# GOODNESS-OF-FIT TESTS WITH NUISANCE REGRESSION AND SCALE

by

JANA JUREČKOVÁ
and
PRANAB K. SEN

Department of Biostatistics
University of North Carolina

Institute of Statistics
Mimeo Series No. 2190

May 1998

# GOODNESS-OF-FIT TESTS WITH NUISANCE REGRESSION AND SCALE*†

By Jana Jurečková and Pranab Kumar Sen‡
*Charles University, Prague, and University of North Carolina at Chapel Hill*

Consider a linear model $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{E}$, where $\mathbf{X}$ is a known $(n \times p)$ matrix, $\boldsymbol{\beta}$ is a vector of unknown parameters, and the components of $\mathbf{E}$ are independent and identically distributed with an absolutely continuous distribution $F$. For a given symmetric and absolutely continuous distribution $F_0$, we want to test the goodness-of-fit hypothesis $\mathbf{H}_0$ : $F(z) \equiv F_0(z/\sigma)$, $\boldsymbol{\beta} \in \mathbb{R}^p$, $\sigma > 0$ unspecified, against the alternative $\mathbf{K}$ : $F(z) \neq F_0(z/\sigma)$. The proposed tests are based on regression invariant and scale equivariant regression rank scores scale statistics. Based on certain first-order representations, asymptotic distributional properties of the proposed tests are presented; both fixed-alternative consistency and contiguous-alternative asymptotic power properties are studied.

**1. Introduction.** Historically, goodness-of-fit (GOF) tests were developed first for categorical data models, and the Pearsonian goodness-of-fit test, having an asymptotic chi square distribution, is a precursor in this domain. For continuous random variables, many GOF tests are now available in the research literature; they include the conventional Kolmogorov-Smirnov and Cramér-von Mises type tests based on the empirical (sample) distribution function and its specified hypothetical population counterpart $F$, as well as the more complex case where the hypothetical d.f. is of known functional form but involves some unknown parameters. For a continuous $F$, the Kolmogorov-Smirnov and Cramér-von Mises tests are exactly distribution-free (under the hypothesis, the distribution of the test criterion does not depend on the actual form of $F$), so that the exact distribution theory under the null hypothesis can be analytically traced, and parallel asymptotics can be neatly presented. We may refer to Durbin (1973) and Hájek and Šidák (1967) for a broad account of exact distribution-free GOF tests and related asymptotics. The situation is more complex in the case of distributions admitting finite-dimensional nuisance parameters. In the special case of testing the normality, one could profit from some invariance stuctures. In a general case, with a nonnormal $F$, the tests may not be exact; moreover, the asymptotic properties of the tests may be highly complex. Led by the pioneering work of Durbin (1973), there has been a surge of development of GOF tests based on the empirical distribution function, comparing it with the hypothetical distribution function with the unknown parameters

---

*Key words and phrases: Contiguity; convex ordering; first order asymptotic representation; heavier tails; regression equivariance; regression quantiles; regression rank scores; scale equivariance; scale statistics.

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

‡This work has been supported by the US-Czech Collaborative Research Grant (NSF INT-9600518 and ES 046/1996). Besides that, the research of the first author was supported by the Czech Republic Grant 201/96/0230

replaced by efficient estimators, acquired from the sample. Applications of these tests may be based only on their asymptotic theory; but this (even under the null hypothesis) may not be very simple or tractable. For some developments along this line we refer to Khmaladze (1993) where other pertinent references are cited. Although the latter approach has added some extra theoretical layers of sophistications in this development, in terms of construction and (asymptotic or finite sample) distribution theory of GOF tests, there is very little insight that may be gained from this theory. Even asymptotic optimality properties of such GOF tests have not been explored to a satisfactory extent. A second approach, mainly advocated by the present authors (Jurečková (1995), Jurečková and Sen (1998)), exploits the second-order asymptotic distributional representations of a general class of estimators of location and scale parameters which are treated as nuisance parameters in their proposed GOF tests.

All these developments relate to the one-sample models where the basic observable random variables are treated as independent and identically distributed (*i.i.d.*). Additional complications arise when these observations are not necessarily identically distributed. This is the case which is treated in the present paper. Specifically, we are interested in GOF tests for a general linear model, treating the regression as well as the scale parameters as nuisance. Our proposed model includes the multi-sample models as special cases. We shall consider the conventional linear model:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{E};\ \mathbf{E} = (E_1, \ldots, E_n)', \tag{1.1}$$

with an $n \times p$ nonrandom regression matrix $\mathbf{X} = \mathbf{X_n}$ with an intercept (*i.e.*, $x_{i1} = 1,\ i = 1, \ldots, n$) and unknown $\boldsymbol{\beta} \in \mathbb{R}^p$. The errors $E_1, \ldots, E_n$ are supposed to be *i.i.d.* with an absolutely continuous distribution function $F$. We intend to test the null hypothesis

$$H_0 : F(z) \equiv F_0(z/\sigma),\ \boldsymbol{\beta},\ \sigma(>0) \text{ nuisance,} \tag{1.2}$$

with $F_0$ being a given symmetric, absolutely continuous distribution function, against the class of alternatives $K : F \neq F_0$, upto any $\sigma(>0)$. We propose a class of tests for $H_0$ against $K$ based on suitable regression-invariant and scale-equivariant scale statistics that are based on *regression rank scores*; these statistics were formulated by the present authors in (1994). In this setup, the proposed GOF tests are different in formulation from the conventional empirical distribution function tests, though the empirical distribution function of the residuals enters implicitly into the formulation of the regression rank scores. In this way, we are able to reduce the functional parameter space to a finite-dimensional one, and end up with a much simpler distribution theory under the null as well as alternative hypotheses. The Durbin-Khamaladze approach, based exclusively on independent, identically distributed random variables and devoted entirely to weak invariance principles or functional central limit theorems, may not be totally adoptable in the current situation; even if it could be employed principally, the associated distributional problems could be paramount, and simple test constructions would thereby be rather obstructive.

Note that the proposed model includes the one-sample location-scale model as a special case where $\mathbf{X}_n = \mathbf{1}_n$ (*i.e.*, $p = 1$). It also includes the two-sample location (nuisance scale) model as a special case where $\mathbf{X}_n$ is of order $n \times 2$ and $\mathbf{X}'_n = \left(\mathbf{1}'_n, \mathbf{x}'_2\right)$, and $\mathbf{x}_2$ has only the elements $-1$ and $+1$. In these special cases, our regression rank score scale statistics reduce to the rank scale statistics used by Jaeckel (1972) as measures of dispersion of residuals. To our knowledge, there has not been any previous work done in the goodness of fit testing problem, incorporating nuisance regression and scale parameters, even for the special one- or two-sample problem.

To achieve our goals, in the next section we introduce the preliminary notions and show the asymptotic properties of some scale statistics that are basic in construction of the GOF criterion. The tests are formulated in Section 3 along with their asymptotic null distribution and consistency against fixed alternatives. Section 4 is devoted to contiguous alternatives of a broad type that need not be of a specific parametric form as in the Durbin-Khamaladze approach. Otherwise speaking, the alternative $K$ could include also some nonparametric alternatives, as we shall illustrate in the same section.

**2. Preliminary notions and scale statistics.** The regression rank scores, developed by Gutenbrunner and Jurečková (1992), occupy a focal point in our developments. The regression rank scores for the model (1.1) are defined as the solution $\hat{\mathbf{a}}_n = (\hat{a}_{n1}, \ldots, \hat{a}_{nn})$ of the parametric linear programming problem

$$\begin{aligned} \mathbf{Y}'\hat{\mathbf{a}}(\alpha) : &= \max \\ \mathbf{X}'\hat{\mathbf{a}}(\alpha) &= (1-\alpha)\mathbf{X}'\mathbf{1}_n \\ \hat{\mathbf{a}}(\alpha) \in [0,1]^n, & \quad 0 < \alpha < 1. \end{aligned} \tag{2.1}$$

The current authors in (1994) defined a scale statistic, based on regression rank scores, which is regression invariant and scale equivariant, and hence convenient for a studentization and for all procedures requiring these properties. The statistic is defined in the following way:
Select a score function $\varphi : (0,1) \mapsto \mathbb{R}$, nondecreasing, non-constant, square-integrable and skew-symmetric on $(0,1)$, *i.e.*, $\varphi(1-u) = -\varphi(u), u \in (0,1)$. Denote

$$A^2(\varphi) = \int_0^1 \varphi^2(u)du. \tag{2.2}$$

Calculate the *regression scores* generated by $\varphi$ in the following way:

$$\hat{b}_{ni} = -\int_0^1 \varphi(u)d\hat{a}_{ni}(u), \quad i = 1, \ldots, n \tag{2.3}$$

and put

$$S_n = S_n(\mathbf{Y}) = n^{-1}\sum_{i=1}^{n} Y_i\hat{b}_{ni} = n^{-1}\mathbf{Y}'\hat{\mathbf{b}}_n. \tag{2.4}$$

Then, it follows from the definition of $\hat{\mathbf{a}}_n$ and $\hat{\mathbf{b}}_n$ (*see* (2.1) and (2.3)) that $S_n$ is *regression-invariant* and *scale-equivariant* in the sense that

$$\begin{aligned} S_n(\mathbf{Y}+\mathbf{X}\mathbf{d}) &= n^{-1}(\mathbf{Y}+\mathbf{X}\mathbf{d})'\hat{\mathbf{b}}_n(\mathbf{Y}+\mathbf{X}\mathbf{d}) \\ &= n^{-1}\mathbf{Y}'\hat{\mathbf{b}}_n \\ &= S_n(\mathbf{Y}) \quad \forall \mathbf{d} \in \mathbb{R}^p \end{aligned} \tag{2.5}$$

and

$$S_n(c\mathbf{Y}) = cS_n(\mathbf{Y}) \quad \forall c > 0. \tag{2.6}$$

This, among others, implies that $S_n$ is also *translation invariant*; moreover, because $\sum_{i=1}^n \hat{b}_{ni} = 0$, we see that $S_n > 0$ with probability 1.

In (1994), the authors studied the asymptotic behavior of $S_n$ for functions $\varphi$ satisfying

$$\varphi(u) = 0 \quad \text{if} \quad \begin{array}{ll} \text{either} & 0 < u < \alpha_0 \\ \text{or} & 1 - \alpha_0 < u < 1, \end{array} \qquad 0 < \alpha_0 < \frac{1}{2}$$

and, using the results of Gutenbrunner and Jurečková (1992), they showed that

$$S_n \xrightarrow{p} S(F) \quad \text{as } n \to \infty \tag{2.7}$$

where

$$S(F) = \int_0^1 \varphi(u) F^{-1}(u) du. \tag{2.8}$$

Moreover, $\sqrt{n}(S_n - S(F))$ was shown to be asymptotically normal $\mathcal{N}(0, \tau^2(\varphi, F))$ with

$$\tau^2(\varphi, F) = \int_0^1 \int_0^1 (u \wedge v - uv)\varphi(u)\varphi(v))dF^{-1}(u)dF^{-1}(v) \tag{2.9}$$

which also follows from the asymptotic representation

$$\begin{aligned} S_n - S(F) &= n^{-1} \sum_{i=1}^n \psi(E_i) + o_p(n^{-1/2}) \\ &= \int_0^1 \varphi(u)(u - F_n(F^{-1}(u))dF^{-1}(u) + o_p(n^{-1/2}) \end{aligned} \tag{2.10}$$

where

$$\psi(z) = \int_0^1 \frac{t - I[F(z) \le u]}{f(F^{-1}(u))} \varphi(u) du, \; z \in \mathbb{R}, \tag{2.11}$$

and where $F_n(z)$ is the empirical *d.f.* of $E_1, \ldots, E_n$.

The scale statistics of $S_n$ type will form a basis of goodness-of-fit criteria of the hypothesis

$$\mathbf{H}_0: \; F(z) \equiv F_0(z/\sigma), \quad \beta \in \mathbb{R}^p, \; \sigma > 0 \text{ unspecified.} \tag{2.12}$$

If the underlying *d.f.* has exponential tails under the hypothesis as well under the alternative, then $\varphi(u) = F_0^{-1}(u)$, $0 < u < 1$, is an appropriate choice of the score function, $F_0$ being the hypothetical distribution function. Hence, our first goal is to extend the propositions (2.7) - (2.11) to possibly unbounded score fucntions. The trimming of the $\varphi$-function will be necessary when either we test $\mathbf{H}_0$ for heavy-tailed $F_0$ or against heavy-tailed alternatives.

We shall treat the asymptotic behavior of the scale statistics under the following conditions on the model:

**(F1)** The distribution function $F$ of the errors has symmetric, positive, absolutely continuous density $f$ and finite Fisher's information $\mathcal{I}(f) = \int_{\mathbf{R}} (\frac{f'(x)}{f(x)})^2 f(x) dx < \infty$. $f$ has bounded derivative $f'$ for $|x| > K_f \ge 0$.

**(F2)** $f$ is monotonically decreasing to 0 as $x \to \pm\infty$ and

$$\lim_{x\to-\infty} \frac{-\log\ F(x)}{c|x|^r} = \lim_{x\to\infty} \frac{-\log\ (1-F(x))}{c|x|^r} = 1$$

for $c > 0,\ r > 0$.

**(X1)** $x_{i1} = 1,\ i = 1, \ldots, n$.

**(X2)** $\lim_{n\to\infty} \mathbf{D}_n = \mathbf{D}$ where $\mathbf{D}_n = n^{-1}\mathbf{X}_n'\mathbf{X}_n$ and $\mathbf{D}$ is a positive definite $p \times p$ matrix.

**(X3)** $\max_{1\le i\le n} \|\mathbf{x}_i\| = o(n^{1/4})$ and $n^{-1}\sum_{i=1}^n \|\mathbf{x}_i\|^4 = O(1)$ as $n \to \infty$ where $\mathbf{x}_i'$ is the $i$-th row of $\mathbf{X}_n,\ i = 1, \ldots, n$.

**($\Phi$1)** $\varphi(u),\ 0 < u < 1$ is a nondecreasing, square integrable function, $\varphi(1-u) = -\varphi(u),\ 0 < u < 1$, such that its derivative $\varphi'(u)$ exists and satisfies

$$|\varphi'(u)| \le C(u(1-u))^{-1-\delta},\ C > 0,\ 0 < \delta < \frac{1}{4}$$

for $0 < u < \alpha_0, \quad 1 - \alpha_0 < u < 1$ with given $\alpha_0 \in (0, \frac{1}{2})$.

Fix $b \in \mathbb{R},\ \frac{\delta}{2} < b < \frac{1}{4} - \frac{\delta}{2}$ for some $\delta \in (0, 1/4)$ and put

$$\alpha_n^* = n^{-\frac{1}{1+4b}} \quad \text{and} \quad \sigma_u = \frac{(u(1-u))^{1/2}}{f(F^{-1}(u))}, \quad 0 < u < 1. \tag{2.13}$$

**Theorem 2.1** *Under the conditions* **(F1)**, **(F2)**, **(X1)**- **(X3)**, **($\Phi$1)**, *the statistic* $S_n(\mathbf{Y})$ *admits the asymptotic representation*

$$\begin{aligned} S_n - S(F) &= n^{-1}\sum_{i=1}^n \psi(E_i) + o_p(n^{-1/2}) \\ &= \int_0^1 \varphi(u)(u - F_n(F^{-1}(u))dF^{-1}(u) + o_p(n^{-1/2}) \end{aligned} \tag{2.14}$$

*where* $\psi(z)$ *is given in (2.11),* $F_n(z)$ *is the empirical d.f. of* $E_1, \ldots, E_n$ *and*

$$S(F) = \int_0^1 \varphi(u)F^{-1}(u)du. \tag{2.15}$$

*Consequently,* $\sqrt{n}(S_n - S(F))$ *has an asymptotically normal distribution* $\mathcal{N}(0, \tau^2(\varphi, F))$ *with*

$$\tau^2(\varphi, F) = \int_0^1\int_0^1 (u \wedge v - uv)\varphi(u)\varphi(v))dF^{-1}(u)dF^{-1}(v). \tag{2.16}$$

PROOF. Let $\hat{a}_{ni}(u),\ i = 1, \ldots, n,\ 0 < u < 1$, be the regression rank scores corresponding to the model (1.1). Moreover, denote

$$E_{iu} = E_i - F^{-1}(u), \quad \tilde{a}_i(u) = I[E_i \geq F^{-1}(u)],\ i = 1, \ldots, n,\ 0 < u < 1. \tag{2.17}$$

The duality of regression rank scores and regression quantiles implies

$$\begin{aligned} \textstyle\sum_{i=1}^n Y_i(\hat{a}_{ni}(u) - (1-u)) &= \sum_{i=1}^n E_i(\hat{a}_{ni}(u) - (1-u)) \\ &= \textstyle\sum_{i=1}^n \rho_u(Y_i - \mathbf{x}_i'\hat{\beta}(u)), \qquad 0 < u < 1, \end{aligned} \tag{2.18}$$

where $\rho_u(x) = x(u - I[x < u]),\ x \in \mathbb{R},\ 0 < u < 1$ and $\hat{\beta}(u)$ is the $u$-regression quantile. By Lemma 3.1 and Theorem 3.1 in Gutenbrunner et al. (1993)

$$\begin{aligned} &\Big|(u(1-u))^{-1/2}\sigma_u^{-1}\sum_{i=1}^n \Big[\rho_u(Y_i - \mathbf{x}_i'\hat{\beta}(u)) - \rho_u(E_{iu})\Big] \\ &+ (u(1-u))^{-1/2}\sigma_u^{-1}(\hat{\beta}(u) - \beta(u))'\sum_{i=1}^n \mathbf{x}_i(u - I[E_{iu} < 0]) \\ &- \frac{1}{2}n\sigma_u^{-2}(\hat{\beta}(u) - \beta(u))'\mathbf{D_n}(\hat{\beta}(u) - \beta(u))\Big| = o_p(1) \end{aligned} \tag{2.19}$$

uniformly in $\alpha_n^* \leq u \leq 1 - \alpha_n^*$. Applying the uniform Bahadur representation of regression quantile (Theorem 3.1 in Gutenbrunner et al. (1993))

$$\begin{aligned} & n^{1/2}\sigma_u^{-1}(\hat{\beta}(u) - \beta(u)) \\ = \ & n^{-1/2}(u(1-u))^{-1/2}\mathbf{D}_n^{-1}\sum_{i=1}^n \mathbf{x}_i(u - I[E_{iu} < 0]) + \mathbf{o}_p(1) \end{aligned} \tag{2.20}$$

which holds uniformly in $\alpha_n^* \leq u \leq 1 - \alpha_n^*$, and inserting it in (2.19), we obtain

$$\begin{aligned} & \frac{f(F^{-1}(u))}{u(1-u)}\Big\{\sum_{i=1}^n E_i(\hat{a}_{ni}(u) - \tilde{a}_i(u)) - F^{-1}(u)n(F_n(F^{-1}(u)) - u)\Big\} \\ = \ & \frac{1}{u(1-u)}n^{-1}\sum_{i=1}^n\sum_{j=1}^n (u - I[E_{iu} < 0])(u - I[E_{ju} < 0])\mathbf{x}_i'\mathbf{D}_n^{-1}\mathbf{x}_j + o_p(1) \end{aligned} \tag{2.21}$$

uniformly in $\alpha_n^* \leq u \leq 1 - \alpha_n^*$, while, due to **(F2)**,

$$\frac{f(F^{-1}(u))}{u(1-u)} = O(-\log\ u)^{1-\frac{1}{r}}) \quad \text{as } u \to 0. \tag{2.22}$$

By virtue of (2.22), the approximation (2.21) simplifies as

$$\begin{aligned} A_n(u) &= n^{-1}\sum_{i=1}^n E_i(\hat{a}_{ni}(u) - \tilde{a}_i(u)) - F^{-1}(u)(F_n(F^{-1}(u)) - u) \\ &= \frac{1}{f(F^{-1}(u))}n^{-2}\sum_{i=1}^n\sum_{j=1}^n (u - I[E_{iu} < 0])(u - I[E_{ju} < 0])\mathbf{x}_i'\mathbf{D}_n^{-1}\mathbf{x}_j + o_p(n^{-1/2}) \end{aligned} \tag{2.23}$$

uniformly in $\alpha_n^* \leq u \leq 1-\alpha_n^*$. Integrating $-\varphi(u)$ with respect to the left-hand side of (2.23) and using the per partes integration we obtain

$$\begin{aligned} & S_n - S(F) - \int_0^1 \varphi(u)(F_n(F^{-1}(u)) - u)dF^{-1}(u) \\ (2.24) \quad &= n^{-1}\sum_{i=1}^n E_i \int_0^1 (\hat{a}_{ni}(u) - \tilde{a}_i(u))d\varphi(u) - \int_0^1 (F_n(F^{-1}(u)) - u)F^{-1}(u)d\varphi(u) \\ &= \int_0^1 A_n(u)d\varphi(u). \end{aligned}$$

Let us split the integration domain $(0,1)$ into five parts,

$$\begin{aligned} \mathcal{I}_{-2} &= (0, \alpha_n^*) \\ \mathcal{I}_{-1} &= (\alpha_n^*, \alpha_0) \\ \mathcal{I}_0 &= (\alpha_0, 1-\alpha_0) \\ \mathcal{I}_1 &= (1-\alpha_0, 1-\alpha_n^*) \\ \mathcal{I}_2 &= (1-\alpha_n^*, 1) \end{aligned}$$

and split the integral $\int_0^1 A_n(u)d\varphi(u)$ accordingly. First, by (2.23) and by **(X3)**,

$$\left|n^{1/2}\int_{\mathcal{I}_0} A_n(u)du\right| \leq Kn^{-3/2}\sum_{i=1}^n \|\mathbf{x}_i\|^2 = O(n^{-1/2})$$

where $0 < K < \infty$ is a constant. By virtue of **($\Phi$1)**, we could write for the second integral

$$\begin{aligned} & n^{1/2}\left|\int_{\mathcal{I}_{-1}} A_n(u)d\varphi(u)\right| \\ \leq\ & n^{1/2}\int_{\mathcal{I}_{-1}} c(u(1-u))^{-2-\delta}n^{-2}\left|\sum_{i=1}^n\sum_{j=1}^n (u - I[E_i < F^{-1}(u)])(u - I[E_j < F^{-1}(u)])h_{ij}\right|du \\ +\ & o_p(n^{-1/2}) \end{aligned}$$

where $\mathbf{H}_n = \left\|h_{ij}\right\|_{i,j=1}^n = \mathbf{X_n}(\mathbf{X}_n'\mathbf{X}_n)^{-1}\mathbf{X}_n'$ is the projection matrix. Dominating the quadratic form in the integrand by the square norm of the vector multiplied by the maximal eigenvalue of $\mathbf{H}_n$ (which is equal to 1), we obtain

$$\begin{aligned} & n^{1/2}\left|\int_{\mathcal{I}_{-1}} A_n(u)d\varphi(u)\right| \\ \leq\ & Kn^{-3/2}\int_{\alpha_n^*}^{\alpha_0}(u(1-u))^{-2-\delta}\sum_{i=1}^n (u - I[E_i < F^{-1}(u)])^2 du + o_p(n^{-1/2}) \\ =\ & Kn^{-1/2}\int_{\alpha_n^*}^{\alpha_0}(u(1-u))^{-2-\delta}|(F_n(F^{-1}(u)) - u)(1-2u) + u(1-u)|du + o_p(n^{-1/2}). \end{aligned}$$

Noting that $|(F_n(F^{-1}(u)) - u)(1-2u) + u(1-u)| = O_p(n^{-1/2}(u(1-u))^{1/2}) + O(u(1-u))$ as $u \downarrow 0$ and $n \to \infty$, we arrive at the conclusion that $n^{1/2}\int_{\mathcal{I}_{-1}} A_n(u)d\varphi(u) \to 0$ as $n \to \infty$. Similarly we obtain $n^{1/2}\int_{\mathcal{I}_1} A_n(u)d\varphi(u) \to 0$ as $n \to \infty$.

Finally, let us consider $n^{1/2}\int_{\mathcal{I}_{-2}} A_n(u)d\varphi(u)$. By (2.24), we could write

$$\begin{aligned}
&\left|n^{1/2}\int_{\mathcal{I}_{-2}} A_n(u)d\varphi(u)\right| \\
=& \left|n^{-1/2}\sum_{i=1}^{n} E_i \int_{\mathcal{I}_{-2}} (\hat{a}_{ni}(u)-\tilde{a}_i(u))d\varphi(u) - \int_{\mathcal{I}_{-2}} n^{1/2}(F_n(F^{-1}(u))-u)F^{-1}(u)d\varphi(u)\right| \\
\le& \ |I_1|+|I_2|+|I_3|
\end{aligned}$$

where

$$\begin{aligned}
I_1 &= n^{-1/2}\sum_{i=1}^{n} E_i \int_{\mathcal{I}_{-2}} (1-\hat{a}_{ni}(u))d\varphi(u) \\
I_2 &= n^{-1/2}\sum_{i=1}^{n} E_i \int_{\mathcal{I}_{-2}} (1-\tilde{a}_i(u))d\varphi(u) \\
I_3 &= \int_{\mathcal{I}_{-2}} n^{1/2}(u-F_n(F^{-1}(u)))F^{-1}(u)d\varphi(u)
\end{aligned}$$

and we shall treat $I_1, I_2, I_3$ separately. Then, using (2.1), **(X1)** and the behavior of extreme observations of a population with exponentially tailed density satisfying **(X2)**, we obtain

$$\begin{aligned}
|I_1| &\le n^{-1/2}\max_{1\le i\le n}|E_i|\int_0^{\alpha_n^*}|\varphi'(u)|\sum_{i=1}^{n}(1-\hat{a}_{ni}(u))du \\
&\le n^{1/2}\int_0^{\alpha_n^*} u^{-\delta}du . O_p((\log\ n)^{\frac{1}{r}+\frac{\delta}{2}}) = O_p\Big(n^{\frac{1}{2}-(1-\delta)/(1+4b)}(\log\ n)^{\frac{1}{r}+\frac{\delta}{2}}\Big) = o_p(1)
\end{aligned}$$

for $\frac{\delta}{2} < b < \frac{1}{4}-\frac{\delta}{2}$. Similarly,

$$\begin{aligned}
I_2 &= n^{-1/2}\int_0^{\alpha_n^*}\sum_{i=1}^{n} E_i(1-\tilde{a}_{ni}(u))d\varphi(u) \\
&= n^{-1/2}\sum_{i=1}^{n} E_i[\varphi(\alpha_n^*)-\varphi(F(E_i))]I[F(E_i)<\alpha_n^*].
\end{aligned}$$

Because $\varphi$ is nondecreasing and skew-symmetric, this further implies

$$\begin{aligned}
&\mathrm{Var}(I_2) \le \mathbb{E}\Big(E_1^2(2\varphi(F(E_1)))^2 I[F(E_i)<\alpha_n^*]\Big) \\
=&\ 4\int_{-\infty}^{F^{-1}(\alpha_n^*)} x^2\varphi^2(F(x))dF(x) = 4\int_0^{\alpha_n^*}(F^{-1}(u))^2\varphi^2(u)du \to 0
\end{aligned}$$

as $n\to\infty$. Similarly we treat $I_3$, regarding that $\mathbb{E}\Big(n^{1/2}(u-F_n(F^{-1}(u))\Big)^2 = u(1-u)$. Hence, $n^{1/2}\int_{\mathcal{I}_{-2}} A_n(u)d\varphi(u) = o_p(1)$ and analogously $n^{1/2}\int_{\mathcal{I}_2} A_n(u)d\varphi(u) = o_p(1)$ and finally we conclude that

$$n^{1/2}\int_0^1 A_n(u)d\varphi(u) = o_p(1). \tag{2.25}$$

This together with (2.24) implies

$$n^{1/2}(S_n - S(F)) = n^{1/2}\int_0^1 \varphi(u)(F_n(F^{-1}(u)) - u)dF^{-1}(u) + o_p(1) \tag{2.26}$$

and this, in turn, proves the theorem. □

**4. Goodness-of-fit test and its consistency.** Let $Y_1 \dots, Y_n$ be observations satisfying the model (1.1) and assume that the distribution of the errors $E_1, \dots, E_n$ fulfills the hypothesis $\mathbf{H}_0 : F(x) \equiv F_0(x/\sigma)$ with $\boldsymbol{\beta}, \sigma$ unknown.

Put $\varphi_0(t) = F_0^{-1}(t),\ 0 < t < 1$; then $\varphi_0$ satisfies the condition **($\Phi$1)** provided $F_0$ satisfies **(F1)** and **(F2)**. Denote $S_{n0}$ the scale statistic (2.4) generated by $\varphi_0$. Then under $F \equiv F_0(./\sigma)$, Theorem 2.1 implies

$$\sqrt{n}(S_n - \sigma S_0(F_0)) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2(\tau(F_0^{-1}, F_0))^2)$$

where

$$(\tau(F_0^{-1}, F_0))^2 = \int_0^1 \int_0^1 (u \wedge v - uv) F_0^{-1}(u) F_0^{-1}(v) dF_0^{-1}(u) dF_0^{-1}(v)$$

is in correspondence with (2.9). Next select $\varphi_1 \neq \varphi_0$ and consider the scale statistic $S_{n1}$ (2.4) generated by $\varphi_1$. Then, by Theorem 2.1 and by Cramér-Slutsky theorem, we get

$$\sqrt{n}\Big\{\frac{S_{n0}}{S_{n1}} - \frac{S_0(F_0)}{S_1(F_0)}\Big\} \xrightarrow{\mathcal{D}} \mathcal{N}\Big(0, \frac{(\tau(F_0^{-1}, F_0))^2}{(S_1(F_0))^2}\Big),$$

what is already independent of $\sigma$.

The test criterion for $\mathbf{H}_0$ which we propose is the ratio

$$T_n = \sqrt{n}\Big[\frac{S_{n0}}{S_{n1}} - \xi(F_0)\Big] \frac{S_1(F_0)}{\tau(F_0^{-1}, F_0)} \tag{3.1}$$

where $S_{n0},\ S_{n1}$ are the scale statistics of type (2.4) with $\varphi \equiv \varphi_0$ or $\varphi \equiv \varphi_1$, respectively, $\xi(F_0) = S_0(F_0)/S_1(F_0)$, $S_0(F_0) = \int_0^1 (F_0^{-1}(u))^2 du$, $S_1(F_0) = \int_0^1 \varphi_1(u)(F_0^{-1}(u))du$. Under $\mathbf{H}_0$,

$$T_n \xrightarrow{\mathcal{D}} \mathcal{N}(0,1), \tag{3.2}$$

hence, we reject $\mathbf{H}_0$ provided

$$\sqrt{n}\Big|\frac{S_{n0}}{S_{n1}} - \xi(F_0)\Big| \frac{S_1(F_0)}{\tau(F_0^{-1}, F_0)} \geq \Phi^{-1}\Big(1 - \frac{\alpha}{2}\Big). \tag{3.3}$$

Let $F_0$ and $F$ be two distribution functions, both symmetric around 0 and satisfying conditions **(F1)**, **(F2)**. Put $\varphi_0(u) = F_0^{-1}(u)$, and let $\varphi_1(u)$ be another nondecreasing function satisfying the condition ($\Phi$**1**). It follows from Theorem 2.1 that, under the alternative $F$,

$$\begin{aligned} S_{n1} &\xrightarrow{p} S_1(F) = \int_0^1 \varphi_1(u) F^{-1}(u) du \\ T_n &\xrightarrow{\mathcal{D}} \mathcal{N}\Big(0, \tau^2(F_0^{-1}, F)/(S_1(F))^2\Big) \end{aligned} \tag{3.4}$$

hence

$$\sqrt{n}\Big\{\frac{S_{n0}}{S_{n1}} - \frac{S_0(F)}{S_1(F)}\Big\} \frac{S_1(F)}{\tau(F_0^{-1}, F)} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1). \tag{3.5}$$

We are interested in the consistency of the test against the alternative that the true distribution of the errors has distribution function $F(x/\sigma)$ with some $\sigma > 0$ where $F$ is symmetric and satisfying **(F1)**, **(F2)**, $F \neq F_0$. Combining (3.1), (3.2) and (3.5), we obtain

$$
\begin{aligned}
(3.6)\qquad & P_F\Big(|T_n| \geq \Phi^{-1}\Big(1-\frac{\alpha}{2}\Big)\Big) \\
&= P_F\Big\{\Big|\Big[\sqrt{n}\Big(\frac{S_{n0}}{S_{n1}} - \frac{S_0(F)}{S_1(F)}\Big)\frac{S_1(F)}{\tau(F_0^{-1},F)}\Big]\cdot\frac{S_1(F_0)}{S_1(F)}\cdot\frac{\tau(F_0^{-1},F)}{\tau(F_0^{-1},F_0)} \\
&+ \sqrt{n}\Big[S_0(F)\frac{S_1(F_0)}{S_1(F)} - S_0(F_0)\Big]\Big/\tau(F_0^{-1},F_0)\Big| \geq \Phi^{-1}\Big(1-\frac{\alpha}{2}\Big)\Big\}.
\end{aligned}
$$

The first term in the probability on the right-hand side of (3.6) is of order $O_p(1)$ under $F$, while the second term is $O(n^{1/2})$ provided

$$(3.7)\qquad A(F_0,F) = S_0(F)\frac{S_1(F_0)}{S_1(F)} - S_0(F_0) \neq 0.$$

If it is the case then

$$(3.8)\qquad P_F\Big(|T_n| \geq \Phi^{-1}\Big(1-\frac{\alpha}{2}\Big)\Big) \to 1 \quad \text{as } n \to \infty$$

for fixed $F_0, F$. We shall show that $A(F_0,F) > 0$ provided $F$ has heavier tails than $F_0$ in the sense of Hájek (1974). Following Hájek, Section 34, we say that the distribution $F$ *has heavier (longer) tails than* $F_0$, if there exists a nondecreasing nonconstant function $b(u) : (0,1) \mapsto \mathbb{R}^+$ such that

$$(3.9)\qquad F^{-1}(u) = \begin{cases} b(1-u)F_0^{-1}(u) & \dots\ 0 < u < 1/2 \\ b(u)F^{-1}(u) & \dots\ 1/2 < u < 1. \end{cases}$$

Obviously the relation that $F$ has heavier tails than $F_0$ is invariant to the change of scales. On the other hand, consider the function

$$(3.10)\qquad h_F(x) = F^{-1}(F_0(x)),\ x \geq 0.$$

If $h_F$ is convex, then

$$(3.11)\qquad \frac{d}{dx}h_F(x) = \Big[\frac{f_0(F_0^{-1}(u))}{f(F^{-1}(u))}\Big]_{u=F_0(x)}$$

is nondecreasing for $x > 0$, hence the function $\frac{f_0(F_0^{-1}(u))}{f(F^{-1}(u))}$ is nondecreasing for $1/2 < u < 1$. We shall show that the convexity of function $h_F(x)$ of (3.10) is equivalent to the relation that $F$ has heavier tails than $F_0$ in the sense of Hájek.

**Lemma 3.1** *Let $F_0$ and $F$ be two symmetric absolutely continuous distribution functions. Then the function $h_F(x)$ defined in (3.10) is convex for $x > 0$ if and only if $F$ has heavier tails than $F_0$.*

PROOF. I. Denote

$$a(u) = \frac{f_0(F_0^{-1}(u))}{f(F^{-1}(u))},\quad 1/2 < u < 1.$$

If $h_F(x)$ in convex, $a(u)$ is nondecreasing and

$$[f(F^{-1}(u))]^{-1} = a(u)[f_0(F_0^{-1}(u))]^{-1}, \ 1/2 < u < 1.$$

Hence, for any $u < v < 1$,

$$\begin{aligned} \frac{F^{-1}(u)}{F_0^{-1}(u)} &= \frac{\int_{\frac{1}{2}}^{u}[f_0(F_0^{-1}(t))]^{-1}a(t)dt}{\int_{\frac{1}{2}}^{u}[f_0(F_0^{-1}(t))]^{-1}dt} \leq a(u) \\ &\leq \frac{\int_u^v[f_0(F_0^{-1}(t))]^{-1}a(t)dt}{\int_u^v[f_0(F_0^{-1}(t))]^{-1}dt} = \frac{\int_u^v[f(F^{-1}(t))]^{-1}dt}{\int_u^v[f_0(F_0^{-1}(t))]^{-1}dt} \\ &= \frac{F^{-1}(v) - F^{-1}(u)}{F_0^{-1}(v) - F_0^{-1}(u)}. \end{aligned}$$

Hence, $b(u) = F^{-1}(u)/F_0^{-1}(u)$ is nondecreasing for $1/2 < u < 1$ and $F^{-1}(u) = b(u)F_0^{-1}(u)$ for $1/2 < u < 1$.

II. Let $F$ have heavier tails than $F_0$. Then $F^{-1}(u) = b(u)F_0^{-1}(u), \ 1/2 \leq u < 1$, where $b(u)$ is nondecreasing in $1/2 \leq u < 1$. Then $F^{-1}(F_0(x)) = [xb(F_0(x))]$, what is a convex function in $x \geq 0$. □

If $F$ has heavier tails than $F_0$, then $F$ also has heavier tails than $F_0(x/\sigma)$ for $\sigma > 0$, because $F^{-1}(u) = \frac{b(u)}{\sigma}\sigma F_0^{-1}(u)$. If the function $b(u)$ is integrable, then the inequality $\bar{b} > S_1(F)/S_1(F_0)$ with $\bar{b} = \int_0^1 b(u)du$ implies that $A(F_0, F) > 0$, and this all together confirms the consistency of the test, stated in the following theorems.

**Theorem 3.1** *Assume that $F$ have heavier tails than $F_0$ with an nondecreasing integrable function $b(u), 0 < u < 1$ such that*

(3.12) $$\bar{b} > S_1(F)/S_1(F_0).$$

*Then $A(F_0, F) > 0$ and the test (3.3) is consistent against the alternative $F$ where $F(.) \neq F_0(./\sigma)$ for any $\sigma > 0$.*

PROOF. Denote $c(F_0, F) = \frac{S_1(F)}{S_1(F_0)}$. If $F$ has heavier tails than $F_0$ and (3.12) holds, then

$$\begin{aligned} F^{-1}(u) &= b(u)F_0^{-1}(u), \ 1/2 < u < 1, \text{ hence} \\ S_0(F) - c(F_0, F)S_0(F_0) &= 2\int_{1/2}^{1}(F_0^{-1}(u))^2(b(u) - c(F_0, F))du \\ &\geq \int_0^1(F_0^{-1}(u))^2du\int_0^1(b(u) - c(F_0, F))du \geq 0. \end{aligned}$$

□

**Remark.** The simple condition in (3.12) could be easily verified in the particular case $\varphi_1 = -1, 0, 1$ for $0 \leq u < 1/4$, $1/4 \leq u \leq 3/4$ and $3/4 < u \leq 1$, when $S_1(F)$ relates to the interquartile range. However, (3.12) is sufficient but not necessary for $A(F_0, F) > 0$ and not easily verifiable for arbitrary monotone $\varphi_1(u)$. Keeping some other important special $\varphi_1$'s in mind, such as $\varphi_1(u) = \text{sign}\,(u - \frac{1}{2})$, leading to $S_1(F) \equiv \mathbb{E}_F|X|$, we consider an alternative approach that may work out better in some cases.

Assume that the distribution functions $F_0$, $F$ and the score function $\varphi_1$ satisfy the following conditions:

**(A)** $F^{-1}(u) = b(u)F_0^{-1}(u)$, $1/2 < u < 1$, where $b(u)$ is nondecreasing in $(1/2, 1)$ and $b(u) = b(1-u) \geq 0\ \forall u \in (0,1)$.

**(B)** $F_0^{-1}(u) = b_0(u)\varphi_1(u)$, $1/2 < u < 1$, where $b_0(u)$ is nondecreasing in $(1/2, 1)$, $\lim_{u \to 1} b_0(u) = \infty$ and $b_0(u) = b_0(1-u) \geq 0\ \forall u \in (0,1)$.

The following theorem shows that the conditions **(A)** and **(B)** guarantee the consistency of the test against the alternative $F$.

**Theorem 3.2** *Under the conditions* **(A)** *and* **(B)**, $A(F_0, F) > 0$.

PROOF. The condition **(B)** implies that, for every $c > 0$, there exists a $u_0 \in (1/2, 1)$ such that

$$(F_0^{-1}(u))^2 \begin{cases} < & cF_0^{-1}(u)\varphi_1(u) \quad if \quad u < u_0 \\ = & cF_0^{-1}(u)\varphi_1(u) \quad if \quad u = u_0 \\ > & cF_0^{-1}(u)\varphi_1(u) \quad if \quad u > u_0. \end{cases}$$

Consider two auxiliary distribution functions

$$\begin{aligned} G_0(u) &= \Big(\textstyle\int_{1/2}^{u}(F_0^{-1}(v))^2dv\Big)\Big/\Big(\int_{1/2}^{1}(F_0^{-1}(v))^2dv\Big), & 1/2 \leq u \leq 1, \\ G_1(u) &= \Big(\textstyle\int_{1/2}^{u}(F_0^{-1}(v))\varphi_1(v)dv\Big)\Big/\Big(\int_{1/2}^{1}(F_0^{-1}(v))\varphi_1(v)dv\Big), & 1/2 \leq u \leq 1. \end{aligned}$$

Actually, $G_0$ and $G_1$ are nondecreasing and $G_0(1/2) = G_1(1/2) = 0$ and $G_0(1) = G_1(1) = 1$. Choosing

$$c = \Big(\int_{1/2}^{1} F_0^{-1}(u)\varphi_1(u)du\Big)\Big/\Big(\int_{1/2}^{1}(F_0^{-1}(u))^2du\Big),$$

we have for $u \geq u_0$,

$$\begin{aligned} 1 - G_0(u) &= \Big(\int_u^1 (F_0^{-1}(v))^2dv\Big)\Big/\Big(\int_{1/2}^1 (F_0^{-1}(v))^2dv\Big) \\ &\geq c\Big(\int_u^1 F_0^{-1}(v)\varphi_1(v)dv\Big)\Big/\Big(\int_{1/2}^1 (F_0^{-1}(v))^2dv\Big) \\ &= \Big(\int_u^1 F_0^{-1}(v)\varphi_1(v)dv\Big)\Big/\Big(\int_{1/2}^1 F_0^{-1}(v)\varphi_1(v)dv\Big) \\ &= 1 - G_1(u). \end{aligned}$$

Treating the relations analogously for $u \leq u_0$, we arrive at

$$G_0(u) \leq G_1(u) \quad \forall u \in (1/2, 1),$$

This further implies that

$$(3.13)\qquad \frac{\int_{1/2}^{1} F_0^{-1}(u)F^{-1}(u)du}{\int_{1/2}^{1}(F_0^{-1}(u))^2du} = \frac{\int_{1/2}^{1} b(u)(F_0^{-1}(u))^2du}{\int_{1/2}^{1}(F_0^{-1}(u))^2du}$$

$$= \int_{1/2}^{1} b(u)dG_0(u) = \int_{1/2}^{1}[1-G_0(u)]db(u) \geq \int_{1/2}^{1}[1-G_1(u)]db(u) = \int_{1/2}^{1} b(u)dG_1(u)$$

$$= \Big(\int_{1/2}^{1}(F^{-1}(u))\varphi_1(u)du\Big)\Big/\Big(\int_{1/2}^{1}(F_0^{-1}(u))\varphi_1(u)du\Big)$$

where the last equality is due to the condition **(B)**. (3.13) implies that $A(F, F_0) \geq 0$. □

**4. Contiguous alternatives and asymptotic efficiency of the test.** Consider the family of alternatives to $\mathbf{H}_0$

$$(4.1)\qquad \mathbf{K}_n : E_1, \ldots, E_n \text{ have distribution function } G_n(x/\sigma),\ \sigma,\ \beta \text{ unspecified,}$$

where $G_n$ has the density of the form

$$(4.2)\qquad g_n(x) = f_0(x)(1 + n^{-1/2}\lambda u(x)),\ x \in \mathbb{R},\ \lambda > 0,$$

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

$$(4.3)\qquad \int_{-\infty}^{\infty} u(x)f_0(x)dx = 0, \quad \int_{-\infty}^{\infty}(u(x))^2 f_0(x)dx < \infty.$$

By the third LeCam's lemma, the sequence of distributions with densities $\prod_{i=1}^{n} g_n(x_i/\sigma)$ is contiguous with respect to the sequence $\prod_{i=1}^{n} f_0(x_i/\sigma)$, $\sigma > 0$. The distribution function $G_n$ corresponding to $g_n$ could be written in the form

$$(4.4)\qquad G_n(x) = F_0(x) + n^{-1/2}\Lambda(x), \quad \Lambda(x) = \lambda\int_{-\infty}^{x} u(x)dF_0(x).$$

Notice that $\Lambda$ is bounded, $\Lambda(-\infty) = \Lambda(\infty) = 0$. The function $u(.)$ in (4.1) is typically either even or odd, according to the context.

Choose two score functions $\varphi_0, \varphi_1$, both satisfying ($\boldsymbol{\Phi}$**1**), and calculate two corresponding scale statistics $S_{n0}, S_{n1}$ according to (2.4). Then the ratio $S_{n0}/S_{n1}$ is regression and scale invariant.

The asymptotic representation (2.14) holds also under contiguous alternatives, so holds also the convergence $S_{n1} \xrightarrow{p} S_1(F_0)$. Regarding that, we are able to derive the asymptotic distribution of $S_{n0}/S_{n1}$ under $\mathbf{K}_n$, given in the following theorem.

**Theorem 4.1** *Let $Y_1, \ldots, Y_n$ follow the model (1.1) with the regression matrix satisfying* **(X1)** *-* **(X3)**. *Let $S_{n0}$ and $S_{n1}$ be the scale statistic defined in (2.4) with the score functions $\varphi_0, \varphi_1$*

*satisfying* ($\boldsymbol{\Phi 1}$). *Assume that the hypothetical distribution function* $F_0$ *satisfies conditions* **F1** - **F2**. *Then, under* $\mathbf{K}_n$,

$$\sqrt{n}\Big\{\frac{S_{n0}}{S_{n1}} - \frac{S_0(F_0)}{S_1(F_0)}\Big\} \tag{4.5}$$

*is asymptotically normally distributed*

$$\mathcal{N}\Big((S_1(F_0))^{-1}\int_{\mathbf{R}}\Lambda(x)\varphi_0(F_0(x))dx, (S_1(F_0))^{-2}\tau^2(\varphi_0, F_0)\Big) \tag{4.6}$$

*with* $\tau^2(\varphi_0, F_0)$ *given in (2.9), with* $\varphi$, $F$ *replaced by* $\varphi_0$, $F_0$, *respectively.*

PROOF. First, using the per partes integration, ($\boldsymbol{\Phi 1}$), (4.3) and (4.4), we could show that the expectation of the asymptotic distribution (4.6) is finite:

$$\begin{aligned} &|\int_0^1 \varphi_0(u)\Lambda(F_0^{-1}(u))dF_0^{-1}(u)| = |\int_{\mathbf{R}}\Lambda(x)\varphi_0(F_0(x))dx| \\ = &\ \lambda|\int_{\mathbf{R}} xu(x)\varphi_0(F_0(x))f_0(x)dx| + \lambda|\int_{\mathbf{R}} x\Lambda(x)\varphi_0'(F_0(x))f_0(xdx)| < \infty. \end{aligned} \tag{4.7}$$

Due to contiguity, the representation (2.14) holds also under $G_n$ as the distribution function of $E_i$, $i = 1\ldots,n$. Then we could write

$$\begin{aligned} &\psi(E_i) = \psi(G_n^{-1}(U_i)) \\ = &\int_0^1 [f_0(F_0^{-1}(u))]^{-1}(u - I[U_i \le G_n(F_0^{-1}(u))])\varphi_0(u)du \\ = &\int_0^1 [f_0(F_0^{-1}(u))]^{-1}\{u - I[U_i \le u] + I[u < U_i \le u + n^{-1/2}\Lambda(F_0^{-1}(u))]\}\varphi_0(u)du \\ = &\ \psi(F_0(U_i)) + \int_0^1 [f_0(F_0^{-1}(u))]^{-1}\varphi_0(u)I[u < U_i \le u + n^{-1/2}\Lambda(F_0^{-1}(u))]du, \ i = 1,\ldots,n \end{aligned} \tag{4.8}$$

where $U_1, U_2, \ldots$ are independent observations from the uniform $R(0,1)$ distribution. This implies that

$$\begin{aligned} &\sqrt{n}(S_{n0} - S_0(F_0)) = n^{-1/2}\sum_{i=1}^n \psi(F_0(U_i)) + \int_{\mathbf{R}}\Lambda(x)\varphi_0(F_0(x))dx \\ + &\int_0^1 n^{1/2}[U_n(u + n^{-1/2}\Lambda(F_0^{-1}(u)) - U_n(u) - n^{-1/2}\Lambda(F_0^{-1}(u))]\varphi_0(u)dF_0^{-1}(u) + o_p(1) \end{aligned} \tag{4.9}$$

where

$$U_n(v) = n^{-1}\sum_{i=1}^n I[U_i \le v].$$

Denoting

$$V_n(u) = n^{1/2}[U_n(u + n^{-1/2}\Lambda(F_0^{-1}(u)) - U_n(u) - n^{-1/2}\Lambda(F_0^{-1}(u))], \quad 0 < u < 1,$$

we have

$$\begin{aligned}(4.10)\qquad & \mathbb{E}\Big[\int_0^1 V_n(u)\varphi_0(u)dF_0^{-1}(u)\Big]^2 \\ = & \ \mathbb{E}\int_0^1\int_0^1 V_n(u)V_n(v)\varphi_0(u)\varphi_0(v)dF_0^{-1}(u)dF_0^{-1}(v) \\ = & \int_0^1\int_0^1 \mathbb{E}(V_n(u)V_n(v))\varphi_0(u)\varphi_0(v)dF_0^{-1}(u)dF_0^{-1}(v) \\ \le & \int_0^1\int_0^1 [\mathbb{E}(V_n(u))^2\mathbb{E}(V_n(u))^2]^{1/2}|\varphi_0(u)\varphi_0(v)|dF_0^{-1}(u)dF_0^{-1}(v) \\ = & \Big[\int_0^1 [\mathbb{E}(V_n(u))^2]^{1/2}|\varphi_0(u)|dF_0^{-1}(u)\Big]^2\end{aligned}$$

It follows from (4.3) and (4.4) that

$$\begin{aligned}(4.11)\qquad |\Lambda(F_0^{-1}(v))| & \le \lambda v^{1/2}\Big[\int_{-\infty}^{\infty} u^2(x)f_0(x)dx\Big]^{1/2} \\ & = \lambda v^{1/2}C^{1/2}(F_0)\ \text{[say]}, \quad 0 \le v \le 1.\end{aligned}$$

Let us split the integration domain of last integral in (4.10) in three parts: $(0,\alpha_0),[\alpha_0,1-\alpha_0],(1-\alpha_0,1)$ and consider the three integrals separately. Regarding (4.11), **($\Phi$1)** and **(F1)** - **(F2)**, we get

$$\begin{aligned} & \int_0^{\alpha_0} [\mathbb{E}(V_n(u))^2]^{1/2}|\varphi_0(u)|dF_0^{-1}(u) \\ \le & \ n^{-1/4}C^{1/4}(F_0)\int_0^{\alpha_0} v^{(1/4)-\delta}[f_0(F_0^{-1}(v))]dv \\ (4.12)\qquad \le & \ Kn^{-1/4}\int_0^{\alpha_0} v^{(-3/4)-\delta}(-\log\ v)^{(1/r)-1}dv = O(n^{-1/4}).\end{aligned}$$

Similarly we treat $\int_{1-\alpha_0}^1[\mathbb{E}(V_n(u))^2]^{1/2}|\varphi_0(u)|dF_0^{-1}(u)$, while the middle integral is obviously $O(n^{-1/4})$. This together with (4.9) implies, under $E_1,\ldots,E_n$ distributed according to $G_n$,

$$(4.13)\qquad \sqrt{n}(S_{n0}-S_0(F_0)) = n^{-1/2}\sum_{i=1}^n \psi(F_0(U_i)) + \int_{\mathbf{R}}\Lambda(x)\varphi_0(F_0(x))dx + o_p(1).$$

and this, in turn, implies the theorem. □

Let $F$ satisfy **(F1)**, **(F2)** and let the function $u(x)$ satisfying (4.2) and (4.3) be odd. Then the tails of the distribution function $G_n$, defined in (4.4), behave in the following way:

$$\begin{aligned}(4.14)\qquad & \lim_{x\to-\infty}\ \frac{\log(G_n(x))}{-c|x|^r - n^{-1/2}\lambda u(|x|)} = 1 \\ & \lim_{x\to\infty}\ \frac{\log(1-G_n(x))}{-cx^r + n^{-1/2}\lambda u(x)} = 1.\end{aligned}$$

Hence, if $u(x)$ is nondecreasing, then the left hand tail of $G_n$ is lighter and the right hand tail heavier than those of $F_0$; the situation is opposite when $u(x)$ is odd and nonincreasing.

Actually, denoting $v = F_0(x)$, we shall conclude from (4.4) for $x \to -\infty$

$$\Lambda(F_0^{-1}(v)) = \lambda v.u(F_0^{-1}(v)) + o(v) = \lambda F_0(x)u(x) + o(F(x)), \tag{4.15}$$

and hence, as $x \to -\infty$,

$$\begin{aligned}\log\ G_n(x) &\approx \log\ F(x) + \log(1 - n^{-1/2}\lambda u(|x|)) \\ &\approx -c|x|^r - n^{-1/2}\lambda u(|x|);\end{aligned}$$

we proceed analogously for $x \to \infty$.

On the other hand, if $u(x)$ is even and positive in the tails, then we could prove analogously that both tails of $G_n$ are heavier than those of $F_0$.

Though we have introduced a general class of alternatives that may as well be of nonparametric nature, it is interesting to note that a parametric alternative (allowing nuisance regression/scale) may also be treated in this setup. For an illustration, consider the usual Edgeworth type expansion for the density $f(x)$, standardized to null location and unit scale parameters. As in Cramér (1946), p.223, we may write

$$f(x) = \phi(x)\Big\{1 + \frac{c_3}{3!}H_3(x) + \frac{c_4}{4!}H_4(x) + \dots\Big\}$$

where the $H_k(x)$ are the Hermite polynomials, and $c_3 = -\gamma_1,\ c_4 = \gamma_2, \dots$ with $\gamma_1$ and $\gamma_2$ being the Pearson measures of skewness and kurtosis. Thus, under the symmetry condition on $f(x)$, $\gamma_1 = 0$ while $\gamma_2$ (or $c_4$) reflects the kurtosis of $f$, compared to the standard normal density $\phi(x)$. Similarly, in the case of $c_4 = 0$, $c_3$ reflects the skewness.

In the case of contiguous departures along the skewness and kurtosis paths we let

$$\frac{c_3}{3!} = n^{-1/2}\lambda_S \quad \text{or} \quad \frac{c_4}{4!} = n^{-1/2}\lambda_K.$$

This form of contiguous alternatives for infinitely divisible laws may be justified by the basic Theorem 2 in Feller (1966, p.508).

On the other hand, if we take a convenient function $u(x)$ (*e.g.*, odd or even and bounded) in the general setup, we avoid the restrictive moment conditions which are generally associated with the Edgeworth type of expansions.

**Remark.** As another class of of contiguous alternatives we may consider the *Lehmann alternatives* (Lehmann 1953) with the distribution functions $G_n(x)$, where

$$G_n(x) = \frac{1}{2}\Big\{[F_0(x/\sigma)]^{1+\lambda/\sqrt{n}} + 1 - [1 - F_0(x/\sigma)]^{1+\lambda/\sqrt{n}}\Big\},\ \sigma > 0 \text{ nuisance.}$$

The hypothesis $\mathbf{H}_0$ could be then rewritten as $\mathbf{H}_0 :\ \lambda = 0$.

## REFERENCES

CRAMÉR, H. (1946). *Mathematical Methods of Statistics.* Princeton University Press.

DURBIN, J. (1973). *Distribution Theory for Tests Based on Sample Distribution Function.* SIAM, Philadelphia.

FELLER, W. (1966). *An Introduction to Probability Theory and Its Applications, Vol.II.* J.Wiley, New York.

GUTENBRUNNER, C. AND JUREČKOVÁ, J. (1992). Regression rank scores and regression quantiles. *Ann. Statist.* **20** 305-330.

GUTENBRUNNER, C., JUREČKOVÁ, J., KOENKER, R. AND PORTNOY, S. (1993). Tests of linear hypotheses based on regression rank scores. *J. Nonpar. Statist.* **2** 307-331.

HÁJEK, J. (1969). *A Course in Nonparametric Statistics.* Holden-Day, San Francisco.

HÁJEK, J. AND ŠIDÁK, Z. (1967). *Theory of Rank Tests.* Academia, Prague.

HOEFFDING, W. (1948). The strong law of large numbers for U-statistics. *Univ. North Carolina Inst. Statist. Mimeo Series* No 302.

JAECKEL, L.A. (1972). Estimating regression coefficients by minimizing the dispersion of the residuals. *Ann. Math. Statist.* **43** 1449-1458.

JUREČKOVÁ, J. (1995). Trimmed mean and Huber's estimator: Their difference as a goodness-of-fit criterion. *Journal of Statistical Research* **29** 31-35.

JUREČKOVÁ, J. AND SEN, P.K. (1994). Regression rank scores scale statistics and studentization in linear models. *Proc. 4th Prague Symp. on Asymptotic Statistics* (P.Mandl and M. Hušková, eds.), 112-121.

JUREČKOVÁ, J. AND SEN P.K. (1996). *Robust Statistical Procedures. Asymptotics and Interrelations.* J.Wiley, New York.

JUREČKOVÁ, J. AND SEN P.K. (1998). Second-order asymptotic relations and goodness-of-fit tests. Submitted.

KHMALADZE, E.V. (1993). Goodness of fit problem and scanning innovation martingales. *Ann. Statist.* **21** 798-829.

LEHMANN, E.L. (1953). The power of rank tests. *Ann. Math. Statist.* **24** 23-43.

DEPARTMENT OF PROBABILITY
AND STATISTICS
CHARLES UNIVERSITY
SOKOLOVSKÁ 83
CZ-186 75 PRAGUE 8, CZECH REPUBLIC

DEPARTMENTS OF BIOSTATISTICS AND STATISTICS
UNIVERSITY OF NORTH CAROLINA
3101 MCGAVRAN-GREENBERG HALL, CB #7400
CHAPEL HILL, NC 27599-7400
USA