Probab. Th. Rel. Fields 82, 19–31 (1989)

Probability Theory and Related Fields
© Springer-Verlag 1989

Estimation of the Slope in a Linear Functional Relationship

Philip Milasevic
UER de Mathématiques, BFSH-2, Université de Lausanne, CH-1015 Lausanne, Switzerland

Summary. Two nonparametric estimators of the slope of a regression line with error on both variables are considered, each of them being defined as the zero-crossing of a stochastic process whose sample paths are monotone. Their asymptotic behaviour is derived from the local asymptotic behaviour of the underlying processes. One of the estimators is a nonparametric version of Wald's (1940) estimator.

0. Introduction

This paper deals with the estimation of the coefficient a in the following model:

$(X_1, Y_1), \ldots, (X_n, Y_n)$ are observed, with

$$Y_i = Y_i^* + \eta_i$$
$$X_i = X_i^* + \varepsilon_i$$
$$Y_i^* = aX_i^* + b, \quad i=1, \ldots, n, a \neq 0,$$

where a, b are unknown, $X_1^*, \ldots, X_n^*, Y_1^*, \ldots, Y_n^*$ are unobservable real numbers (the "true" values) and $\varepsilon_1, \ldots, \varepsilon_n, \eta_1, \ldots, \eta_n$ are real-valued random "errors." We shall assume that the ε's are i.i.d., the η's are i.i.d. and that the two sets of errors are independent.

In 1940 Wald proposed an estimator which turns out to be consistent and asymptotically normal under a rather restrictive assumption about the model but without assumptions about the error variances. See Wald (1940) and Chap. 29 of Kendall and Stuart (1973). In the present paper a nonparametric version of Wald's estimator is proposed. Under an assumption (assumption A below) analogous to that of Wald consistency and asymptotic normality are obtained. The role of assumption A is then illustrated in a special case. When the errors are symmetrically distributed about the origin a rank estimator is also investigated.

This research was supported by the Swiss National Science Foundation

Each one of the estimators is defined as the zero-crossing of a monotone stochastic process and their asymptotic normality is derived from a local asymptotic linearity property of the underlying processes.

A consistent estimate for the asymptotic variance of one of the estimators is also given and its rate of convergence is obtained by establishing a central limit theorem for the related process. More precisely a sequence $T_n(t)$ of processes in $D[-1, 1]$ and a sequence d_n of real numbers are obtained such that $n^{1/2}(T_n(t)-d_n t)$ converges weakly to a linear process with quadratic drift.

The paper is divided in six sections: 1. The estimators; 2. Asymptotic linearity; 3. Asymptotic behaviour of the estimators; 4. Relative efficiences; 5. About assumption A; 6. Variance estimation and rate of convergence.

Notation

The indicator of a set C is denoted by $I(C)$. Given a vector $(z_1, \ldots, z_n)$, z_{ij} means the difference $z_j - z_i$ and $R(z_i)$ the rank of z_i in $\{z_1, \ldots, z_n\}$. The integer part of a real number z is denoted by $[z]$. All distribution functions defined below are assumed to be absolutely continuous with respect to Lebesgue's measure on the real line. We denote by F, G, F^*, G^*, H_a the respective c.d.f.s of η_1, ε_1, $\eta_1 - \eta_2$, $\varepsilon_1 - \varepsilon_2$, $\eta_1 - a\varepsilon_1$, with respective densities f, g, f^*, g^*, h_a.

1. The Estimators

(1.1) The estimator Δ_{1n} is defined as any value of Δ minimizing the absolute value of

$$R_n(\Delta) = \sum_{\substack{i=1,\ldots,m \\ j=m+1,\ldots,n}} \left\{ I\left(\frac{Y_{\tau_j} - Y_{\tau_i}}{X_{\tau_j} - X_{\tau_i}} \leq \Delta \right) - 1/2 \right\}$$

where m is the integer part of $n/2$ and τ is the vector of antiranks of $(X_1, \ldots, X_n)$. We shall investigate the asymptotic behavior of Δ_{1n} under the crucial

Assumption A. If $M = \{i=1, \ldots, n \mid R(X_i^*) \leq m\}$ and $N = M^c = \{i=1, \ldots, n \mid R(X_i^*) > m\}$, then $\{\tau_1, \ldots, \tau_m\} \overset{\text{a.s.}}{=} M$ and $\{\tau_{m+1}, \ldots, \tau_n\} \overset{\text{a.s.}}{=} N$.

This assumption means that we can divide the observations according to their size in two groups independently of the errors. See also Wald (1940), Moran (1971). Its importance shall be illustrated in Sect. 5.

Note that Δ_{1n} could more generally be defined by taking $m = [\alpha n]$ for some $0 < \alpha \leq 1/2$ in the definition of $R_n(\Delta)$. If the set N in assumption A is replaced by $\{i=1, \ldots, n \mid R(X_i^*) > [(1-\alpha)n]\}$ the results below, which are stated for $\alpha = 1/2$, can be straightforwardly generalized to any α. Not surprisingly the efficiency of the estimator increases with α.

Instead of working with the variable Δ and the process $R_n(\Delta)$ (see for example Bhattacharya, Chernoff, Yang (1983)), we shall work with a new variable t, connected with Δ by the relation $\Delta = a + tn^{-1/2}$ (where a is the slope) and the process $Y_n(t) := n^{-3/2} R_n(a + t n^{-1/2})$. Using the structure of the model and assumption A we get

$$Y_n(t) = n^{-3/2} \sum_{(i,j) \in M \times N} \{ I(\eta_{ij} - a\varepsilon_{ij} \leqq t n^{-1/2}(X_{ij}^* + \varepsilon_{ij})) - \tfrac{1}{2} \}.$$

(1.2) The estimators Δ_{2n} and Δ_{2n}^*.
Suppose that the intercept b is known (we assume WLOG that $b=0$). Then we define Δ_{2n} as any value of Δ minimizing the absolute value of

$$\sum_{i=1}^{n} R(|Y_i - \Delta X_i|) \operatorname{sign}(Y_i - \Delta X_i).$$

Suppose that b is unknown but that we have a $\sqrt{n}$-consistent (preliminary) estimate b_n for b (see the end of Sect. 3). We then define Δ_{2n}^* as any value of Δ minimizing the absolute value of

$$\sum_{i=1}^{n} R(|Y_i - \Delta X_i - b_n|) \operatorname{sign}(Y_i - \Delta X_i - b_n).$$

By the same transformation as in (1.1) and since b_n is $\sqrt{n}$-consistent, we are then led to the process

$$U_n(t, s) = n^{-3/2} \sum_{i=1}^{n} R(|\eta_i - a\varepsilon_i - t n^{-1/2}(X_i^* + \varepsilon_i) - s n^{-1/2}|)$$
$$\cdot \operatorname{sign}(\eta_i - a\varepsilon_i - t n^{-1/2}(X_i^* + \varepsilon_i) - s n^{-1/2}).$$

If $b=0$, the process of interest is $U_n(t, 0)$.

We shall investigate the asymptotic behaviour of Δ_{2n}^* for f and g symmetric about the origin and under the following

Assumption B. $\forall i \in \{1, \ldots, n\}$, $X_i^* > 0$ and $X_i \overset{\text{a.s.}}{\geqq} 0$.

Assumption B entails the monotonicity of the process $U_n(t, s)$ (see Sect. 2) and hence the fact that Δ_{2n} and Δ_{2n}^* are well defined.

Note that each of the two assumptions A, B implies that ε_1 has a bounded support.

Remark. For practical purposes assumptions A, B appear to be rather restrictive. However, in many physical experiments the quantities of interest are positive and assumption B would then cause no trouble. Concerning assumption A there are physical situations such as some calibration problems where the points of measurement are choosen by the experimenter and where the assumption could be met.

2. Asymptotic Linearity

Notation. For $t, s \in \mathbb{R}^k$, $t \leqq s$ means that the inequality holds componentwise. We write $\langle , \rangle$ for the usual scalar product on $\mathbb{R}^k$. For $\delta > 0$, a δ-net of $A \subset \mathbb{R}$ is a finite set of points $a_1 < a_2 < \ldots < a_N$ of A such that

$$\max_{i=1,\ldots,N-1} (a_{i+1} - a_i) \leqq \delta \quad \text{and} \quad \bigcup_i [a_i, a_{i+1}] = A.$$

For $f\colon \mathbb{R}^k \to \mathbb{R}$ and $A \subset \mathbb{R}^k$, we write $\|f\|_A$ for $\sup_{x \in A} |f(x)|$. For $A \subset \mathbb{R}$, $A^k \subset \mathbb{R}^k$ is the k-fold cartesian product of A. Let $I \subset \mathbb{R}$ be a compact interval and $\mathbb{R}_+$ ($\mathbb{R}_-$) the set of strictly positive (negative) real numbers. For $u, v \in \mathbb{R}^k$, $u \leqq v$, $[u, v]$ means the product $\prod_{j=1}^{k} [u_j, v_j]$.

(2.1) **Lemma.** *Let $(X(t))_{t \in \mathbb{R}^k}$ be a process whose sample paths are nonincreasing (nondecreasing) with probability one and such that $\|X\|_{[u,v]}$ is measurable for each $[u, v] \subset I^k$. Let $c \in \mathbb{R}^k_-$ ($\mathbb{R}^k_+$) and define $Z(t) = X(t) - X(0) - \langle c, t \rangle$. Let $\delta \in \mathbb{R}^k_+$ and Γ_i be a δ_i-net of I, $i = 1, \ldots, k$. Then*

$$\|Z\|_{I^k} - \|Z\|_{\prod_{i=1}^{k} \Gamma_i} \overset{\text{a.s.}}{\leqq} \langle |c|, \delta \rangle \quad \text{with } |c| = (|c_1|, \ldots, |c_k|).$$

Proof. Let $X(t)$ be nonincreasing (say). It is enough to show that for $[u_j, v_j] \subset I$, $j = 1, \ldots, k$, we have

$$\|Z\|_{[u,v]} - \|Z\|_{\{u,v\}} \overset{\text{a.s.}}{\leqq} \langle |c|, v - u \rangle$$

where $\{u, v\} = \prod_{j=1}^{k} \{u_j, v_j\}$ and that the exception set does not depend on u or v. We distinguish two cases

1. $\|Z\|_{[u,v]} = \sup_{t \in [u,v]} Z(t)$ 2. $\|Z\|_{[u,v]} = \sup_{t \in [u,v]} (-Z(t))$

I shall consider the first case. The second is analogous.

(i) Suppose that $Z(u) \geqq 0$ and let $t \in \{s \in [u, v] \mid Z(s) \geqq 0\} =: [u, v]_+$. Then $|Z(t)| - |Z(u)| \overset{\text{a.s.}}{\leqq} \langle |c|, v - u \rangle$ by the monotonicity of $X(\cdot)$ and the fact that $t \geqq u$. Note that the exception set does not depend on u, v or t.

(ii) Suppose that $Z(u) < 0$. For $t \in [u, v]_+$, $X(t) - X(0) \geqq \langle c, t \rangle \geqq \langle c, v \rangle$ and by the monotonicity of $X(\cdot)$, $\langle c, v \rangle \leqq \langle c, t \rangle \leqq X(t) - X(0) \leqq X(u) - X(0) < \langle c, u \rangle$, which implies $Z(t) \overset{\text{a.s.}}{\leqq} \langle c, u \rangle - \langle c, v \rangle = \langle |c|, v - u \rangle$.

Therefore, for all $t \in [u, v]_+$, $Z(t) - |Z(u)| \overset{\text{a.s.}}{\leqq} \langle |c|, v-u \rangle$ where the exception set is independent of u, v and t.

Now we can write

$$0 \leqq \|Z\|_{[u,v]} - \|Z\|_{\{u,v\}} \leqq \|Z\|_{[u,v]} - |Z(u)| = \sup_{t \in [u,v]_+} Z(t) - |Z(u)| \overset{\text{a.s.}}{\leqq} \langle |c|, v-u \rangle. \quad \square$$

(2.2) **Lemma.** *Let* $(X_n(t))_{t \in \mathbb{R}^k}$ *be a sequence of monotone processes as in* (2.1). *Suppose that* $\|\mathrm{Var}(X_n(\cdot) - X_n(0))\|_{I^k} = o(1)$ *as* $n \to \infty$ *and that there exists* $c \in \mathbb{R}^k_-$ $(\mathbb{R}^k_+)$ *such that* $\|E(X_n(\cdot) - X_n(0)) - \langle c, \cdot \rangle\|_{I^k} = o(1)$. *Then* $\|Z_n\|_{I^k} = o_P(1)$ *as* $n \to \infty$.

Proof. Assume WLOG that $I = [0, 1]$ and let, for $m \in \mathbb{N}^*$, $\Gamma_m = \{jm^{-1}; j = 0, \ldots, m\}$. By Lemma (2.1),

$$\|Z_n\|_{I^k} - \|Z_n\|_{\Gamma_m^k} \overset{\text{a.s.}}{\leqq} m^{-1} \sum_{j=1}^{k} |c_j|.$$

Given $\varepsilon > 0$ let $m_0 \in \mathbb{N}$ be such that $m_0^{-1} \sum_{j=1}^{k} |c_j| < \varepsilon/2$. Then $P\{\|Z_n\|_{I^k} > \varepsilon\} \leqq P\{\|Z_n\|_{\Gamma_{m_0}^k} > \varepsilon/2\}$ and it is enough to show that $\|Z_n\|_{\Gamma_m^k} = o_P(1)$ for any fixed $m \in \mathbb{N}^*$. Using triangle's inequality and the second hypothesis we have for n large enough

$$\begin{aligned} P\{\|Z_n\|_{\Gamma_m^k} > \varepsilon\} &\leqq P\{\|X_n(\cdot) - X_n(0) - E(X_n(\cdot) - X_n(0))\|_{\Gamma_m^k} > \varepsilon/2\} \\ &\leqq \sum_{t \in \Gamma_m^k} P\{|X_n(t) - X_n(0) - E(X_n(t) - X_n(0))| > \varepsilon/2\} = o(1) \end{aligned}$$

by Chebyshev's inequality and the first hypothesis. $\square$

Asymptotic Linearity of $Y_n(t)$ and $U_n(t, s)$

Henceforth we shall assume that the X^*'s are uniformly bounded and that $\mu_2 = \lim_{n \to \infty} n^{-2} \sum_{(i,j) \in M \times N} X^*_{ij}$ and $\mu_1 = \lim_{n \to \infty} n^{-1} \sum_{i=1}^{n} X^*_i$ exist whenever used. Under assumption A, $Y_n(t)$ is for each n a nondecreasing function of t with probability one. If moreover f satisfies a Lipschitz condition of order > 0, it is a matter of straightforward calculation to show that

$$\sup_{t \in I} \mathrm{Var}(Y_n(t) - Y_n(0)) = O(n^{-1/2}) \quad \text{and} \quad \sup_{t \in I} |E\,Y_n(t) - dt| = o(1),$$

with $d = \mu_2 \int f^*(ax)\, dG^*(x)$. As a consequence we get by lemma (2.2)

(2.3) **Proposition.** *If assumption A is fulfilled and if f satisfies a Lipschitz condition of order* > 0, *then* $\sup_{t \in I} |Y_n(t) - Y_n(0) - dt| = o_P(1)$. $\square$

By a result of Van Eeden (1972), under assumption B, $U_n(t, s)$ is for each n a nonincreasing function of (t, s) with probability one. If moreover f satisfies a Lipschitz condition of order >0, then for f and g symmetric about zero, $\sup_{t,s\in I} \operatorname{Var}(U_n(t, s)-U_n(0, 0))=O(n^{-1/2})$ and $\sup_{t,s\in I}|EU_n(t, s)-d_1 t-d_2 s|=o(1)$, with $d_1=-2\mu_1 \int h_a^2(x)\,dx$ and $d_2=-2\int h_a^2(x)\,dx$. Therefore we get by lemma (2.2)

(2.4) **Proposition.** *If assumption B is fulfilled, if f and g are symmetric about the origin and if f satisfies a Lipschitz condition of order >0, then*

$$\sup_{t,s\in I}|U_n(t, s)-U_n(0, 0)-d_1 t-d_2 s|=o_P(1). \quad \square$$

(2.5) **Corollary.** *Under the hypotheses of* (2.4) *and if b_n is a $\sqrt{n}$-consistent estimator of b,*

$$\sup_{t\in I}|U_n(t, n^{1/2}(b_n-b))-U_n(0, 0)-d_1 t-d_2 n^{1/2}(b_n-b)|=o_P(1).$$

The proof of the corollary follows from (2.4) and the fact that $\sqrt{n}(b_n-b)=O_P(1)$. $\square$

Remark. Propositions (2.3) and (2.4) still hold if, for example, F is everywhere differentiable with bounded derivative f, or if F is the uniform distribution on some interval and g is square-integrable.

3. Asymptotic Behavior of the Estimators

We first give a result about zero-crossings of monotone processes. This result is a modification of a theorem by Jurečková (1971) and can be proved much the same way.

(3.1) Let $(X_n(t, s))_{(t,s)\in\mathbb{R}^2}$ be a sequence of processes satisfying the conditions of lemma (2.2) for some $c=(c_1, c_2)$ and such that $(X_n(0, 0))_{n\geq 1}$ is tight. Let $L_n(t, s)=X_n(0, 0)+c_1 t+c_2 s$ and $t_n(s)$ be defined by $L_n(t_n(s), s)=0$. For $(s_n)_{n\geq 1}$ a tight sequence of real-valued r.v.'s such that $X_n(t, s_n)$ is measurable for each n and t, let $t_n=t_n(s_n)$ and define

$$D_n=\{t\in\mathbb{R}\,|\,t \text{ minimizes } |X_n(\cdot, s_n)|\}.$$

Then $\sup_{t\in D_n}|t-t_n|=o_P(1)$.

Note. This result is of course also valid in the case of one-parameter processes (with $s_n\equiv 0$).

Propositions (3.2) and (3.4) below are now immediate consequences of (3.1) and (2.3), (2.4), (2.5).

(3.2) **Proposition.** *Under the hypotheses of* (2.3) $n^{1/2}(\Delta_{1n}-a)=-d^{-1}Y_n(0)+o_P(1)$ *with* $d=\mu_2\int f^*(ax)\,dG^*(x)$. $\square$

By Hájek's projection method (Hájek (1968)) it is easy to see that $Y_n(0) \xrightarrow[n\to\infty]{\mathscr{L}} \mathcal{N}(0,48^{-1})$ and we get

(3.3) **Corollary.** *Under the hypotheses of* (2.3) $n^{1/2}(\Delta_{1n}-a) \xrightarrow[n\to\infty]{\mathscr{L}} \mathcal{N}(0,48^{-1}d^{-2})$. □

(3.4) **Proposition.** (a) *Under the hypotheses of* (2.4) *and if* b_n *is a* $\sqrt{n}$*-consistent estimator of* b,

$$n^{1/2}(\Delta_{2n}^*-a) = -d_1^{-1}\,U_n(0,0) - d_1^{-1}\,d_2\,n^{1/2}(b_n-b) + o_P(1)$$

with $d_1 = -2\mu_1 \int h_a^2(x)\,dx$, $d_2 = -2\int h_a^2(x)\,dx$.

(b) *Under the hypotheses of* (2.4) *and if* $b=0$,

$$n^{1/2}(\Delta_{2n}-a) = -d_1^{-1}\,U_n(0,0) + o_P(1). \quad \square$$

Since $U_n(0,0) \xrightarrow[n\to\infty]{\mathscr{L}} \mathcal{N}(0,\tfrac{1}{3})$ we get

(3.5) **Corollary.** *Under the hypotheses of* (2.4) *and if* $b=0$, $n^{1/2}(\Delta_{2n} - a) \xrightarrow[n\to\infty]{\mathscr{L}} \mathcal{N}(0,3^{-1}d_1^{-2})$. □

Remark. In the case where b is unknown we used a consistent estimator of b to define Δ_{2n}^*. It seems difficult to estimate b without estimating the slope as well. Moreover it seems impossible to estimate a and b consistently without any additional information (see also Moran (1971)). We give below an estimator of (a, b) when replicate observations are available.

Let $\xi_1, \ldots, \xi_n$ be positive r.v's independent of the ε's and the η's s.t. $\xi_1 - X_1^*, \ldots, \xi_n - X_n^*$ are i.i.d. with zero expectation. If we define $(\tilde{a}_n, \tilde{b}_n)$ as any value of (Δ_1, Δ_2) minimizing

$$\left|\sum_{i=1}^{n} \{I(Y_i-\Delta_1 X_i-\Delta_2 \leqq 0)-\tfrac{1}{2}\}\right| + \left|\sum_{i=1}^{n} \xi_i\{I(Y_i-\Delta_1 X_i-\Delta_2 \leqq 0)-\tfrac{1}{2}\}\right|,$$

it can be shown by the method used earlier (see also Jurečková (1971)) that under some conditions $n^{1/2}\begin{pmatrix}\tilde{a}_n-a\\ \tilde{b}_n-b\end{pmatrix}$ follows asymptotically a bivariate normal distribution and that Δ_{2n}^* with $\tilde{b}_n$ used as a preliminary estimate for b is asymptotically normal with asymptotic variance larger than the one of $\tilde{a}_n$.

4. Relative Efficiencies

4.1. *ARE of Δ_{1n} with respect to Wald's estimator*

Assume for simplicity that $n=2m$. Under assumption A, Wald's estimator can be written as

$$W_n=\frac{\sum_{j\in N} Y_j-\sum_{i\in M} Y_i}{\sum_{j\in N} X_j-\sum_{i\in M} X_i}.$$

Therefore, provided $\int x^2 h_a(x)\,dx<\infty$, we have

$$n^{1/2}(W_n-a)\xrightarrow[n\to\infty]{\mathscr{L}}\mathcal{N}(0,4^{-1}\mu_2^{-2}\int x^2 h_a(x)\,dx).$$

Since $\int f^*(ax)\,dG^*(x)=\int h_a^2(x)\,dx$ we obtain from corollary (3.3) that

$$ARE(\Delta_{1n}, W_n)=12\int x^2 h_a(x)\,dx(\int h_a^2(x)\,dx)^2.$$

Since the latter expression is invariant under scale transformations of the form $(x, y)\mapsto\lambda(x, y)$ one can prove as in Lehmann (1983) that $ARE(\Delta_{1n}, W_n)\geqq 0.864$.

4.2. *ARE of Δ_{1n} with Respect to Δ_{2n} when $b=0$*

Under assumptions A and B, we get from (3.3) and (3.5) that $ARE(\Delta_{1n}, \Delta_{2n})=4\mu_2^2\mu_1^{-2}$. If $m^{-1}\sum_{j\in N} X_j^*\to b_2$ and $m^{-1}\sum_{i\in M} X_i^*\to b_1$ as $n\to\infty$ with $b_2>b_1$, $ARE(\Delta_{1n}, \Delta_{2n})=(b_2-b_1)^2(b_2+b_1)^{-2}<1$.

5. About Assumption *A*

The estimator Δ_{1n} has been defined in Sect. 1 as zero-crossing of the monotone process $R_n(\Delta)$. Under assumption A, $ER_n(a)=0$ and this definition seems to be reasonable. Would $ER_n(a)$ be different from zero, one should consider instead zero-crossings of $R_n(\Delta)-ER_n(a)$. We investigate in this section a particular situation in which assumption A is not fulfilled. We obtain that $ER_n(a)\sim\kappa n^2$ for n large, where unfortunately κ depends upon the parameter a that we want to estimate. The alternative mentioned above is therefore not practicable and, on the other hand, one cannot expect that Δ_{1n} be consistent for a.

(5.1) **Proposition.** *Let X_i^* be 0 or 1 according to $i\leqq m=[n/2]$ or $i>m$ and let ε_1 be uniformly distributed on $[0, 1+\delta]$, with $\delta>0$. Assume that f^* is strictly* *positive on a neighborhood of the origin. Then* $n^{-2} E R_n(a) \xrightarrow[n\to\infty]{} \kappa$, *where* $\kappa = \kappa(a, f, \delta)$ *has the same sign as* a.

In the proof of this result we shall use a corollary of the following lemma.

(5.2) **Lemma.** *Let* $\varepsilon_1, \varepsilon_2, \ldots$ *be independent and uniformly distributed on* $[0, 1+\delta]$, *with* $\delta > 0$. *Let* $X_i = \varepsilon_i$ *for* $i = 1, \ldots, m$ *and* $X_i = 1 + \varepsilon_i$ *for* $i = m+1, \ldots, n$. *Then* $n^{1/2}(X_{(m)} - z_0) \xrightarrow[n\to\infty]{\mathscr{L}} \mathscr{N}(0, 2^{-1}\delta z_0)$, *where* $X_{(m)}$ *is the* m^{th} *order statistics of* $(X_1, \ldots, X_n)$ *and* $z_0 = 1 + 2^{-1}\delta$.

Proof. We assume for simplicity that n is even, i.e. $n = 2m$. Define

$$T_n^*(t; X_1, \ldots, X_n) = \sum_{i=1}^{n} \{I(X_i \leqq t) - 1/2\}.$$

Then $X_{(m)} = \inf\{t \mid T_n^*(t; X_1, \ldots, X_n) \geqq 0\}$. Since $T_n^*(t; X_1, \ldots, X_n)$ is a right-continuous function of t we have

$$P\{n^{1/2}(X_{(m)} - z_0) \leqq z\} = P\{T_n(z n^{-1/2}) \geqq 0\}$$

where $T_n(z) = T_n^*(z; X_1 - z_0, \ldots, X_n - z_0)$. From Liapunov's CLT and since

$$E\, T_n(z n^{-1/2})(\mathrm{Var}\, T_n(z n^{-1/2}))^{1/2} = \sqrt{2}(\delta z_0)^{-1/2} z + \Delta_n(z) \quad \text{with } \Delta_n(z) \xrightarrow[n\to\infty]{} 0,$$

we obtain $P\{T_n(z n^{-1/2}) \geqq 0\} \xrightarrow[n\to\infty]{} \Phi(\sqrt{2}(\delta z_0)^{-1/2} z)$ where Φ is the c.d.f. of the standard normal distribution. □

Proof of (5.1). We shall consider $a > 0$ (say). Define $Z(i,j) = I(\eta_{ij} \leqq a\varepsilon_{ij}) - 1/2$. Then $R_n(a) = \sum Z(\tau_i, \tau_j)$, where the sum extends to all (i,j) in $M \times N = \{1, \ldots, m\} \times \{m+1, \ldots, n\}$ (here) and τ is the vector of antiranks of $(X_1, \ldots, X_n)$. We now have

$$\begin{aligned} R_n(a) = &\sum Z(\tau_i, \tau_j) I(\tau_i \in M, \tau_j \in M) + \sum Z(\tau_i, \tau_j) I(\tau_i \in N, \tau_j \in N) \\ &+ \sum Z(\tau_i, \tau_j) I(\tau_i \in N, \tau_j \in M) + \sum Z(\tau_i, \tau_j) I(\tau_i \in M, \tau_j \in N). \end{aligned}$$

WLOG, we may assume that $(X_1, \ldots, X_n) = (\varepsilon_1, \ldots, \varepsilon_m, 1 + \varepsilon_{m+1}, \ldots, 1 + \varepsilon_n)$.

(i) *The first term.* Let $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)$, $\eta = (\eta_1, \ldots, \eta_n)$ and, for $(i,j) \in M \times N$,

$$A_{ij} = \{u \in \mathbb{R}^n \mid \rho_i, \rho_j \leqq m,$$

where ρ is the vector of antiranks of $(u_1, \ldots, u_m, 1 + u_{m+1}, \ldots, 1 + u_n)\}$. Then, for $(i,j) \in M \times N$,

$$\begin{aligned} E[Z(\tau_i, \tau_j) I(\tau_i \in M, \tau_j \in M)] &= \int_{A_{ij}} E[Z(\tau_i, \tau_j) \mid \varepsilon = u]\, dG(u_1) \ldots dG(u_n) \\ &= \int_{A_{ij}} [F^*(a(u_{\rho_j} - u_{\rho_i})) - 1/2]\, dG(u_1) \ldots dG(u_n) \geqq 0 \end{aligned}$$

since η is independent of ε, τ and because $u \in A_{ij}$ entails $u_{\rho_j} - u_{\rho_i} \geqq 0$.

It follows that the expectation of the first term is nonnegative. Using the same procedure, one can see that this is also the case for the second and third terms. Concerning the latter, note that $(\tau_i, \tau_j) \in N \times M$ entails $\varepsilon_{\tau_j} - \varepsilon_{\tau_i} \geqq 1$.

(ii) *The fourth term.* Let's call it S_n. We have

$$S_n = \sum_{(i,j) \in M \times N} Z(\tau_i, \tau_j) I(\tau_i \in M, \tau_j \in N) = \sum_{(i,j) \in M \times N} Z(i,j) I(R(X_i) \leqq m, R(X_j) > m).$$

Let $(\psi_n)_{n \geqq 1}$ be a sequence of positive real numbers converging to 0 and such that $P\{|X_{(m)} - z_0| > \psi_n\} \xrightarrow[n \to \infty]{} 0$. The existence of such a sequence is guaranteed by lemma (5.2). Now

$$\begin{aligned} E S_n &= E[I(X_{(m)} \in [z_0 - \psi_n, z_0 + \psi_n]) \sum Z(i,j) I(X_i \leqq z_0 - \psi_n, X_j \geqq z_0 + \psi_n)] + o(n^2) \\ &= E[\sum Z(i,j) I(X_i \leqq z_0 - \psi_n, X_j \geqq z_0 + \psi_n)] + o(n^2) \\ &= E[\sum Z(i,j) I(X_i \leqq z_0, X_j \geqq z_0)] + o(n^2) \\ &= n^2 4^{-1} (1+\delta)^{-2} \iint_{A(\delta)} [F^*(a(y-x)) - 1/2] \, dy \, dx + o(n^2) \end{aligned}$$

by the symmetry of F^* about the origin, where

$$\begin{aligned} A(\delta) = &\{(x, y) \in \mathbb{R}^2 \mid x \in [0, 2^{-1}\delta], y \in [2^{-1}\delta, 1+\delta]\} \\ &\cup \{(x, y) \in \mathbb{R}^2 \mid x \in [2^{-1}\delta, 1 + 2^{-1}\delta], y \in [1 + 2^{-1}\delta, 1+\delta]\}. \end{aligned}$$

Our hypothesis on f^* entails that this integral is strictly positive and thus the result follows, with $\kappa > 4^{-1}(1+\delta)^{-2} \iint_{A(\delta)} [F^*(a(y-x)) - 1/2] \, dy \, dx$. □

Note that $\kappa = 0$ if and only if $\delta = 0$.

6. Variance Estimation and Rate of Convergence

Under assumption A and for f smooth enough Δ_{1n} is asymptotically normal with asymptotic variance $48^{-1} d^{-2}$, where

$$d = \lim_{n \to \infty} n^{-2} \sum_{(i,j) \in M \times N} X_{ij}^* \int f^*(ax) \, dG^*(x) \text{ (see Sect. 3).}$$

To construct a confidence interval for a based on Δ_{1n} we need a consistent estimate for $\underset{\sim}{d}$. Such an estimate is provided by

$$\tilde{d}_n = n^{-3/2} [R_n(\Delta_{1n} + n^{-1/2}) - R_n(\Delta_{1n})].$$

The consistency of $\tilde{d}_n$ follows from proposition (2.3) and the fact that $n^{1/2}(\Delta_{1n} - a) = O_P(1)$.

Now write $\sum$ for $\sum_{(i,j)\in M\times N}$ and define

$$\Delta_n(t)=n^{1/2}(Y_n(t)-Y_n(0))=n^{-1}\sum \Delta_{ij}(t)$$

with

$$\Delta_{ij}(t)=I(\eta_{ij}-a\varepsilon_{ij}\leqq tn^{-1/2}(\varepsilon_{ij}+X^*_{ij}))-I(\eta_{ij}-a\varepsilon_{ij}\leqq 0),$$

and

$$Z_n(t)=\Delta_n(t)-E\Delta_n(t).$$

If we assume that assumption A is fulfilled, that

$$\gamma^*_n:=n^{-3}\Big[\sum_{k\in M}\sum_{\substack{j\neq l\\ j,l\in N}} X^*_{kj}X^*_{kl}+\sum_{l\in N}\sum_{\substack{i\neq p\\ i,p\in M}} X^*_{il}X^*_{pl}\Big]\to\gamma^* \quad \text{as } n\to\infty,$$

in which case $\gamma^*>0$ by definition of M, N (Sect. 1), we can state the following results.

(6.1) **Proposition.** *Let f be differentiable with bounded derivative. Then the sequence of processes $(Z_n(t))_{t\in[-r,r]}$, with $r>0$, converges weakly in $D[-r,r]$ (see Billingsley (1968)) to the process $(tZ)_{t\in[-r,r]}$, where Z is a centered normal variable with variance $\sigma_0^2=\frac{1}{4}J+\gamma^*(J_2-J_1^2)$, where $J_1=\int h_a^2(x)\,dx$, $J_2=\int h_a^3(x)\,dx$ and $J=E(\varepsilon_{12}\varepsilon_{13}\int f(y+\varepsilon_{12})f(y+\varepsilon_{13})\,dF(y))$.*

(6.2) **Proposition.** *Let f be differentiable with continuous bounded derivative. Then the sequence of processes*

$$n^{1/2}(Y_n(t)-Y_n(0)-d_n t)_{t\in[-r,r]} \quad \text{with } d_n=n^{-2}\sum X^*_{ij}\int f^*(ax)\,dG^*(x)$$

converges weakly to the process $(tZ+ct^2)_{t\in[-r,r]}$, with Z as in (6.1) and $c=\mu_2\int xf^{*\prime}(ax)\,dG^*(x)$.

Proof of (6.2). Note that by Lebesgue's dominated convergence theorem the assumptions on f imply that the same properties hold for f^*. Now, using the facts that $f^{*\prime}$ is continuous, that the X^*'s are uniformly bounded, that G^* has a compact support and that f^* is symmetric about the origin, we get

$$E\Delta_n(t)=n^{1/2}d_n t+t^2n^{-2}\sum X^*_{ij}\int xf^{*\prime}(ax)\,dG^*(x)+t^2o(1)$$

and since $n^{1/2}(Y_n(t)-Y_n(0)-d_nt)=Z_n(t)+E\Delta_n(t)-n^{1/2}d_nt$, the result follows from (6.1). □

(6.3) **Corollary.** *Under the hypotheses of* (6.2), $n^{1/2}(\tilde{d}_n-d_n)=O_P(1)$. □

Proof of (6.1). It is enough to show (see Billingsley (1968)) that

I. The finite-dimensional distributions of $(Z_n(t))_t$ converge weakly to those of $(tZ)_t$.

II. $\forall\varepsilon>0$, $\lim_{\delta\to 0}\limsup_{n\to\infty} P\{\omega_{Z_n(\cdot)}(\delta)>\varepsilon\}=0$ where for $x\in D[-r,r]$, $\omega_x(\delta)=\sup_{|t-s|\leqq\delta}|x(t)-x(s)|$.

We first approximate $Z_n(t)$ by its Hájek projection $\hat{Z}_n(t)$ with respect to $(\varepsilon_1, \eta_1), \ldots, (\varepsilon_n, \eta_n)$ (see Hájek (1968)). We have

$$\hat{Z}_n(t) = n^{-1} \sum_{k \in M} \sum_{j \in N} \{H_{a,n,t}(\eta_k - a\varepsilon_k - tn^{-1/2}(\varepsilon_k - X^*_{kj})) - H_a(\eta_k - a\varepsilon_k) - E_{kj}(t)\}$$
$$+ n^{-1} \sum_{l \in N} \sum_{i \in M} \{H_a(\eta_l - a\varepsilon_l) - H_{a,n,t}(\eta_l - a\varepsilon_l - tn^{-1/2}(\varepsilon_l + X^*_{il})) - E_{il}(t)\}$$

where $E_{ij}(t)$ stands for $E\Delta_{ij}(t)$ and $H_{a,n,t}$ is the c.d.f. of $\eta_1 - (a + tn^{-1/2})\varepsilon_1$. For f as in (6.1), there are constants $A > 0$, $B > 0$ such that for all $t, s \in [-r, r]$ and for each $n \in \mathbb{N}$

$$\operatorname{Var}(Z_n(t) - Z_n(s)) = (t-s)^2 [\gamma^*_n (J_2 - J_1^2) + 4^{-1} J] + r_n(t, s)$$

with $|r_n(t, s)| \leqq A|t-s| n^{-1/2}$, and

$$\operatorname{Var}(\hat{Z}_n(t) - \hat{Z}_n(s)) = (t-s)^2 [\gamma^*_n (J_2 - J_1^2) + 4^{-1} J] + \hat{r}_n(t, s)$$

with $|\hat{r}_n(t, s)| \leqq B(t-s)^2 n^{-1/2}$.

Since $\operatorname{Var}(Z_n(t) - \hat{Z}_n(t)) = \operatorname{Var}(Z_n(t)) - \operatorname{Var}(\hat{Z}_n(t))$ (see Hájek (1968)), it follows that there is a constant $D > 0$ s.t. for all $t, s \in [-r, r]$ and each n $\operatorname{Var}(X_n(t) - X_n(s)) \leqq D n^{-1/2} |t-s|$, where $X_n(t) = Z_n(t) - \hat{Z}_n(t)$.

In particular, $\operatorname{Var} X_n(t) \leqq D n^{-1/2} |t|$ for each t.

Next, define

$$Z_n = n^{-3/2} \sum_{k \in M} \sum_{j \in N} [\alpha_k + X^*_{kj}(h_a(\eta_k - a\varepsilon_k) - J_1)]$$
$$+ n^{-3/2} \sum_{l \in N} \sum_{i \in M} [\alpha'_l + X^*_{il}(h_a(\eta_l - a\varepsilon_l) - J_1)]$$

where $\alpha_k = \int y f(\eta_k - a\varepsilon_k + ay)\, dG(y) - \varepsilon_k h_a(\eta_k - a\varepsilon_k)$, $\alpha'_k = -\alpha_k$.

If f is bounded (it is the case under the hypothesis of (6.1)) and considering that $E\alpha_k = 0$, $\operatorname{Var}\alpha_k = J$, the X^*'s are uniformly bounded and G has compact support, the Liapunov CLT yields $Z_n \xrightarrow[n \to \infty]{\mathscr{L}} \mathcal{N}(0, \sigma_0^2)$ with σ_0^2 as in (6.1). Moreover, for f as in (6.1),

$$\operatorname{Var}(\hat{Z}_n(t) - tZ_n) = t^2 O(n^{-1}) \qquad \text{for each } t.$$

The convergence of the finite-dimensional distributions follows now from the Cramér-Wold device (e.g. see Billingsley (1968)).

To prove II one can proceed as in Antille (1972). □

Acknowledgements. I wish to thank A. Antille, R. Beran and L. LeCam for the pleasant discussions which took place during the realization of this paper.

References

1. Antille, A.: Linéarité asymptotique d'une statistique de rang. Z. Wahrscheinlichkeitstheor. Verw. Geb. **24**, 309–324 (1972)
2. Bhattacharya, P.K., Chernoff, H., Yang, S.S.: Nonparametric estimation of the slope of a truncated regression. Ann. Stat. **11**, 505–514 (1983)
3. Billingsley, P.: Convergence of probability measures. New York: Wiley 1968
4. Hájek, J.: Asymptotic normality of simple linear rank statistics under alternatives. Ann. Math. Stat. **39**, 325–346 (1968)
5. Jurecková, J.: Nonparametric estimate of regression coefficients. Ann. Math. Stat. **42**, 1328–1338 (1971)
6. Kendall, M., Stuart, A.: The advanced theory of statistics. Vol. 2. London: Griffin 1973
7. Lehmann, E.L.: Theory of point estimation. New York: Wiley 1983
8. Moran, P.A.P.: Estimating structural and functional relationships. J. Multivariate Anal. **1**, 232–255 (1971)
9. Van Eeden, C.: An analogue, for signed-rank statistics of Jurecková's asymptotic linearity theorem for rank statistics. Ann. Math. Stat. **43**, 791–802 (1972)
10. Wald, A.: The fitting of straight lines if both variables are subject to error. Ann. Math. Stat. **11**, 284–300 (1940)

Received October 21, 1987; in revised form September 1, 1988