for GOF tests: estimate $\boldsymbol{\theta}$ by $\hat{\boldsymbol{\theta}}$, (asymptotically) optimal under $F_0$, consider the *d.f.* $\hat{F}_n = F(.,\hat{\theta})$ and base a GOF test on $\hat{F}_n - F_n$. Unfortunately, the asymptotic representation for $n^{1/2}(\hat{F}_n - F_n)$, consists of two terms, one weakly converging to a Brownian bridge (under $F_0$), but the other one is generally a complex functional of a Gaussian function (and cannot be treated as a drift function, linear or not, even under $\mathbf{H}_0$). For a broad overview of such GOF tests, we may refer to Durbin (1973).

Motivated by these GOF tests (in the presence of nuisance parameters), and having the robustness aspects in mind, we advocate the use of robust $(L-, M-, R-)$ estimators of $\boldsymbol{\theta}$; we put special emphasis on the location parameter in this context, and a similar case can be worked out for scale parameter oriented GOF tests. We motivate our proposed testing procedure as follows. Let $T_1$, $T_2$ be two estimators of $\theta_1$, translation equivariant, such that each of them admits a second order asymptotic distributional representation (SOADR)

$$T_{nj} - \theta_1 = \frac{1}{n}\sum_{i=1}^{n} \psi_{j,F}(X_i - \theta_1) + \xi_{n,j,F} + o_p(n^{-1}),\ j = 1, 2, \tag{1.2}$$

where the score function $\psi_{j,F}$ generally depends on $F$ and is so standardized that $\mathbb{E}_{F_0}\psi_j(X) = 0$ and $\xi_{n,j,F}$ are the second order terms (typically $O_p(n^{-1})$), $j = 1, 2$. The $\xi_{n,j,F}$ are typically functionals reflecting quadratic variation of $F_n - F_0$, and hence they are natural candidates for GOF test statistics. Even if $T_1, T_2$ belong to different classes of estimators, they could be *first order asymptotically equivalent under* $F_0$ in the sense that

$$\psi_{1,F}(x) \equiv \psi_{2,F}(x) \quad \text{iff} \quad F \equiv F_0. \tag{1.3}$$

Or, in other words, the influence functions of $T_{n1}$ and $T_{n2}$ coincide at and only at $F_0$. First order asymptotically equivalent M-, L- and R-estimators were studied by a host of researchers and reported in Chapter 7 of Jurečková and Sen (1996). Generally, $\psi_{1,F}(x) \equiv \psi_{2,F}(x)$ holds in the family of logconcave densities if and only if $F \equiv F_0$. If $T_1$, $T_2$ are first order asymptotically equivalent under $F_0$, then we have at $F_0$

$$n(T_1 - T_2) = n(\xi_{n,1,F_0} - \xi_{n,2,F_0}) + o_p(1). \tag{1.4}$$

Then $n(T_1 - T_2)$, independent of the nuisance $\theta_1$, may be used as a test criterion. Whenever $T_1, T_2$ are scale-equivariant, $n(T_1 - T_2)$ may as well be studentized by any (robust) translation-invariant estimator of $\theta_2$. Under the null hypothesis, it has an asymptotic (non-normal) distribution. On the other hand, if $F \neq F_0$, then $\psi_{1,F}(x)$ and $\psi_{2,F}(x)$ are different and, by (1.2), $n^{1/2}(T_1 - T_2)$ has an asymptotic normal distribution; hence, $n(T_1 - T_2) = O_p(n^{1/2})$, what insures a high order consistency of GOF test based on (1.4) under non-local alternatives. Under contiguous alternatives, which in our case could be formulated in a reasonably broad (non- or semi-)parametric way, (1.2) leads to a nondegenerate non-null distribution of $n(T_1 - T_2)$.

The main tools used in the present paper are the SOADR of the type (1.2) supplemented with the asymptotic distribution of $\xi_{n,F}$. The SOADR results, often, could be simplified with the aid of the Hadamard expansions of the pertaining functionals, though SOADR is not restricted to only Hadamard differentiable functionals. Remark that *L-estimators* are location-scale equivariant and also Hadamard differentiable under fairly general regularity conditions. The *R-estimators* are location and scale equivariant and Hadamard differentiable for bounded score generating functions: this is the case of the $R$-estimator based on the Wilcoxon scores (viz., Hodges and Lehmann (1963), Sen (1963)), and we shall mostly apply this in our GOF tests. In Section 2, we incorporate the Wilcoxon score $R$-estimator and its dual $L$-estimator for constructing a GOF test statistic that can be used when $\boldsymbol{\theta}$ is nuisance (vector). Since both of these estimators are Hadamard differentiable, by appealing to the allied SOADR results on the difference of two Hadamard-differentiable functionals outlined in the Appendix, we are able to study the asymptotic null distribution.

The $M$-estimators of location are generally not scale-equivariant. Nevertheless, if we have only a nuisance location parameter in the formulation of $\mathbf{H}_0$, such $M$-estimators can be used to provide robust GOF tests. Led by this motivation, in Section 3 we consider the pair of $M$- and $L$-estimators, say $M_n$ and $L_n$, first-order asymptotically equivalent under $\mathbf{H}_0$. Our GOF test criterion is based on the difference $M_n^{(1)} - L_n$ where $M_n^{(1)}$ is the one-step version of $M_n$, starting at $L_n$. We shall derive asymptotic null distribution of the criterion, as well as the asymptotic distribution under the local alternatives.

**2. GOF tests based on Wilcoxon scores $R$-estimator and its dual $L$-estimator.** Let $X_1, \ldots, X_n$ be *i.i.d.* observations with the distribution function $F(x-\theta)$ where $F(x) + F(-x) = 1 \quad \forall x \in \mathbb{R}^1$, $F$ has an absolutely continuous density $f > 0$ and finite Fisher's information $I(F)$.

**I.** As an inversion of the Wilcoxon one-sample signed-rank test, we define the Wilcoxon score $R$-estimator $T_n$ of $\theta$ as the solution of the estimating equation

$$Q_n(t) = \int_{\mathbf{R}} [F_n(x) - F_n(2t-x)]dF_n(x) = 0 \tag{2.1}$$

where $F_n(x) = \frac{1}{n}\sum_{i=1}^n I[X_i \le x]$ is the empirical distribution function based on $X_1, \ldots, X_n$. The solution of (2.1) could be written in an explicit form

$$T_n = \text{med}\left\{\frac{X_i + X_j}{2} : 1 \le i \le j \le n\right\}. \tag{2.2}$$

The corresponding functional $T(F) = \theta$ may be defined implicitly as the root of the equation

$$\int_{\mathbf{R}} [F(x) - F(2t-x)]dF(x) = 0. \tag{2.3}$$

The following theorem presents the allied SOADR result; by virtue of location and scale equivariance, without loss of generality, we let $\theta_1 = 0, \theta_2 = 1$, for simplicity of notation.

**Theorem 2.1** *Let $X_1, \ldots, X_n$ be i.i.d. observations with the distribution function $F$ where $F(x) + F(-x) = 1$, $x \in \mathbb{R}$, and $F$ has two bounded derivatives $f$, $f'$. Then the R-estimator $T_n$ defined in (2.1) admits the SOADR :*

$$\begin{aligned} T_n &= \frac{1}{n\gamma}\sum_{i=1}^n \Big(F(X_i) - \frac{1}{2}\Big) + \frac{1}{n\gamma}\sum_{i=1}^n \Big(F(X_i) - \frac{1}{2}\Big)\frac{1}{n\gamma}\sum_{i=1}^n \Big(f(X_i) - \gamma\Big) \\ &\quad -\frac{1}{2\gamma}U_n^{(2)} + o_p(n^{-1}) \end{aligned} \tag{2.4}$$

*where*

$$\gamma = \int_{\mathbb{R}} f^2(x)dx \tag{2.5}$$

*and*

$$U_n^{(2)} = \frac{1}{n(n-1)} \sum_{1 \le i \ne j \le n} \{I(X_i + X_j \le 0) - F(-X_i) - F(-X_j) + \frac{1}{2}\}. \tag{2.6}$$

**Proof.** We shall use the notation $O_E(.)$ and $o_E(.)$ in the sense that $A_n = O_E(a_n)$ means $\mathbb{E}|A_n| = O(a_n)$ and $A_n = o_E(a_n)$ means $\mathbb{E}|A_n| = o(a_n)$ as $n \to \infty$. Denote

$$Y_n(u) = \int_{\mathbb{R}} \Big[F_n(x) - F_n(\frac{2u}{\sqrt{n}} - x)\Big] dF_n(x). \tag{2.7}$$

Then noting that $Y_n(\sqrt{n}T_n) = 0$, we have

$$Y_n(0) - Y_n(\sqrt{n}T_n) = Y_n(0). \tag{2.8}$$

Then we could write

$$\begin{aligned} Y_n(0) &= \int_{\mathbb{R}} [F_n(x) - F_n(-x)] dF_n(x) \\ &= n^{-2} \sum_{i=1}^n \sum_{j=1}^n \left(I[X_j \le X_i] - I[X_j \le -X_i]\right) \\ &= (2n)^{-1} - (1 - n^{-1})(U_n - \mathbb{E}U_n) - n^{-1}\big(n^{-1} \sum_{i=1}^n I(X_i \le 0) - \frac{1}{2}\big) \\ &= (2n)^{-1} - (U_n - \mathbb{E}U_n) + O_E(n^{-3/2}), \end{aligned} \tag{2.9}$$

where

$$U_n = \frac{1}{n(n-1)} \sum_{i=1}^n \sum_{j \ne i} I[X_i + X_j \le 0], \quad \mathbb{E}U_n = \frac{1}{2}. \tag{2.10}$$

Therefore, by the Hoeffding decomposition of $U$-statistics,

$$\begin{aligned} U_n - \tfrac{1}{2} &= \frac{2}{n} \sum_{i=1}^n \Big(\mathbb{E}\{I[X_i + X_j \le 0] | X_i\} - \frac{1}{2}\Big) \\ &+ \frac{1}{n(n-1)} \sum_{1 \le i \ne j \le n} \Big(I[X_i + X_j \le 0] - \mathbb{E}\{I[X_i + X_j \le 0] | X_i\} \\ &\qquad - \mathbb{E}\{I[X_i + X_j \le 0] | X_j\} + \frac{1}{2}\Big) \\ &= \frac{2}{n} \sum_{i=1}^n \big(F(-X_i) - \frac{1}{2}\big) \\ &+ \frac{1}{n(n-1)} \sum_{i \le j \ne \le n} \{I[X_i + X_j \le 0] - F(-X_i) - F(-X_j) + \frac{1}{2}\} \\ &= 2U_n^{(1)} + U_n^{(2)} \quad \text{(say)} \end{aligned} \tag{2.11}$$

where $\sqrt{n}U_n^{(1)} \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1/12)$ and $(n-1)U_n^{(2)} \xrightarrow{\mathcal{D}} \sum_{k=0}^{\infty} \lambda_k (Z_k^2 - 1)$ where $Z_k$'s are *i.i.d.* standard normal variables, and $\{\lambda_k\}$ is a sequence of eigenvalues of the functional $U_n^{(2)}(.) \in L_2(F)$ corresponding to pertaining orthonomal functions. Moreover,

$$
\begin{aligned}
Y_n(0) - Y_n(u) &= \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} I[X_i + X_j \le \frac{2u}{\sqrt{n}}] \\
&= \frac{1}{n^2}\Big(\sum_{i=1}^{n} I[0 < X_i \le \frac{u}{\sqrt{n}}] + \sum_{1\le i\ne j\le n} I[X_i + X_j \le \frac{2u}{\sqrt{n}}]\Big) \\
&= \frac{1}{n}[F_n(u/\sqrt{n}) - F_n(0)] + \frac{n-1}{n}U_n^*(u) \\
&= \frac{n-1}{n}U_n^*(u) + O_E(n^{-7/4}) \qquad (2.12)
\end{aligned}
$$

uniformly for $|u| \le K$, $0 < K < \infty$ any fixed positive number, where

$$
U_n^*(u) = \frac{1}{n(n-1)}\sum_{1\le i\ne j\le n} I[0 < X_i + X_j \le \frac{2u}{\sqrt{n}}], \qquad (2.13)
$$

and

$$
\begin{aligned}
\mathbb{E}U_n^*(u) &= \int_{\mathbb{R}}\Big[F(\frac{2u}{\sqrt{n}} - x) - F(-x)\Big]dF(x) \\
&= \frac{2u}{\sqrt{n}}\int_{\mathbb{R}} f(-x)dF(x) + \frac{2u^2}{n}\int_{\mathbb{R}} f'(-x)dF(x) + o_E(n^{-1}) \\
&= \frac{2u}{\sqrt{n}}\gamma + o_E(n^{-1}) \qquad (2.14)
\end{aligned}
$$

uniformly in $|u| \le K$. Again, the Hoeffding decomposition to $U_n^*(u)$ leads to

$$
U_n^*(u) - \mathbb{E}U_n^*(u) = 2U_n^{(1)}(u) + U_n^{(2)}(u) \qquad (2.15)
$$

where we write $U_n^{(1)}(u)$ as

$$
\begin{aligned}
&\frac{1}{n(n-1)}\sum_{1\le i\ne j\le n}\{\mathbb{E}(I[0 < X_i + X_j \le 2un^{-1/2}|X_i) - \mathbb{E}U_n^*(u)\} \\
&= \int_{\mathbb{R}}\Big[F\Big(\frac{2u}{\sqrt{n}} - x\Big) - F(-x)\Big]d(F_n(x) - F(x)) \\
&= \frac{2u}{\sqrt{n}}\int_{\mathbb{R}} f(-x)d(F_n(x) - F(x)) + \frac{2u^2}{n}\int_{\mathbb{R}} f'(-x)d(F_n(x) - F(x)),
\end{aligned}
$$

and hence,

$$
nU_n^{(1)}(u) = 2uZ_n^* + u^2 O_E(n^{-1/2}) + o_E(1) \qquad (2.16)
$$

as $n \to \infty$, uniformly in $|u| \le K$, where

$$
Z_n^* = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}(f(-X_i) - \gamma) \qquad (2.17)
$$

Similarly we could show that

$$
\mathbb{E}U_n^{(2)}(u) = 0, \quad \mathbb{E}(U_n^{(2)}(u))^2 = \frac{u^2}{n^3}O_E(1) \quad \forall u \in [-K, K]. \qquad (2.18)
$$

Combining (2.16), (2.17) and (2.15), we obtain

$$n^2\mathbb{E}\Big[U_n^*(u) - \frac{2u}{\sqrt{n}}\gamma\Big(1 + \frac{1}{\gamma\sqrt{n}}Z_n^*\Big)\Big]^2 = O\Big(\frac{u^2}{n}\Big) \text{ for } |u| \le K \tag{2.19}$$

and this implies, by Theorem 12.1 of Billingsley (1968), that

$$n \sup_{|u|\le K} \Big|U_n^*(u) - \frac{2u}{\sqrt{n}}\gamma\Big(1 + \frac{1}{\gamma\sqrt{n}}Z_n^*\Big)\Big| = O_p(n^{-1/2}), \tag{2.20}$$

hence

$$n \sup_{|u|\le K} \Big|\frac{n}{n-1}(Y_n(0) - Y_n(u)) - \frac{2u}{\sqrt{n}}\gamma\Big(1 + \frac{1}{\gamma\sqrt{n}}Z_n^*\Big)\Big| = o_p(1). \tag{2.21}$$

Inserting $u \to \sqrt{n}T_n$ into (2.21) and regarding (2.9) and (2.11) leads to the SOADR result:

$$\begin{aligned} T_n &= \frac{n}{n-1}(-U_n^{(1)} - \frac{1}{2}U_n^{(2)})\Big[\gamma\Big(1 + \frac{1}{\gamma\sqrt{n}}Z_n^*\Big)\Big]^{-1} + o_p(n^{-1}) \\ &= \frac{1}{n\gamma}\sum_{i=1}^n \Big(F(X_i) - \frac{1}{2}\Big) + \gamma^{-2}n^2\sum_{i=1}^n \Big(F(X_i) - \frac{1}{2}\Big)\sum_{j=1}^n \Big(f(-X_j) - \gamma\Big) \\ &\quad -\frac{1}{2\gamma}U_n^{(2)} + o_p(n^{-1}) \end{aligned} \tag{2.22}$$

□

**II.** Let $L_n$ be the $L$-estimator based on the weight function $J(t) = \gamma^{-1}f(F^{-1}(t))$, $\gamma = \int_{\mathbf{R}} f^2(x)dx)$, where under $\mathbf{H}_0$, $F \equiv F_0$, $J(t)$ is completely known. Then

$$L_n = \sum_{i=1}^n c_{ni}X_{n:i} \tag{2.23}$$

with

$$c_{ni} = \int_{(i-1)/n}^{i/n} J(t)dt,\ i = 1, \ldots, n, \tag{2.24}$$

and we may set

$$L_n = \int_{\mathbf{R}} xJ(F_n(x))dF_n(x). \tag{2.25}$$

$L_n$ admits a SOADR, characterized in the following theorem (where again, without loss of generality, we set $\theta_1 = 0, \theta_2 = 1$):

**Theorem 2.2** *Under the above conditions, $L_n$ admits the expansion*

$$L_n = \theta(F) + \bar{L}_n^{(1)} + \hat{L}_n^{(2)} + o_p(n^{-1}) \tag{2.26}$$

*where* $\theta(F) = \theta_1 (= 0)$,

$$\bar{L}_n^{(1)} = \frac{1}{n\gamma}\sum_{i=1}^n (F(X_i) - \frac{1}{2}) \tag{2.27}$$

*and*

$$\hat{L}_n^{(2)} = \frac{1}{2\gamma}\int_{\mathbf{R}} (-f'(x)/f(x))(F_n(x) - F(x))^2 dF(x). \tag{2.28}$$

*Moreover, as* $n \to \infty$,

$$\sqrt{n}\bar{L}_n^{(1)} \xrightarrow{\mathcal{D}} \mathcal{N}\Big(0, \frac{1}{12\gamma^2}\Big) \tag{2.29}$$

*and*

$$2n\gamma\hat{L}_n^{(2)}(\theta) \xrightarrow{\mathcal{D}} \int_0^1 \varphi(t)(B(t))^2 dt \tag{2.30}$$

*as* $n \to \infty$, *where* $B = \{B(t) : 0 \le t \le 1\}$ *is the Brownian bridge and*

$$\varphi(t) = \frac{-f'(F^{-1}(t)}{f(F^{-1}(t)}, \quad 0 < t < 1.$$

PROOF. For the sake of brevity, we let $\mathcal{B}_n(x) = F_n(x) - F(x),\ x \in \mathbb{R}$, and write

$$\begin{aligned} L_n &= \int_{\mathbb{R}} xJ(F_n(x))dF_n(x) \\ &= \int_{\mathbb{R}} xJ(F(x) + \mathcal{B}_n(x))d(F(x) + \mathcal{B}_n(x)) \end{aligned} \tag{2.31}$$

and this, after some arithmetics and using the special form of $J(.)$, leads to (2.26) with

$$\bar{L}_n^{(1)} = \frac{1}{\gamma}\int_{\mathbb{R}} F(x)d\mathcal{B}_n(x) = \frac{1}{n\gamma}\sum_{i=1}^n (F(X_i) - \frac{1}{2}) \tag{2.32}$$

and

$$\hat{L}_n^{(2)} = \frac{1}{2\gamma}\int_{\mathbb{R}} (-f'(x)(F_n(x) - F(x))^2 dx. \tag{2.33}$$

The remaining propositions follow from (2.32) and (2.33). □

Let's now consider the GOF testing problem based on the pair of statistics $(T_n, L_n)$, studied above. Let $X_1, \ldots, X_n$ be independent observations with d.f. $F(x; \boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is a nuisance parmeter (vector). Under the null hypothesis, $F \equiv F_0$, so for the $L$-estimator, we put $J(t) = \gamma_0^{-1} f_0(F_0^{-1}(t)),\ 0 < t < 1$ with $\gamma_0 = \int_{\mathbb{R}} f_0^2(x)dx$. Since both the estimators are location and scale equivariant, the distribution of $n(T_n - L_n)/\theta_2$ is independent of $\boldsymbol{\theta}$, and without loss of generality (WLOG), we put $\theta_1 = 0, \theta_2 = 1$. It follows from Theorems 2.1 and 2.2 that under $H_0$,

$$n(T_n - L_n) = \frac{1}{\gamma_0\sqrt{n}}\sum_{i=1}^n \Big(F_0(X_i) - \frac{1}{2}\Big).\frac{1}{\gamma_0\sqrt{n}}\sum_{i=1}^n \Big(f_0(X_i) - \gamma_0\Big) - \frac{n}{2\gamma_0}U_n^{(2)} - n\hat{L}_n^{(2)} + o_p(1) \tag{2.34}$$

where $U_n^{(2)}$ and $\hat{L}_n^{(2)}$ are defined as in before, but for the d.f. $F_0$. Since $\theta_2$ is unknown, we use any consistent estimator of a scale parameter (viz., the inter-quartile range), say $\hat{\theta}_{2,n}$, and consider the test statistic

$$\mathcal{Z}_n = n(T_n - L_n)/\hat{\theta}_{2,n} \tag{2.35}$$

which will have the same asymptotic distribution as of $n(T_n - L_n)/\theta_2$. Further note that by (2.34), the asymptotic null distribution of $\mathcal{Z}_n$ is nondegenerate, and hence the asymptotic critical levels are all $O(1)$. On the other hand, if $F \ne F_0$, $F$ symmetric, then the distribution of $n^{1/2}(T_n - L_n)$ will still be independent of $\theta_1$, and

$$n^{1/2}(T_n - L_n) = n^{-1/2}\gamma^{-1}\sum_{i=1}^n [F(X_i) - \frac{1}{2}] - n^{-1/2}\sum_{i=1}^n \psi(X_i) + o_p(1), \tag{2.36}$$

where

$$\psi(x) = -\int_{-\infty}^{\infty} J(F(y))\{I[x \le y] - F(y)\}dy, \quad x \in \mathbb{R}^1$$

and $J(t) = \gamma_0^{-1} f_0(F_0^{-1}(t)),\ 0 < t < 1$. Therefore, under such a fixed alternative, $\mathcal{Z}_n$ is $O_p(n^{1/2})$, so that the test based on $\mathcal{Z}_n$ will be consistent. When the unknown $F$ is not symmetric, the centering constants for $T_n$ and $L_n$ may not be the same, so that $\mathcal{Z}_n$ would be $O_p(n)$, and thereby, the consistency property would remain in tact. Granted this consistency property, we could consider some local alternatives and study the behavior of the proposed test. Toward this end, in the last section, we will consider some general types of contiguous alternatives that are not necessarily of the parametric type, and indicate the merits and demerits of the proposed GOF test relative to some other alternative ones proposed earlier.

In order to use the GOF test in practice, we need to have a handle over the null hypothesis distribution of $\mathcal{Z}_n$, at least in the asymptotic case. Though (2.34) can be used to study the asymptotic null distribution, it is not very simple. Recall that $U_N^{(2)}$ is a $U$-stastistic that is stationary of order 1, while $\hat{L}_n^{(2)}$ can also be expressed as a $U$-statistic (or more precisely a von Mises (1947) functional) that is also stationary of order 1; the first term on the right hand side of (2.34) is the product of two normal variables, and its distribution theory can be handled in a more manageable way. Therefore, to study the asymptotic null distribution of $\mathcal{Z}_n$, we need to study the asymptotic distributions of linear combinations of $U$-statistics that are first order stationary. This has been accomplished in the Appendix by reference to some existing SOADR results on statistical functionals that are Hadamard differentiable or differentiable in the von Mises (1947) sense. In that way, the results would apply to a bigger class of GOF test statistics that are characterizable in the same sense. Based on the advent of modern computational techniques and algorithms, we may consider the following simple simulation (Monte Carlo) procedures for estimating the (asymptotic) null distribution of $\mathcal{Z}_n$.

Generate $\{Y_1, \ldots, Y_n\}$ as independent copies of a r.v. from the hypothesized distribution $F_0$. Based on this set, compute the corresponding $T_n$, $L_n$, $\hat{\theta}_{2,n}$, and hence $\mathcal{Z}_N$. Repeat this random experiment $M$ times (with independent samples of size $n$), where $M$ is a large positive integer, say 500. Plot these simulated $\mathcal{Z}_n$ values, and obtain the corresponding empirical distribution. By the Glivenko-Cantelli lemma, this simulated empirical distribution converges almost surely (uniformly) to the true limiting null distribution of $\mathcal{Z}_n$, the existence of which is ensured by its SOADR representation in (2.34). This simulation technique also applies to other GOF test statistics that can be expressed in a form similar to $\mathcal{Z}_n$.

In the rest of this section, we illustrate the proposed GOF test for some specific $F_0$.

First, consider the classical Normal $F_0$. In this case, $\gamma = \int_{-\infty}^{\infty} f_0^2(x)dx = (2\sqrt{\pi})^{-1}$ and $f_0(x) = (2\pi)^{-1/2}\exp\{-\frac{1}{2}x^2\}$, so that for $L_n$ we can use the normal quantile densities as tabulated extensively in the literature. In this case, we can use the interquartile range for the estimation of $\theta_2$. Moreover note that $J(t) = f_0(F_0^{-1}(t))$ is bounded, and differentiable, so that all the regularity assumptions are satisfied here. To obtain the critical levels for $\mathcal{Z}_n$ (by simulation), all we need to generate random standard normal deviates and proceed as sketeched in the preceding paragraph.

Next consider the case of a logistic $F_0$, so that

$$F_0(x) = \{1 + \exp(-x)\}^{-1},\ x \in \mathbb{R}. \tag{2.37}$$

For this model, $f_0(x) = F_0(x)(1 - F_0(x))$, so that for $L_n$, we have a simple $J(t) = t(1-t),\ 0 \le t \le 1$, and $\gamma_0 = 1/6$. In addition, in this case, the Wilcoxon scores estimator $T_n$ is asymptotically optimal (for the location parameter $\theta_1$), so that we have a natural appeal for using this particular $R$-estimator for this model.

Consider the case of a Laplace $F_0$ where $f_0(x) = \frac{1}{2}\exp\{-|x|\},\ x \in \mathbb{R}$, so that $f_0(x) = F_0(x)$ or $1 - F_0(x)$ according as $x$ is negative or not. For $L_n$ we have a simple $J(t) = t$ or $1 - t$, according as $t \le 1/2$ or not, so that $\gamma_0 = 1/4$. For the Laplace distribution, an asymptotically optimal (location-scale equivariant) estimator of the location parameter is the sample median which is simultaneously an $L-, M-$ and $R$-estimator. So it might sound more appealing to use the sample median instead of the Wilcoxon scores estimator $T_n$. However, there is an awkward feature of the sample median: the SOADR for the sample median (or any quantile) involves the second-order term of the order $(n^{-3/4})$, not $(n^{-1})$ [viz., Jurečková and Sen, 1996].This slower rate of convergence will naturally affect the performance characteristics of the GOF tests that are based on them.

Finally, consider the case of the Cauchy $F_0$, where the density function $f_0$ is given by

$$f_0(x) = \pi^{-1}\{1 + x^2\}^{-1}, \ x \in \mathbb{R}. \tag{2.38}$$

This leads us to the quantile function $F_0^{-1}(t) = \tan\{\pi(2t-1)/2\}, \ t \in [(0,1)$, so that

$$f_0(F_0^{-1}(t)) = \{\pi \sec^2\{\pi(2t-1)/2\}\}^{-1}, \ t \in (0,1). \tag{2.39}$$

Therefore $J(t)$ is bounded and continuous, $\gamma_0 = (2\pi)^{-1}$, and we have no difficulty in constructing a GOF test statistic $\mathcal{Z}_n$ based on $T_n$, its dual $L_n$ and the interquartile range $\hat{\theta}_{2,n}$.

Remark that the Shapiro-Wilk (1965) type of GOF tests may require considerable modifications for $F_0$ other than a normal d.f., and its distribution problem may have similar complexities. In that sense, the proposed GOF tests have greater robustness, flexibility and adaptiveness for a larger class of $F_0$. A similar lack of robustness property is ascribable to the other GOF tests based on the estimated d.f. $\hat{F}_n$ or the Rao-Blackwell version, mentioned in Section 1. We shall make more comments on it from alternative hypotheses perspectives in Section 4. Actually, as we shall see later in Section 5, the choice of a suitable $R$-estimator may be linked to such plausible alternatives, particularly in the local case.

**3. GOF tests for the nuisance location parameter case.** Primarily guided by robustness considerations, we consider here some $M$-estimators (of $\theta_1$) that are not generally scale equivariant, and formulate suitable GOF tests for $H_0 : F(x) = F_0(x - \theta)$ where only the location parameter $\theta$ is treated as nuisance. Let $X_1, \ldots, X_n$ be *i.i.d.* observations with the distribution function $F(x-\theta)$ where $F(x) + F(-x) = 1 \ \ \forall x \in \mathbb{R}^1$, $F$ has an absolutely continuous density $f > 0$ and finite Fisher's information. Let $L_n$ be the $L$-estimator generated by a smooth weight function $J(u) : [0,1] \mapsto \mathbb{R}^1_+, \quad \int_0^1 J(u)du = 1, \quad J$ has an absolutely continuous derivative $J'$ in $(0,1)$, *i.e.*

$$L_n = \sum_{i=1}^{n} c_{ni} X_{n:i} \tag{3.1}$$

where $c_{ni}, \ i = 1, \ldots, n$ are generated by some $J_n : \ [0,1] \mapsto \mathbb{R}^1_+$ such that

$$\lim_{n\to\infty} J_n(t) = J(t), \ 0 < t < 1, \quad \left| \int_0^1 F^{-1}(t)\{J_n(\hat{G}_n(t)) - J(G_n(t))\} d\hat{G}_n(t) \right| = o_p(n^{-1}), \tag{3.2}$$

in the following way:

$$c_{ni} = \int_{(i-1)/n}^{i/n} J_n(t)dt, \quad i = 1, \ldots, n. \tag{3.3}$$

Note that $L_n$ can be equivalently expressed as

$$L_n = \int_0^1 F^{-1}(t) J_n(\hat{G}_n(t)) d\hat{G}_n(t) \tag{3.4}$$

where $\hat{G}_n(t) = F_n(F^{-1}(t)), t \in (0,1)$, and we may set

$$L_n = \int_0^1 F^{-1}(t)J(\hat{G}_n(t))d\hat{G}_n(t) + o_p(n^{-1}). \tag{3.5}$$

Then $L_n$ admits an asymptotic expansion, characterized in the following theorem:

**Theorem 3.3** *Under the above conditions, $L_n$ admits the expansion*

$$L_n = \theta(F) + \bar{L}_n^{(1)} + \hat{L}_n^{(2)} + o_p(n^{-1}) \tag{3.6}$$

*where*

$$\bar{L}_n^{(1)} = -\frac{1}{\sqrt{n}}\int_0^1 J(t)B_n(t)dF^{-1}(t) \tag{3.7}$$

*and*

$$\hat{L}_n^{(2)} = -\frac{1}{2n}\int_0^1 J'(t)(B_n(t))^2 dF^{-1}(t) \tag{3.8}$$

*where $B_n(t) = \sqrt{n}(\hat{G}_n(t) - t),\ 0 < t < 1$. Moreover, as $n \to \infty$,*

$$\sqrt{n}\bar{L}_n^{(1)} \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma_L^2); \quad \sigma_L^2 = 2\int\int_{0<s<t<1} J(s)J(t)s(1-t)dF^{-1}(s)dF^{-1}(t) \tag{3.9}$$

*and, provided $\hat{L}_n^{(2)} \in L_2(F)$, then*

$$-2n\hat{L}_n^{(2)} \xrightarrow{\mathcal{D}} \int_0^1 J'(t)[B(t)]^2 dF^{-1}(t) \tag{3.10}$$

*where $B = \{B(t) : 0 \le t \le 1\}$ is the Brownian bridge.*

PROOF. We write

$$\begin{aligned}
L_n &= \int_0^1 F^{-1}(t)J(t + \frac{1}{\sqrt{n}}B_n(t))d[t + \frac{1}{\sqrt{n}}B_n(t)] + o_p(n^{-1}) \\
&= \int_0^1 F^{-1}(t)J(t)dt + \int_0^1 F^{-1}(t)\{J(t + \frac{1}{\sqrt{n}}B_n(t)) - J(t)\}dt \\
&+ \frac{1}{\sqrt{n}}\int_0^1 F^{-1}(t)J(t + \frac{1}{\sqrt{n}}B_n(t))dB_n(t) + o_p(n^{-1}) \\
&= T(F) + \frac{1}{\sqrt{n}}\int_0^1 F^{-1}(t)J'(t)B_n(t)dt + \frac{1}{2n}\int_0^1 F^{-1}(t)(B_n)^2 J''(t)dt \\
&+ \frac{1}{\sqrt{n}}\int_0^1 F^{-1}(t)J(t)dW_n^0(t) + \frac{1}{n}\int_0^1 F^{-1}(t)B_n(t)J'(t)dB_n(t) + o_p(n^{-1}) \\
&= T(F) + \frac{1}{\sqrt{n}}\int_0^1 F^{-1}(t)d[J(t)B_n(t)] + \frac{1}{2n}\int_0^1 F^{-1}(t)d[J'(t)(B_n(t))^2] + o_p(n^{-1}) \\
&= T(F) - \frac{1}{\sqrt{n}}\int_0^1 J(t)B_n(t)dF^{-1}(t) - \frac{1}{2n}\int_0^1 J'(t)(B_n(t))^2 dF^{-1}(t) + o_p(n^{-1}) \\
&= T(F) + \bar{L}_n^{(1)} + \hat{L}_n^{(2)} + o_p(n^{-1}) \qquad (3.11)
\end{aligned}$$

with $\bar{L}_n^{(1)}$ and $\hat{L}_n^{(2)}$ given in (3.7) and (3.8), respectively. This, in turn, implies (3.9) and (3.10). □

Alternatively, we could express (3.6) in the following way:

$$L_n = \theta + \frac{1}{n}\sum_{i=1}^{n}\phi(X_i - \theta) + \hat{L}_n^{(2)} + o_p(n^{-1}) \tag{3.12}$$

where

$$\phi(x) = -\int_{-\infty}^{\infty} J(F(y))\{I[x \leq y] - F(y)\}dy, \quad x \in \mathbb{R}^1. \tag{3.13}$$

We shall refer to the representation (3.12) in the subsequent text.

Let us now proceed to the case of an $M$-estimator $M_n$ of $\theta$, generated by a smooth nondecreasing skew-symmetric score function $\psi$, $\psi(-x) = -\psi(x)$, as a solution of the equation

$$\sum_{i=1}^{n} \psi(X_i - t) = 0. \tag{3.14}$$

The authors proved in (1990) that $M_n$ admits SOADR:

$$M_n = \theta + \frac{1}{\gamma_1\sqrt{n}}U_{n2} + \frac{1}{\gamma_1 n}U_{n1}U_{n2} + o_p(n^{-1}) \tag{3.15}$$

where

$$U_{n1} = n^{-1/2}\sum_{i=1}^{n}\{\psi'(X_i) - \gamma_1\} = -\int_0^1 B_n(t)d\psi'(F^{-1}(t)), \tag{3.16}$$

$$U_{n2} = n^{-1/2}\sum_{i=1}^{n}\psi(X_i) = \int \psi(x-\theta)d(\sqrt{n}[\hat{F}_n(x) - F(x)] = -\int_0^1 B_n(t)d\psi(F^{-1}(t)) \tag{3.17}$$

and

$$\gamma_1 = \int_{-\infty}^{\infty} \psi'(x)dF(x) = \mathbb{E}_0\psi'(X_i). \tag{3.18}$$

where WOLG we put $\theta = 0$, and assume the following regularity conditions on $\psi$ and $F$ :

**(M1)** $0 < \mathbb{E}_0\psi^2(X_1) < \infty$.

**(M2)** $\psi$ has an absolutely continuous derivative, $0 < \gamma_1 < \infty$ and there exist $\delta > 0,\ K_1, K_2 > 0$ such that

$$\mathbb{E}_0|\psi'(X_1 - t)|^2 \leq K_1, \quad \mathbb{E}_0|\psi''(X_1 - t)|^2 \leq K_2 \quad \text{for} \quad |t| \leq \delta.$$

**(M3)** There exists a function $H(x)$ such that $\mathbb{E}_0 H(X_1) < \infty$ and

$$|\psi''(x-t) - \psi''(x)| \leq |t|^\alpha H(x) \quad \text{a.s.} \quad [F] \quad \text{for} \quad |t| \leq \delta$$

for some $\alpha > 0,\ \delta > 0$.

Assume that the function $\phi$ in (3.13) is proportional to $\psi$, more precisely.

$$\phi(x) = \gamma_1^{-1}\psi(x), \quad x \in \mathbb{R}^1 \tag{3.19}$$

and that the conditions (M1) - (M3) as well as the conditons of Theorem 3.1 are fullfilled. Then $\sqrt{n}(M_n - L_n) = o_p(1)$ and $n(M_n - L_n)$ has a nondegerate (nonnormal) asymptotic distribution, characterized in the folowing theorem:

**Theorem 3.4** *Let $L_n$ be an L-estimator (2.23) satisfying (3.2) and $M_n$ be an M-estimator defined as a solution of (3.14). Assume the conditions (M1) - (M3) and the conditions of Theorem 3.1. Then, under (3.19), as $n \to \infty$,*

$$\begin{aligned} n(M_n - L_n) &= \frac{1}{2\gamma_1}\int_{-\infty}^{\infty} \frac{\psi''(x)}{f(x)}[B(F(x))]^2 dx \\ &+ \frac{1}{\gamma_1}\int_{-\infty}^{\infty} \psi'(x)B(F(x))dx \int_{-\infty}^{\infty} \psi''(x)B(F(x))dx + o_p(1). \end{aligned} \tag{3.20}$$

*If (3.19) is not satisfied, then $\sqrt{n}(M_n - L_n)$ is asymptotically normally distributed $\mathcal{N}(0, \sigma^2)$ with*

$$\sigma^2 = \mathbb{E}_0\Big(\frac{\psi(X_1)}{\gamma_1} - \phi(X_1)\Big)^2. \tag{3.21}$$

PROOF. Under general $\phi$, the representations (3.12) and (3.15) imply that $\sqrt{n}(M_n - L_n) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2)$. Under (3.19), combining (3.15) and (3.12), we obtain

$$n(M_n - L_n) = \frac{\sqrt{n}}{\gamma_1}(U_{n2} - \gamma_1 U_{n3}) + \Big(\frac{1}{\gamma_1}U_{n1}U_{n2} - n\hat{L}_n^{(2)}\Big) + o_p(1) \tag{3.22}$$

where

$$U_{n3} = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\phi(X_i - \theta) = -\int_0^1 J(t)B_n(t)dF^{-1}(t). \tag{3.23}$$

Thus, under (3.19), $U_{n2} - \gamma_1 U_{n3} = 0$, while $n\hat{L}_n^{(2)}$ has the representation (3.8); this implies

$$n(M_n - L_n) = \frac{1}{\gamma_1}U_{n1}U_{n2} - n\hat{L}_n^{(2)} + o_p(1). \tag{3.24}$$

Actually, by (3.23) we obtain

$$U_{n2} - \gamma_1 U_{n3} = -\int_0^1 B_n(t)[d\psi(F^{-1}(t)) - \gamma_1 J(t)dF^{-1}(t)] \tag{3.25}$$

hence (3.19) is equivalent to

$$\gamma_1 J(t) \equiv \frac{d\psi(F^{-1}(t))}{dF^{-1}(t)} = \psi'(F^{-1}(t)). \tag{3.26}$$

Then, combining (3.16), (3.17) and (3.8), we get that, under (3.26),

$$\begin{aligned} U_{n1}U_{n2} &- n\gamma_1\hat{L}_n^{(2)} \\ &= \Big(\int_0^1 B_n(t)d\psi'(F^{-1}(t))\Big)\Big(\int_0^1 B_n(t)d\psi(F^{-1}(t))\Big) \\ &+ \frac{\gamma_1}{2}\int_0^1 (B_n(t))^2 J'(t)dF^{-1}(t) \\ &= \Big(\int_0^1 B_n(t)\psi'(F^{-1}(t))dF^{-1}(t)\Big)\Big(\int_0^1 B_n(t)\psi''(F^{-1}(t))dF^{-1}(t)\Big) \\ &+ \frac{1}{2}\int_0^1 (B_n(t))^2\frac{\psi''(F^{-1}(t))}{f(F^{-1}(t))}dF^{-1}(t) \\ &= \Big(\int_{-\infty}^{\infty} (B_n(F(x))\psi'(x)dx\Big)\Big(\int_{-\infty}^{\infty} B_n(F(x))\psi''(x)dx\Big) \\ &+ \frac{1}{2}\int_{-\infty}^{\infty} (B_n(F(x)))^2\frac{\psi''(x)}{f(x)}dx. \end{aligned} \tag{3.27}$$

Theorem 3.2 sparks an idea to use $n(M_n - L_n)$ as an test criterion for $\mathbf{H}_0$. Indeed, we shall describe such tests in the next section. However, while $L_n$ is computationally appealing. $M_n$ may be awkward. We eliminate this problem by replacing $M_n$ by its *one-step version* $M_n^{(1)}$ starting with $L_n$ in the role of the initial estimator.

The one-step $M$-estimator is defined as follows:

$$M_n^{(1)} = \begin{cases} M_n^{(0)} & \text{if } \hat{\gamma}_n = 0 \\ M_n^{(0)} + \frac{1}{n\hat{\gamma}_n}\sum_{i=1}^n \psi(X_i - M_n^{(0)}) & \text{if } \hat{\gamma}_n \neq 0 \end{cases} \tag{3.28}$$

where

$$\hat{\gamma}_n = \frac{1}{n}\sum_{i=1}^n \psi'(X_i - M_n^{(0)}). \tag{3.29}$$

Under the conditions (M1) - (M3),

$$\gamma_2 = \mathbb{E}_\theta \psi''(X_1 - \theta) = 0 \quad \text{and} \quad \mathbb{E}_\theta \psi''(X_1 - \theta - \eta) \to 0 \quad \text{when } \eta \to 0.$$

Then, by Theorem 2.2 of Jurečková and Sen (1990), letting (WLOG) $\theta = 0$,

$$n(M_n - M_n^{(1)}) = o_p(1), \tag{3.30}$$

hence

$$M_n^{(1)} = \frac{1}{\gamma_1\sqrt{n}}U_{n2} + \frac{1}{n\gamma_1}U_{n1}U_{n2} + o_p(n^{-1}). \tag{3.31}$$

Put $M_n^{(0)} - L_n$; then, using the representations (3.31), (3.12) and $U_{n3} = n^{-1/2}\sum_{i=1}^n \phi(X_i)$, we have

$$n(M_n^{(1)} - L_n) = \frac{\sqrt{n}}{\gamma_1}(U_{n2} - \gamma_1 U_{n3}) + (\frac{1}{\gamma_1}U_{n1}U_{n2} - n\hat{L}_n^{(2)}) + o_p(1). \tag{3.32}$$

If $\phi(x) \equiv \gamma_1^{-1}\psi(x)$, *i.e.*, the $M$-estimator and the initial $L$-estimator have the same influence functions, then $U_{n2} - \gamma_1 U_{n3} = 0$ and we get

$$n(M_n^{(1)} - L_n) = \frac{1}{\gamma_1}U_{n1}U_{n2} - n\hat{L}_n^{(2)} + o_p(1) \tag{3.33}$$

while $\hat{L}_n^{(2)}$ has the representation (3.8). This, in turn, implies, that $n(M_n^{(1)} - L_n)$ has the asymptotic representation (3.20), identical with the representation of $n(M_n - L_n)$.

Let now $F_0$ be a hypothetical distribution function, symmetric about 0 and we want to test the hypothesis

$$\mathbf{H}_1 : F \equiv F_0,\ \theta \text{ unspecified} \tag{3.34}$$

against the general alternative $\mathbf{K}_1 : F \not\equiv F_0$. We propose a class of tests of $\mathbf{H}_1$ based on the above relations of the M-estimator and of its one-step version described above. Generally, the performance of such a test consists in the following steps:

(i) Select a skew-symmetric function $\psi : \mathbb{R}^1 \mapsto \mathbb{R}^1$ such that $F_0$ and $\psi$ fulfill the conditions (M1) - (M3) of Section 3 and calculate the $L$-estimator $L_n$ of $\theta$ according to (3.4) with the coefficients $c_{ni},\ i = 1, \ldots, n$ generated by the weight function

$$J(t) = \gamma_1^{-1}\psi'(F_0^{-1}(t)),\ 0 < t < 1,\ \gamma_1 = \int_{-\infty}^{\infty} \psi'(x)dF_0(x) \tag{3.35}$$

according to (3.2) and (2.24).

(ii) Calculate the one-step $M$-estimator $M_n^{(1)}$ defined in (3.28), starting with $L_n$ as the initial estimator, *i.e.* $M_n^{(0)} = L_n$.

(iii) Calculate the *test criterion*, which is equal to

$$S_n = n(M_n^{(1)} - L_n) \tag{3.36}$$

and reject $\mathbf{H}_0$ provided

$$|S_n| \geq u_{\alpha/2} \tag{3.37}$$

where $u_{\alpha/2}$ is the $100(\alpha/2)\%$ critical value of the distribution of $|\mathcal{S}|$, and $\mathcal{S}$ is the functional

$$\begin{aligned} \mathcal{S} &= \frac{1}{2\gamma_1}\int_{-\infty}^{\infty} \frac{\psi''(x)}{f_0(x)}[B(F_0(x))]^2 dx \\ &+\frac{1}{\gamma_1}\int_{-\infty}^{\infty} \psi'(x)B(F_0(x))dx \int_{-\infty}^{\infty} \psi''(x)B(F_0(x))dx \end{aligned} \tag{3.38}$$

of the Brownian bridge $B(.)$.

Actually, the weak convergence $\mathcal{S}_n \xrightarrow{\mathcal{D}} \mathcal{S}$ under $\mathbf{H}_1$ follows from Theorem 3.2 and from the considerations after. While the percentiles of $|\mathcal{S}|$ cannot be directly calculated, they should be approximated either by the bootstrap or by the random walk.

On the other hand, if $F \neq F_0$ but both $L_n$ and $M_n^{(1)}$ admit the asymptotic representations, then

$$n^{-1/2}\mathcal{S}_n \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2(F_0, F)) \tag{3.39}$$

with

$$\sigma^2(F, F_0) = \int_{-\infty}^{\infty} (\gamma_1^{-2}\psi(x) - \phi(x))^2 dF(x) \tag{3.40}$$

and

$$\phi(x) = (\gamma_1)^{-1}\int_{-\infty}^{\infty} \psi'(F_0^{-1}(F(x))\{I[x \leq y] - F(y)\}dy. \tag{3.41}$$

Hence, the test based on $\mathcal{S}_n$ has asymptotic power 1 against every fixed alternative $F \neq F_0$.

Under the special choice of M-estimator generated by $\psi(x) = F_0(x) - \frac{1}{2}$, $x \in \mathbb{R}^1$, the criterion (3.36) partially simplifies. The corresponding asymptotically equivalent $L$-estimator is based on the weight function

$$J(t) = \gamma_1^{-1} f_0(F_0^{-1}(t)),\ 0 < t < 1,\ \gamma_1 = \int_{-\infty}^{\infty} f_0^2(x)dx \tag{3.42}$$

and (3.38) takes on the form

$$\mathcal{S} = \frac{1}{2\gamma_1}\int_0^1 [B(t)]^2 \frac{f_0'(F_0^{-1}(t))}{(f_0(F_0^{-1}(t)))^2}dt + \frac{1}{\gamma_1}\int_0^1 B(t)dt. \int_0^1 B(t)\frac{f_0'(F_0^{-1}(t))}{f_0(F_0^{-1}(t))}dt. \tag{3.43}$$

This special case when used for GOF testing for the logistic distribution with nuisance location parameter will be based on the criterion, relates to $\mathcal{S}_n \xrightarrow{\mathcal{D}} \mathcal{S}$ where

$$\mathcal{S} = 3\int_0^1 [B(t)]^2 \frac{(1-2t)}{t(1-t)}dt + 6\int_0^1 B(t)dt \left(\int_0^1 B(t)dt - 2\int_0^1 tB(t)dt\right).$$

**4. Contiguous alternatives and GOF tests.** Consider the family of alternatives to $\mathbf{H}_1$,

$$\mathbf{H}_n : f^{(n)}(x) = f_0(x)(1 + n^{-1/2}\lambda u(x)), \quad \lambda \text{ unspecified} \tag{4.1}$$

where $u(x)$ is a fixed function satisfying

$$\int_{-\infty}^{\infty} u(x) f_0(x) dx = 0, \quad \int_{-\infty}^{\infty} (u(x))^2 f_0(x) dx < \infty. \tag{4.2}$$

By the third LeCam's lemma, the sequence $\{\mathbf{H}_n\}$ is contiguous with respect to $\mathbf{H}_1$. The corresponding distribution function $F^{(n)}$ could be written in the form

$$F^{(n)}(x) = F_0(x) + n^{-1/2}\Lambda(x), \quad \Lambda(x) = \lambda \int_{-\infty}^{x} u(y) dF_0(y). \tag{4.3}$$

Notice that $\Lambda$ is bounded, $\Lambda(-\infty) = \Lambda(\infty) = 0$. It is also possible to consider general local nonparametric alternatives, satisfing contiguity; the following theorem describing the asymptotic behavior of the test criterion (3.36) under $\mathbf{H}_n$ could be easily extended for such alternatives as well.

**Theorem 4.1** *Assume the conditions of Theorem 3.2. Then, under* $\mathbf{H}_n$, *the test criterion* $\mathcal{S}_n = n(M_n^{(1)} - L_n)$, *described in steps (i) - (iii) of Section 3, converges in distribution to the functional* $\mathcal{S}^* = \mathcal{S} + \mathcal{Z}$, *where* $\mathcal{S}$ *is defined in (3.38) and*

$$\begin{aligned} \mathcal{Z} &= -\frac{1}{\gamma_1}\int_{-\infty}^{\infty} \psi(x) d\Lambda(x) \int_{-\infty}^{\infty} B(F_0(x))\psi''(x) dx \\ &- \frac{1}{\gamma_1}\int_{-\infty}^{\infty} \psi'(x) d\Lambda(x) \int_{-\infty}^{\infty} B(F_0(x))\psi'(x) dx \\ &+ \frac{1}{\gamma_1}\int_{-\infty}^{\infty} \psi(x) d\Lambda(x) \int_{-\infty}^{\infty} \psi'(x) d\Lambda(x). \end{aligned} \tag{4.4}$$

PROOF. WLOG, we put $\theta = 0$. Notice that

$$\gamma_1^{(n)} = \int_{-\infty}^{\infty} \psi'(x) f^{(n)}(x) dx = \gamma_1 + n^{-1/2}\lambda \int_{-\infty}^{\infty} \psi'(x) u(x) f_0(x) dx. \tag{4.5}$$

Moreover,

$$\sqrt{n}(\hat{F}_n(x) - F^{(n)}(x)) = B_n(F_0(x)) - n^{-1/2}\Lambda(x) \tag{4.6}$$

and hence

$$B(F^{(n)}(x)) = B(F_0(x)) + |\Lambda(x)|^{1/2} O_p(n^{-1/4}) \tag{4.7}$$

as $n \to \infty$ for the Brownian bridge $B$. By (3.8) and (3.35), under $\mathbf{H}_n$,

$$\begin{aligned} n\hat{L}_n^{(2)} &= -\frac{1}{2\gamma_1}\int_{-\infty}^{\infty} \frac{\psi''(x)}{f_0(x)} n(\hat{F}_n(x) - F^{(n)}(x))^2 dx \\ &\xrightarrow{\mathcal{D}} -\frac{1}{2\gamma_1}\int_{-\infty}^{\infty} \frac{\psi''(x)}{f_0(x)} [B(F_0(x))]^2 dx, \end{aligned} \tag{4.8}$$

as $n \to \infty$. Further, by (3.16) and (3.17), under $\mathbf{H}_n$, as $n \to \infty$,

$$\begin{aligned} U_{n1}U_{n2} = &\Big[\int_{-\infty}^{\infty} (\sqrt{n}(\hat{F}_n(x) - F^{(n)}(x))\psi''(x) dx - \int_{-\infty}^{\infty} \psi'(x) d\Lambda(x)\Big] \cdot \\ &\Big[\int_{-\infty}^{\infty} (\sqrt{n}(\hat{F}_n(x) - F^{(n)}(x))\psi'(x) dx - \int_{-\infty}^{\infty} \psi(x) d\Lambda(x)\Big] \\ \xrightarrow{\mathcal{D}} &\Big(\int_{-\infty}^{\infty} B(F_0(x))\psi''(x) dx - \int_{-\infty}^{\infty} \psi'(x) d\Lambda(x)\Big)\Big(\int_{-\infty}^{\infty} B(F_0(x))\psi'(x) dx - \int_{-\infty}^{\infty} \psi(x) d\Lambda(x)\Big). \end{aligned} \tag{4.9}$$

The asymptotic relation (3.24) remains true under the contiguous alternative; hence, combining (4.5) - (4.9), we arrive at the proposition of the theorem. □

Denote, for the sake of brevity,

$$\mathcal{X} = \frac{1}{2\gamma_1}\int_{-\infty}^{\infty} \frac{\psi''(x)}{f_0(x)}[B(F_0(x))]^2 dx, \quad \mathcal{Y} = \frac{1}{\gamma_1}\int_{-\infty}^{\infty} \psi'(x)B(F_0(x))dx,$$

$$\mathcal{W} = \int_{-\infty}^{\infty} \psi''(x)B(F_0(x))dx$$

and

$$a = \frac{1}{\gamma_1}\int_{-\infty}^{\infty} \psi(x)u(x)f_0(x)dx, \quad b = \int_{-\infty}^{\infty} \psi'(x)u(x)f_0(x).$$

Then the limiting functional under $\mathbf{H}_1$ is

$$\mathcal{S} = \mathcal{X} + \mathcal{Y}\mathcal{W}$$

while that under $\mathbf{H}_n$ is

$$\mathcal{S} + \mathcal{Z} = \mathcal{X} + (\mathcal{Y} - \lambda a)(\mathcal{W} - \lambda b).$$

We could easily verify that $\mathbb{E}\mathcal{X} = \mathbb{E}\mathcal{Y} = \mathbb{E}\mathcal{W} = 0$ and that $\mathcal{X}$, $\mathcal{Y}$ and $\mathcal{W}$ are uncorrelated.

The function $u(.)$ in (4.1) is typically either even or odd. For instance, the function $u(x) = -1 - x\frac{f_0'(x)}{f_0(x)}$, corresponding to the scale alternative, is even for symmetric $f_0$.

If $u$ is even, then $a = 0$, hence, denoting $\sigma^2 = \text{var}\ \mathcal{S}$,

$$\text{var}(\mathcal{S} + \mathcal{Z}) = \sigma^2 + \lambda^2 b^2 \mathbb{E}(\mathcal{Y}^2) \tag{4.10}$$

and

$$P(|\mathcal{S} + \mathcal{Z}| \geq u_{\alpha/2}) - \alpha = \lambda^2 K b^2 \mathbb{E}(\mathcal{Y}^2)(u_{\alpha/2})^2/(2\sigma^3) + o(\lambda^2) \quad \text{as} \quad \lambda \to 0 \tag{4.11}$$

where $K > 0$ is a constant dependent on the distribution of $\mathcal{S}$.

Similar conclusion we obtain in the case of odd $u$ (local alternative of skewness etc.), when $b = 0$.

**Appendix.** Let $X_1, \ldots, X_n$ be *i.i.d* random variable with a *d.f.* $F$ and let $\hat{F}_n$ be the corresponding sample *d.f.*. A natural estimator for a functional $T(F)$ is $T(\hat{F}_n)$. If $T$ is the second order differentiable at $F$ in the Hadamard sense (see Jurečková and Sen (1996) for details) then we could write an expansion

$$\begin{aligned} T(\hat{F}_n) &= T(F + (\hat{F}_n - F)) \\ &= T(F) + \frac{1}{n}\sum_{i=1}^{n} T^{(1)}(F; X_i) + \frac{1}{2n^2}\sum_i \sum_j T^{(2)}(F; X_i, X_j) + o_p(n^{-1}) \\ &= T(F) + \hat{T}_n^{(1)} + \hat{T}_n^{(2)} + o_p(n^{-1}). \end{aligned} \tag{5.1}$$

where

$$\int_0^1 \int_0^1 T^{(2)}(F; x, y)dF(x) = 0 = \int_0^1 \int_0^1 T^{(2)}(F; x, y)dF(y),$$

and

$$2n\hat{T}_n^{(2)} = n\int_0^1 T^{(2)}(F;F^{-1}(s),F^{-1}(t))d\hat{G}_n(s)d\hat{G}_n(t)$$

$$(5.2) \qquad = \int_0^1 T^{(2)}(F;F^{-1}(s),F^{-1}(t))dB_n(s)dB_n(t)$$

and, provided $T^2(.) \in L_2(F)$ (see Jurečková and Sen (1996), Chapter 4),

$$(5.3) \qquad 2n\hat{T}_n^{(2)} \xrightarrow{\mathcal{D}} \sum_{k=0}^{\infty} \lambda_k Z_k^2$$

where $\{Z_k\}_{k=0}^{\infty}$ are *i.i.d* standard normal random variables and $\{\lambda_k\}$ is a sequence of the eigenvalues of $T^{(2)}(.)$ with respect to orthonormal functions $\{\tau_k(.) : k \geq 0\}$ such that

$$\int T^{(2)}(F;x,y,)\tau_k(x)dF(x) = \lambda_k\tau_k(y) \quad a.s.\ [F],\ \forall k \geq 0$$

and

$$\int \tau_j(x)\tau_k(x)dF(x) = \delta_{jk}.$$

A similar expansion holds for $U$-statistics and related von Mises' functionals that need not be Hadamard differentiable.

Suppose now that $T_1(.)$ and $T_2(.)$ are two Hadamard differentiable functionals, estimating the same parameter (*i.e.*, $T_1(F) = T_2(F)$) and such that

$$(5.4) \qquad \bar{T}_{1n}^{(1)} \equiv \bar{T}_{2n}^{(1)} \quad \text{a.e.}$$

Under (1.3), $T_1(.)$ and $T_2(.)$ are *the first order asymptotically equivalent.* Then, denoting $T_0(.) = T_1(.) - T_2(.)$,

$$(5.5) \qquad n(T_1(\hat{F}_n) - T_2(\hat{F}_n)) = n(\hat{T}_{1n}^{(2)} - \hat{T}_{2n}^{(2)}) + o_p(1)$$

where

$$(5.6) \qquad 2n(\hat{T}_{1n}^{(2)} - \hat{T}_{2n}^{(2)}) = \int_0^1\int_0^1 T_0^{(2)}(F;F^{-1}(s),F^{-1}(t))dB_n(s)dB_n(t)$$

and

$$(5.7) \qquad T_0^{(2)}(F;x,y) = T_1^{(2)}(F;x,y) - T_2^{(2)}(F;x,y).$$

Then (5.6) has an asymptotic (nonnormal) distribution which follows from the asymptotic theory of Hoeffding's $U-$statistics. More precisely, we could write

$$(5.8) \qquad 2n(\hat{T}_{1n}^{(2)} - \hat{T}_{2n}^{(2)}) = \frac{1}{n}\sum_{i=1}^{n} T_0^{(2)}(F;X_i,X_i) + (n-1)U_{0n}^{(2)}$$

where

$$U_{0n} = \binom{n}{2}^{-1} \sum_{1\leq i<j\leq n} T_{0n}^{(2)}(F;X_i,X_j).$$

Then, by the Khintchine law of large numbers,

$$(5.9) \qquad \frac{1}{n}\sum_{i=1}^{n} T_0^{(2)}(F;X_i,X_i) \xrightarrow{\text{a.s}} \tau_0^{(2)}(F) = \int T_0^{(2)}(F;x,x)dF(x)$$

and

$$(n-1)U_{0n}^{(2)} \xrightarrow{\mathcal{D}} \sum_{k=0}^{\infty} \lambda_{0k}(Z_k^2 - 1). \tag{5.10}$$

## REFERENCES

Bhattacharyya, B. B. and Sen, P.K. (1977). Weak convergence of the Rao-Blackwell estimator of a distribution function. *Ann. Probability* 5, 500-510.

Billingsley, P. (1968). *Convergence of Probability Measures.* John Wiley, New York.

Durbin, J. (1961). Some methods of constructing exact tests. *Biometrika* 48, 41-55.

Durbin, J. (1973). *Distribution theory for Tests Based on the Sample distribution Function,* SIAM, Phiadelphia.

Hodges, J. L. and Lehmann, E.L. (1963). Estimates of location based on rank tests. *Ann. Math. Statist.* 34, 598-611.

Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. *Ann. Math. Statist.* 18, 293-325.

Jaeckel, L.A. (1971). Robust estimation of location: symmetric and asymmetric contamination. *Ann. Math. Statist.* 42, 1020-1034.

Jurečková, J. (1977). Asymptotic relations of $M$-estimates and $R$-estimates in linear regression models. *Ann. Statist.* 5, 664-672.

Jurečková, J. (1995). Trimmed mean and Huber's estimator: Their difference as a goodness-of-fit criterion. *J. Statist. Res.* 29, 31-35.

Jurečková, J. and Sen, P.K. (1990). Effect of the initial estimator on the asymptotic behavior of one-step $M$-estimator. *Ann. Inst. Statist. Math.* 42, 345-357.

Jurečková, J. and Sen P.K. (1996). *Robust Statistical Procedures: Asymptotics and Interrelations.* John Wiley, New York.

Sen, P. K. (1963). On the estimation of relative potency in dilution (-direct) assays by distribution-free methods. *Biometrics* 19, 532-552.

Sen, P.K. (1996). Statistical functionals, Hadamard differentiability and martingales. *Probability Models and Statistics. A J.Medhi Festschrift.* New Age Publishers, New Delhi, pp. 29-47.

Shapiro, S. S. and Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika* 52, 591-611.

Shorack, G. and Wellner, J. (1986). *Empirical Processes with Applications to Statistics.* John Wiley, New York.

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

*Charles University*
*Department of Probability and Statistics*
*Sokolovská 83*
*CZ-186 75 Prague 8*
*Czech Republic*
*jurecko@karlin.mff.cuni.cz*

*University of North Carolina*
*Departments of Biostatistics and Statistics,*
*3101 MacGavran-Greenberg Hall*
*Chapel Hill, NC 27599-7400*
*U.S.A.*
*pksen@bios.unc.edu*