統計数理研究所　夏の大学院
統計思考院
2012年9月20日

一般化レーマン対立仮説モデルと順位推定

ε

$\mu+\varepsilon$

$\alpha+\beta x+\varepsilon$

三浦良造
統計数理研究所外来研究員
一橋大学名誉教授

: 1. Model.

: Generalized Lehmann's Alternative Model.

: Semi parametric model. $G(x)=h(F(x):\theta)$, $G(x)=h(F(x-\mu):\theta)$

and $G(x)=h(F(y-\alpha-\beta x):\theta)$

: Skew Symmetry . Azzalini (1985). Parametric model. MLE.

:2. Statistics.

: Rank statistics for nonparametric Tests. Lehmann(1953).

: Estimation. Miura (1985,1993)

:3 . Asymptotic Normality.

: Special Construction (Skorokhod's Theorem, or Skorokhod's representation).

: Weak convergence of Empirical Distribution Functions.

: iid case and non-iid case(weakly dependent case.

This work is incomplete yet, but

: Make use of the results in Shao&Yu(1996). Louhichi(2000 and her other papers)).

:4. A direct application to a simple linear regression utilizes an weighted empirical distribution . Wellner&Shorack (1986). We need these for non-iid case.

1. Models

Generalized Lehmann's Alternative Model.

Generalize G(x)= F(x) $^{\theta}$, to G(x)=h(F(x): θ)

Assume F(.) is symmetric about zero.

Assume 関数形h(･: θ) is known, but F(.) unknown.

Definition. Let Θ be an interval in the real line. A function $h(t;\theta)$ for $t \in (0,1)$ and $\theta \in \Theta$ which satisfies the following (1) and (2) is called the generalized Lehmann's alternative model;
(1) $h(0;\theta) = 0$ and $h(1;\theta) = 1$ for any $\theta \in \Theta$. $h(t;\theta)$ is a strictly monotone function of t.
(2) There exists $\theta^* \in \Theta$ such that $h(t;\theta^*) = t$ for $t \in (0,1)$. And for $\theta > \theta'$, $h(t;\theta) < h(t;\theta')$ for all t (or $<$ may be reversed for all t and $\theta > \theta'$).

We shall also call $h(F(\cdot);\theta)$ a generalized Lehmann's alternative model. In terms of random variables, the observations following a generalized Lehmann's

In the one-sample problem, it is not possible to estimate θ for generalized Lehmann's alternative models $h(F;\theta)$, when F is unknown and no restrictions are made on the shape of F. The parameter θ is not even identifiable in that case. Throughout this paper we assume:

$$F \text{ is continuous and } F(x) = 1 - F(-x). \tag{1.1}$$

Also note that (2) in the definition of the generalized Lehmann's alternative model implies

$$h(t;\theta) + h(1-t;\theta) \neq 1 \text{ for } t \in (0,1) \text{ and } \theta \in \Theta - \{\theta^*\}. \tag{1.2}$$

Under (1.1) and (1.2), θ is identifiable and can be estimated.

Examples. Let F and G be d.f.'s which are connected through the generalized Lehmann's alternative model $G = h(F;\theta)$.
(i) If $h(t;\theta) = 1 - (1-t)^{\theta}$ for $\theta \in (0,\infty)$, then

$$\log \Lambda_G = \theta \log \Lambda_F,$$

where Λ_F and Λ_G are cumulative hazard functions corresponding to F and G respectively. This model is the well-known proportional hazards model proposed by Cox (1972).
(ii) Taking $h(t;\theta) = t[(1-t)\theta + t]^{-1}$ for $\theta \in (0,\infty)$ yields the proportional odds model:

$$\frac{G}{1-G} = \theta^{-1}\frac{F}{1-F}.$$

This model has been considered by Ferguson (1967) and in more general regression setting by Pettitt (1984), among others.
The above two models have useful and important applications in survival analysis. Other examples of our model include

applications in survival analysis.

Other examples of our model include

(iii) $h(t;\theta) = (1-\theta)t + \theta t^2$ for $\theta \in [0,1)$ (Contamination),

(iv) $h(t;\theta) = (e^{\theta t} - 1)/(e^{\theta} - 1)$ for $\theta \in (0,\infty)$.

(iii) was considered in Lehmann (1953) and (iv) was found in Ferguson (1967). Both of these are Lehmann alternatives for which the locally most powerful rank test is Wilcoxon.

(v) $h(t;\theta) = t^{\theta}$ for $\theta \in (0,\infty)$ (Lehmann (1953)),

(vi) $h(t;\theta) = \sum_i c_i(\theta)t^i$ with $\sum_i c_i(\theta) = 1$ and $c_i(\theta) \geq 0$ for $\theta \in \Theta$ (Mixture of extremals by a discrete distribution).

(vii) $h(t;\theta) = E(E^{-1}(t) - \log\theta)$ for $\theta \in (0,\infty)$ where E is a known distribution function over the real line. This model can be rewritten as $\psi(X) = \log\theta + \epsilon$ where $X \sim G, \epsilon \sim E$ and $\psi = E^{-1} \circ F$, and includes (i) and (ii).

See Dabrowska, Doksum and Miura (1989) for other examples and Tsukahara (1991) for interesting relations among such models.

(1).Transformation parameter θ,
and
(2). Simultaneous estimation of
Transformation parameter θ and location parameter μ

(1). $X_1,\ldots,X_n$ distributed with $G(x)=h(F(x): \theta)$.

$X_i = \varepsilon_i$,i=1,2,…,n.

(2). $X_1,\ldots,X_n$ distributed with $G(x)=h(F(x- \mu): \theta)$.

$X_i = \mu + \varepsilon_i$,i=1,2,…,n. ε_i follows $G(x)=h(F(x): \theta)$

:iid case. Tsukahara&Miura(1993).

: weakly dependent case (on going)

Generalized Lehmann's Alternative model (GLAM) and skew symmetry:

GLAM

G(x:δ,F)=h(F(x): δ). g(x: δ,F)=h'(F(x): δ)f(x)

Skew symmetry (Azzalini)

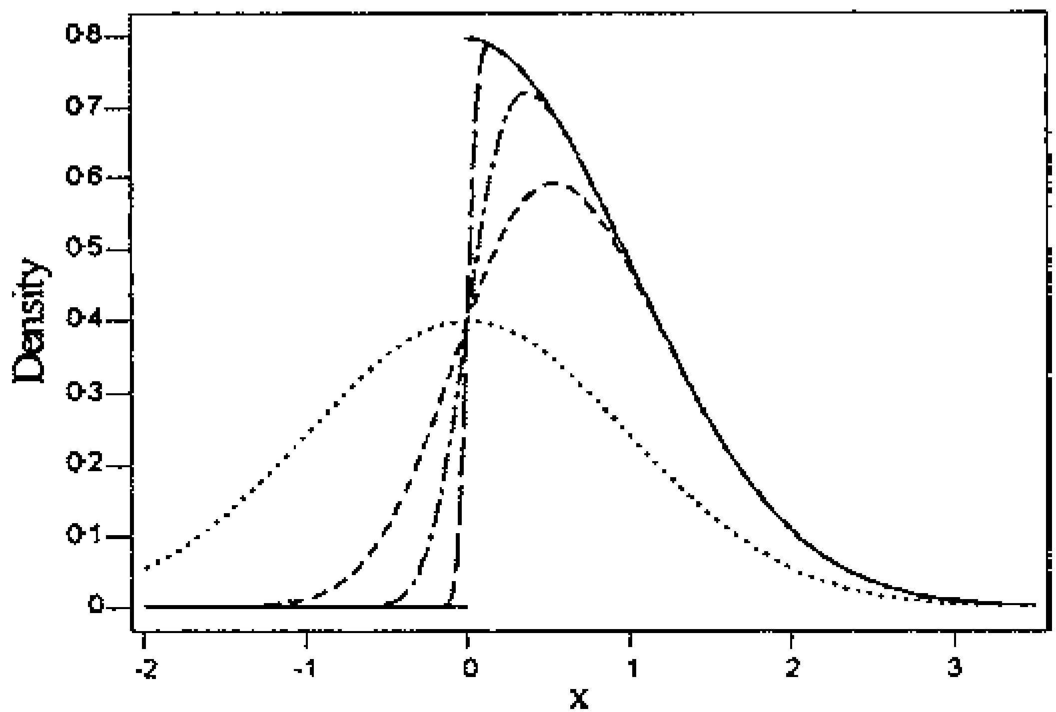
g(x: δ,F)= f(x)2F(δx), or =f(x)2K(δx)

Thus ,2F(δx) or f(x)2K(δx) in skew symmetry corresponds to h'(F(x): δ) in GLAM.

$$h(t:\delta, F, K) = \int_0^t 2K(\delta F^{-1}(u))du$$

and

$$h(F(x):\delta, F, K) = \int_0^{F(x)} 2K(\delta F^{-1}(u))du$$



0·8
0·7
0·6
0·5
0·4
0·3
0·2
0·1
0
Density
-2
-1
0
1
2
3
X

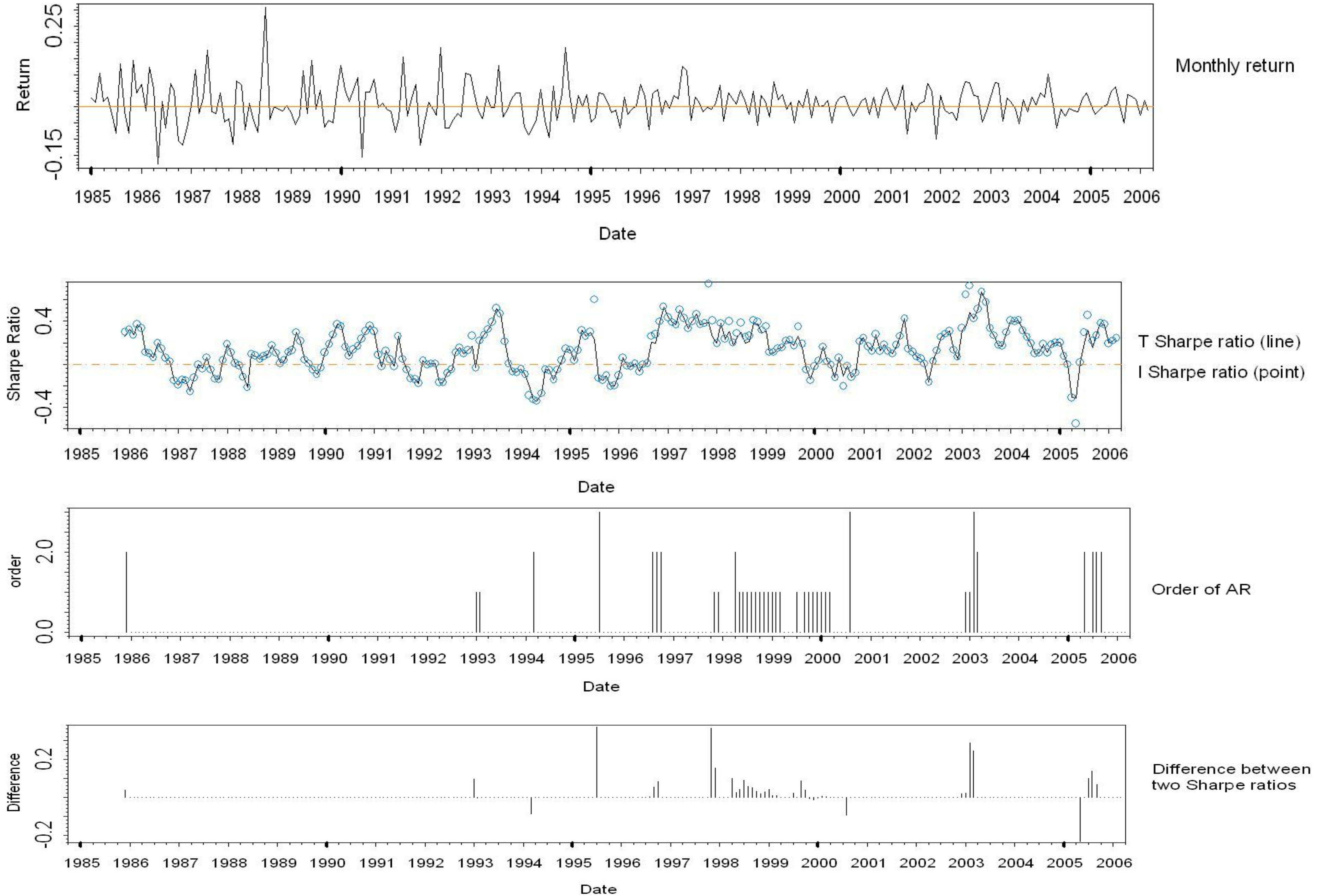


Fig.5-1 A rolling AR structure of a hedge fund's returns with Sharpe ratios

2. Statistics and Estimation for μ and θ in $G(x)=h(F(x-\mu):\theta)$.

*

Statistica Sinica 3(1993), 83-101

ONE-SAMPLE ESTIMATION FOR GENERALIZED LEHMANN'S ALTERNATIVE MODELS

Ryozo Miura and Hideatsu Tsukahara

Hitotsubashi University and University of Illinois at Urbana-Champaign

Abstract: This paper shows that nonparametric estimation of θ for generalized Lehmann's alternative models $h(F;\theta)$ is possible, even in the one-sample problem, when symmetry of the basic distribution function F about zero, $F(x) = 1 - F(-x)$, is assumed. Simultaneous nonparametric estimators of μ and θ for the model $h(F(\cdot - \mu);\theta)$ are also provided under the symmetry of F. The asymptotic normality of these estimators is proved under certain regularity conditions.

(1) Transformation Parameter θ

(1). X_1 ,…, X_n distributed with $G(x)=h(F(x): \theta)$.

$X_i = \varepsilon_i$, $i=1,2,\ldots,n$.

In this section, $X_1, X_2, \ldots, X_n$ are i.i.d. with d.f. $G(x) = h(F(x); \theta_0)$ and θ_0 is to be estimated.

Let $G_n(\cdot)$ be the empirical distribution function of X_i's, that is,

$$G_n(x) \triangleq n^{-1} \sum_{i=1}^{n} I_{[X_i \leq x]},$$

where I_A is an indicator function of a set A and let $\tilde{G}_n(x)$ be a linearized version of G_n: let $X_{(1)} < X_{(2)} < \cdots < X_{(n)}$ be the order statistics of X_i's and define $\tilde{G}_n(x)$ by

$$\tilde{G}_n(x) \triangleq \frac{x + iX_{(i+1)} - (i+1)X_{(i)}}{(n+1)(X_{(i+1)} - X_{(i)})}, \quad x \in \left[X_{(i)}, X_{(i+1)}\right],$$

for $i = 0, 1, \ldots, n$ with $X_{(0)} = X_{(1)} - 1/n$ and $X_{(n+1)} = X_{(n)} + 1/n$. For $i = 1, 2, \ldots, n$, let

$$Z_i(r) \triangleq \tilde{G}_n^{-1}\left(h\left(\frac{i}{n+1}; r\right)\right),$$

and define

$$R_i^+(r) = \text{ the rank of } |Z_i(r)| \text{ among } \{|Z_j(r)| : j = 1, 2, \ldots, n\}.$$

Note that $\tilde{G}_n^{-1}(h(\cdot; \theta_0))$ may be viewed as an estimator of F^{-1} and so $Z_i(\theta_0)$'s can be regarded as an approximation of the ordered sample from F. Also, by

$$G_n(x) \triangleq n^{-1}\sum_{i=1}^{n} I_{[X_i \le x]}, \quad \tilde{G}_n(x) \triangleq \frac{x + iX_{(i+1)} - (i+1)X_{(i)}}{(n+1)(X_{(i+1)} - X_{(i)})}, \quad x \in \left[X_{(i)}, X_{(i+1)}\right]$$

$$Z_i(r) \triangleq \tilde{G}_n^{-1}\left(h\left(\frac{i}{n+1}; r\right)\right), \qquad \int_0^1 J(t)dt = 0.$$

$$R_i^+(r) = \text{ the rank of } |Z_i(r)| \text{ among } \{|Z_j(r)| : j = 1, 2, \ldots, n\}.$$

$$S_n(r) \triangleq \frac{1}{n}\sum_{i:Z_i(r)>0} J\left(\left(1 + \frac{R_i^+(r)}{n+1}\right)/2\right) + \frac{1}{n}\sum_{i:Z_i(r)<0} J\left(\left(1 - \frac{R_i^+(r)}{n+1}\right)/2\right). \quad (2.1)$$

Note that $G(x)=(h(F(x):\theta)$, so that $G^{-1}(t)=F^{-1}(h^{-1}(t:\theta))$,
and $G^{-1}(h(t:\theta))=F^{-1}(h^{-1}(h(t:\theta):\theta))=F^{-1}(t)$.

Thus, $Z_i(\theta)$ is approximately distributed with F.

We need to examine the probability distribution of Ranks
in weakly dependent cases..

Monotonicity in r.

As r tends larger, $Z_i(r)$ are shifted to larger, uniformly in i=1.2,...,n,
and More of $Z_i(r)$ will be positive.

This makes a similar shift to ranks of $|Z_i(r)|$.

$S_n(r)$ as a function of r. It is larger for a large r.

Then the statistic we shall use for inference concerning θ_0 is

$$S_n(r) \triangleq \frac{1}{n} \sum_{i:Z_i(r)>0} J\left(\left(1+\frac{R_i^+(r)}{n+1}\right)/2\right) + \frac{1}{n} \sum_{i:Z_i(r)<0} J\left(\left(1-\frac{R_i^+(r)}{n+1}\right)/2\right). \quad (2.1)$$

If J is symmetric about $\frac{1}{2}$ in the sense that $J(t) = -J(1-t)$, $0 \leq t < 1$, then it is easy to see that

$$S_n(r) = \frac{1}{n} \sum_{i=1}^{n} J^*\left(\frac{R_i^+(r)}{n+1}\right) \mathrm{sign} Z_i(r),$$

where $J^*(t) = J((1+t)/2)$, $0 \leq t \leq 1$. So that the statistic $S_n(r)$ may be regarded as a signed linear rank statistic. The point is that under (1.1) and (1.2) $Z_1(r), Z_2(r), \ldots, Z_n(r)$ are thought of as a sample from a symmetric distribution *only when* $r = \theta_0$, and $S_n(r)$ gives the strongest support to $r = \theta_0$ when it is closest to zero. This makes it possible to estimate θ even in the one-sample situation. Then our estimator $\hat{\theta}_n$ of θ_0 is defined as the value of r which makes $|S_n(r)|$ closest to zero. Such r exists since $S_n(r)$ is nonincreasing in r.

We can write

$$S_n(r) = \int_0^\infty J\left(\frac{1+H_{n,r}(x)}{2}\right) dL_{n,r}(x) + \int_{-\infty}^0 J\left(\frac{1-H_{n,r}(-x)}{2}\right) dL_{n,r}(x),$$

where

$$\begin{aligned} u_n(t) &\triangleq \frac{1}{n}\left(\text{the number of } \left\{i : \frac{i}{n+1} \le t\right\}\right), \quad t \in (0,1), \\ L_{n,r}(x) &\triangleq \frac{1}{n}\left(\text{the number of } \{i : Z_i(r) \le x\}\right) \\ &= u_n\left(h^{-1}(\tilde{G}_n(x); r)\right), \quad x \in \mathbb{R}, \\ H_{n,r}(x) &\triangleq \frac{1}{n+1}\left(\text{the number of } \{i : |Z_i(r)| \le x\}\right), \quad x \in (0,\infty). \end{aligned}$$

We set $H(x) \triangleq F(x) - F(-x)$ for $x \in (0,\infty)$.

Next we shall state the assumptions which are necessary to prove the asymptotic normality of our estimator. Assume that $h(t;\theta)$ is continuously differentiable with respect to t and θ and let

$$h_1(t;\theta) \triangleq \frac{\partial}{\partial t} h(t;\theta), \qquad h_2(t;\theta) \triangleq \frac{\partial}{\partial \theta} h(t;\theta).$$

Let $u(t) = t(1-t)$. Assume, uniformly in θ in a neighborhood of θ_0,

(A.1) $$|J'(t)| \le M\Big[u(h(t;\theta_0))\Big]^{-3/2+\delta}, \quad \text{for } \delta > 0$$

(A.2) $$\frac{1}{|h_1(t;\theta)|} \le M < \infty$$

(A.3) $$|h_2(t;\theta)| \le M\Big[u(h(t;\theta_0))\Big]^{1/2-\delta'}, \quad \text{for } \delta' > 0$$

where M is a universal constant. We require $\rho \triangleq \delta - \delta' > 0$. Further, assume

(A.4) $h_k(t;\theta) \sim h_k(t;\theta_0)$ uniformly in $t \in (0,1)$ as $\theta \to \theta_0$, $(k = 1, 2)$.

Assumptions (A.2)-(A.4) hold for Examples (i) and (v) with $\theta_0 \leq 1$ and Examples (ii)-(iv). Note that (A.2) implies

$$\int_0^1 \Big[h(t;\theta)(1-h(t;\theta))\Big]^{-1+\rho} dt < \infty \tag{2.2}$$

by an easy change of variables. Also note that (A.1) and (A.2) imply

$$|J(t)| \leq M\Big[u(h(t;\theta_0))\Big]^{-1/2+\delta}; \tag{2.3}$$

in fact, letting t_0 be such that $J(t_0) = 0$ and $m \triangleq h(t_0;\theta_0) > 0$, we have

$$\begin{aligned} |J(t)| &= \left|\int_{t_0}^t J'(s)ds\right| \leq M\int_{t_0}^t \Big[u(h(s;\theta_0))\Big]^{-3/2+\delta} ds \\ &\leq m^{-3/2+\delta} M \int_{t_0}^t \Big(1-h(s;\theta_0)\Big)^{-3/2+\delta} h_1(s;\theta_0) ds \\ &\leq M\Big[u(h(t;\theta_0))\Big]^{-1/2+\delta} \end{aligned}$$

for $t > t_0$, and it can be proven similarly for $t < t_0$.

For a function g on I ($I = [0,1]$ or $\mathbb{R}$), define $\|g\| = \sup_{t \in I} |g(t)|$. By Skorohod's representation theorem, there exists a probability space on which a sequence of i.i.d. uniform (0,1) random variables U_{ni}'s and a Brownian bridge U are defined and satisfy

$$\|U_n - U\| \xrightarrow{\text{a.s.}} 0, \quad n \to \infty, \tag{2.4}$$

where

$$\begin{aligned}
\Gamma_n(t) &\triangleq n^{-1} \sum_{i=1}^{n} I_{[U_{ni} \le t]}, \quad t \in (0,1), \\
U_n(t) &\triangleq \sqrt{n}(\Gamma_n(t) - t), \quad t \in (0,1).
\end{aligned}$$

Using these U_{ni}'s, we shall represent the observation as $X_i = G^{-1}(U_{ni})$ for $i = 1, 2, \ldots, n$, which is called the special construction following Shorack and Wellner (1986). We shall then obtain convergence in probability of the estimator, but on the original probability space we can claim convergence in distribution only.

$$S_n(r) = \int_0^\infty J\Big(\frac{1+H_{n,r}(x)}{2}\Big)\,dL_{n,r}(x) + \int_{-\infty}^0 J\Big(\frac{1-H_{n,r}(-x)}{2}\Big)\,dL_{n,r}(x),$$

where

$$\begin{aligned}
u_n(t) &\triangleq \frac{1}{n}\Big(\text{the number of } \Big\{i : \frac{i}{n+1} \le t\Big\}\Big), \quad t \in (0,1), \\
L_{n,r}(x) &\triangleq \frac{1}{n}\Big(\text{the number of } \{i : Z_i(r) \le x\}\Big) \\
&= u_n\Big(h^{-1}(\tilde{G}_n(x); r)\Big), \quad x \in \mathbb{R}, \\
H_{n,r}(x) &\triangleq \frac{1}{n+1}\Big(\text{the number of } \{i : |Z_i(r)| \le x\}\Big), \quad x \in (0,\infty).
\end{aligned}$$

We set $H(x) \triangleq F(x) - F(-x)$ for $x \in (0,\infty)$.

The following lemma is needed.

Lemma 2.1. *Let $r = \theta_0 + b/\sqrt{n}$. Then for the special construction $X_i = G^{-1}(U_{ni})$ and any given positive number B, we have, uniformly in x and $|b| \le B$,*

$$\sqrt{n}\Big[L_{n,r}(x) - F(x)\Big] \xrightarrow{\text{a.s.}} A(F(x)), \quad n \to \infty, \tag{2.5}$$

where

$$A(t) \triangleq \frac{U(h(t;\theta_0))}{h_1(t;\theta_0)} - b \cdot \frac{h_2(t;\theta_0)}{h_1(t;\theta_0)}, \tag{2.6}$$

provided (A.2)-(A.4) *hold.*

Proof. Let $K_{n,r}(x) \triangleq h^{-1}[\Gamma_n(G(x)); r]$. We first prove that, uniformly in x and b,

$$\sqrt{n}\Big|L_{n,r}(x) - K_{n,r}(x)\Big| \xrightarrow{\text{a.s.}} 0 \tag{2.7}$$

as $n \to \infty$. This follows from

$$\begin{aligned}
&\sqrt{n}\Big[K_{n,r}(x) - F(x)\Big] \\
= \;& \sqrt{n}\Big[h^{-1}(\Gamma_n(G(x)); r) - h^{-1}(G(x); r)\Big] + \sqrt{n}\Big[h^{-1}(G(x); r) - h^{-1}(G(x); \theta_0)\Big] \\
= \;& \frac{U_n(G(x))}{h_1(h^{-1}(G^*(x); r); r)} - b \cdot \frac{h_2(h^{-1}(G(x); r^*); r^*)}{h_1(h^{-1}(G(x); r^*); r^*)}.
\end{aligned} \tag{2.8}$$

Theorem 2.1. *Assume that $h(t;\theta)$ is continuously differentiable with respect to t and θ and $\tau(\theta_0) > 0$. Also let the assumptions* (1.1), (A.1)-(A.4) *hold. Then, as $n \longrightarrow \infty$, we have*

$$\sqrt{n}\left(\hat{\theta}_n - \theta_0\right) \xrightarrow{d} N\left(0, \frac{\sigma^2(\theta_0)}{\tau^2(\theta_0)}\right).$$

Proof. Noting that $\int_0^1 J(t)dt = 0$, $\sqrt{n}S_n(r)$ can be expressed as

$$\begin{aligned} & \sqrt{n}\left[\int_0^\infty J\left(\frac{1+H_{n,r}(x)}{2}\right)dL_{n,r}(x) - \int_0^\infty J\left(\frac{1+H(x)}{2}\right)dF(x)\right] \\ + & \sqrt{n}\left[\int_{-\infty}^0 J\left(\frac{1-H_{n,r}(-x)}{2}\right)dL_{n,r}(x) - \int_{-\infty}^0 J\left(\frac{1-H(-x)}{2}\right)dF(x)\right]. \quad (2.9) \end{aligned}$$

Then the first term in (2.9) is decomposed to $\sum_{i=1}^2 B_{in} + \sum_{i=1}^3 C_{in}$ where

$$\begin{aligned} B_{1n} &\triangleq \int J\left(\frac{1+H}{2}\right)d\left\{\sqrt{n}(K_{n,r} - F)\right\}, \\ B_{2n} &\triangleq \frac{1}{2}\int \sqrt{n}(H_{n,r} - H)J'\left(\frac{1+H}{2}\right)dF, \\ C_{1n} &\triangleq \int J\left(\frac{1+H_{n,r}}{2}\right)d\left\{\sqrt{n}(L_{n,r} - K_{n,r})\right\}, \\ C_{2n} &\triangleq \frac{1}{2}\int \sqrt{n}\left(H_{n,r} - H\right)J'\left(\frac{1+H}{2}\right)d\left(K_{n,r} - F\right), \\ C_{3n} &\triangleq \sqrt{n}\int \left[J\left(\frac{1+H_{n,r}}{2}\right) - J\left(\frac{1+H}{2}\right) - \frac{1}{2}\left(H_{n,r} - H\right)J'\left(\frac{1+H}{2}\right)\right]dK_{n,r}. \end{aligned}$$

Note that $(1+H)/2 = F$ due to the symmetry of F, which we shall use repeatedly without mention.

to prove that $\sqrt{n}[K_{n,r}(x) - F(x)]$ converges to the process on the right-hand side of (2.5). Now we have

$$
\begin{aligned}
&\sqrt{n}\Big[K_{n,r}(x) - F(x)\Big] \\
= \;& \sqrt{n}\Big[h^{-1}(\Gamma_n(G(x)); r) - h^{-1}(G(x); r)\Big] + \sqrt{n}\Big[h^{-1}(G(x); r) - h^{-1}(G(x); \theta_0)\Big] \\
= \;& \frac{U_n(G(x))}{h_1(h^{-1}(G^*(x); r); r)} - b \cdot \frac{h_2(h^{-1}(G(x); r^*); r^*)}{h_1(h^{-1}(G(x); r^*); r^*)}.
\end{aligned}
\tag{2.8}
$$

Here $G^*(x)$ is a random function taking values between $G(x)$ and $\Gamma_n(G(x))$, and r^* lies between θ_0 and r. Then

$$
\begin{aligned}
&\left\| \frac{U_n(G(x))}{h_1(h^{-1}(G^*(x); r); r)} - \frac{U(G(x))}{h_1(F(x); \theta_0)} \right\| \\
\leq \;& M \left\| \frac{h_1(F(x); \theta_0)}{h_1(h^{-1}(G^*(x); r); r)} - 1 \right\| \cdot \Big\| U_n(G(x)) \Big\| + M \Big\| U_n(G(x)) - U(G(x)) \Big\|.
\end{aligned}
$$

It follows from Glivenko-Cantelli theorem that $\|G^*(x) - G(x)\| \xrightarrow{\text{a.s.}} 0$. Also $r^* \to \theta_0$ uniformly in b as $n \to \infty$. Thus the first term converges almost surely to 0 by virtue of (A.4) and $\|U_n(G(x))\| \stackrel{\text{a.s.}}{=} O(1)$, which is an easy consequence of (2.4). Next by (2.4) and (A.2), we see that the second term converges almost surely to 0. Furthermore it follows from (A.2)-(A.4) that

$$\left\| \frac{h_2(h^{-1}(G(x); r^*); r^*)}{h_1(h^{-1}(G(x); r^*); r^*)} - \frac{h_2(F(x); \theta_0)}{h_1(F(x); \theta_0)} \right\| \longrightarrow 0.$$

Therefore (2.8) converges almost surely to

$$\frac{U(G(x))}{h_1(F(x); \theta_0)} - b \cdot \frac{h_2(F(x); \theta_0)}{h_1(F(x); \theta_0)},$$

uniformly in x and b, which completes the proof of the lemma.

Thus

$$\sum_{i=1}^{2} B_{in} \xrightarrow{P} -\frac{1}{2}\int_{\frac{1}{2}}^{1}\left[\frac{U(h(t;\theta_0))}{h_1(t;\theta_0)}+\frac{U(h(1-t;\theta_0))}{h_1(1-t;\theta_0)}\right]dJ(t)$$
$$+\frac{b}{2}\int_{\frac{1}{2}}^{1}\left[\frac{h_2(t;\theta_0)}{h_1(t;\theta_0)}+\frac{h_2(1-t;\theta_0)}{h_1(1-t;\theta_0)}\right]dJ(t)+\lambda(1/2). \quad (2.13)$$

Next we show that $\sum_{i=1}^{3} C_{in} \xrightarrow{P} 0$. For C_{1n}, note that $H_{n,r} \leq n/(n+1)$.

It can be seen in the similar way that the second term in (2.9) converges in probability to

$$-\frac{1}{2}\int_{0}^{\frac{1}{2}}\left[\frac{U(h(t;\theta_0))}{h_1(t;\theta_0)}+\frac{U(h(1-t;\theta_0))}{h_1(1-t;\theta_0)}\right]dJ(t)$$
$$+\frac{b}{2}\int_{0}^{\frac{1}{2}}\left[\frac{h_2(t;\theta_0)}{h_1(t;\theta_0)}+\frac{h_2(1-t;\theta_0)}{h_1(1-t;\theta_0)}\right]dJ(t)-\lambda(1/2).$$

we obtain asymptotic linearity: for any $B>0$

$$\sup_{|b|\leq B}\left|\sqrt{n}S_n(r)+\frac{1}{2}T-\frac{1}{2}b\tau(\theta_0)\right| \xrightarrow{P} 0, \quad (2.14)$$

where

$$T \triangleq \int_0^1 \left[\frac{U(h(t;\theta_0))}{h_1(t;\theta_0)} + \frac{U(h(1-t;\theta_0))}{h_1(1-t;\theta_0)}\right] dJ(t).$$

Now let $\epsilon > 0$ be a given number small enough to satisfy $\epsilon < \tau(\theta_0)/2$. Take $B_\epsilon > 1$ so large that

$$P\left\{|T| > \frac{B_\epsilon \tau(\theta_0)}{2}\right\} < \frac{\epsilon}{2}.$$

By asymptotic linearity (2.14), there exists an N_ϵ such that for all $n \geq N_\epsilon$,

$$P\left\{\sup_{|b|\leq B_\epsilon}\left|\sqrt{n}S_n(r) + \frac{1}{2}T - \frac{b}{2}\tau(\theta_0)\right| > \epsilon\right\} < \frac{\epsilon}{2}.$$

Thus for all $n \geq N_\epsilon$, any value b_n of b which minimizes $|\sqrt{n}S_n(r)| = |\sqrt{n}S_n(\theta_0 + b/\sqrt{n})|$ lies in $[-B_\epsilon, B_\epsilon]$ and it follows that

$$|b_n - T/\tau(\theta_0)| < \epsilon/\tau(\theta_0)$$

with probability exceeding $1-\epsilon$ (note that $T/\tau(\theta_0)$ minimizes $|-T/2 + b\tau(\theta_0)/2|$). Noting that $\sqrt{n}(\hat{\theta}_n - \theta_0)$ is a value of b which minimizes $|\sqrt{n}S_n(r)|$, we obtain

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{P} \frac{T}{\tau(\theta_0)}.$$

iid から weakly dependent へ

stationary process : strongly mixing or associated

For a function g on I ($I = [0,1]$ or $\mathbb{R}$), define $\|g\| = \sup_{t \in I} |g(t)|$. By Skorohod's representation theorem, there exists a probability space on which a sequence of i.i.d. uniform (0,1) random variables U_{ni}'s and a Brownian bridge U are defined and satisfy

$$\|U_n - U\| \xrightarrow{\text{a.s.}} 0, \quad n \to \infty, \tag{2.4}$$

where

$$\begin{aligned} \Gamma_n(t) &\triangleq n^{-1} \sum_{i=1}^{n} I_{[U_{ni} \le t]}, \quad t \in (0,1), \\ U_n(t) &\triangleq \sqrt{n}(\Gamma_n(t) - t), \quad t \in (0,1). \end{aligned}$$

The Annals of Probability
1996, Vol. 24, No. 4, 2098–2127

WEAK CONVERGENCE FOR WEIGHTED EMPIRICAL PROCESSES OF DEPENDENT SEQUENCES

BY QI-MAN SHAO[1] AND HAO YU[2]

National University of Singapore and University of Western Ontario

While the problem of weak convergence for weighted empirical processes of independent sequences has been intensively studied in recent years, there are only a few studies concerned with the counterpart for dependent sequences [cf. Yu (1993a)]. In the latter case the limit process is changed from being a Brownian bridge due to the appearance of covariances among observations. Namely, under certain conditions, we have

$$\alpha_n(\cdot) \to_{\mathscr{D}} B^*(\cdot) \quad \text{in } D[0,1], \tag{1.7}$$

where $B^*(\cdot)$ is a zero-mean Gaussian process specified by $B^*(0) = B^*(1) = 1$ and

$$\begin{aligned} EB^*(s)B^*(t) = s \wedge t - st & \\ + \sum_{k=2}^{\infty} \{ \mathrm{Cov}(I(U_1 \le s), I(U_k \le t)) & \\ + \mathrm{Cov}(I(U_k \le s), I(U_1 \le t)) \}. & \end{aligned} \tag{1.8}$$

iid case

Also Pyke and Shorack (1968) show that $\|(U_n(t) - U(t))/q(t)\| \xrightarrow{P} 0$ for $q(t) = [u(t)]^{1/2-\delta}$ for some $\delta > 0$. Note that this implies

Non-iid case (weakly dependent): α-mixing or associated.

Proposition 2.2 (Shao and Yu) *Suppose that $(\xi_n)_{n\in\mathbb{N}}$ is a strictly stationary sequence of $U(0,1)$ random variables which is also strongly mixing with the mixing rate*

$$\alpha(n) = O(n^{-\theta-\eta}) \quad \text{for some } \theta \geq 1+\sqrt{2}, \text{ and } \eta > 0.$$

Then $\mathbb{U}_n/q \xrightarrow{\mathcal{L}} \mathbb{U}/q$ in $D[0,1]$, where $q(u) := [u(1-u)]^{1/2-1/(2\theta)}$ and $\mathbb{U}$ is a Gaussian process with $\mathbb{U}(0) = \mathbb{U}(1) = 0$ and

$$\mathrm{E}[\mathbb{U}(u)] = 0, \quad \mathrm{E}[\mathbb{U}(u)\mathbb{U}(v)] = \sigma(u,v).$$

$$\alpha(n) = \sup_{k\geq 1} \alpha(\mathscr{F}_1^k, \mathscr{F}_{n+k}^\infty). \qquad \mathscr{F}_n^m = \sigma(X_i, n \leq i \leq m)$$

$$\alpha(\mathscr{F}_1, \mathscr{F}_2) = \sup_{\substack{A\in\mathscr{F}_1 \\ B\in\mathscr{F}_2}} |P(A\cap B) - P(A)\,P(B)|.$$

$1+\sqrt{2} \leq \theta$

$2(1+\sqrt{2}) \leq 2\theta$, ==> $0 < \dfrac{1}{2\theta} \leq \dfrac{1}{2(1+\sqrt{2})} = 0.208......$

Ann. Inst. Henri Poincaré, Probabilités et Statistiques 36, 5 (2000) 547–567

Weak convergence for empirical processes of associated sequences

by

Sana LOUHICHI [1]

UMR C 8628 CNRS, Université de Paris-Sud, Bât. 425, Mathématiques,
F-91405 Orsay Cedex, France

Manuscript received in 29 October 1998, revised 22 March 2000

2. MAIN RESULT AND APPLICATION

THEOREM 1. – *Let $(X_n)_{n\in\mathbb{Z}}$ be a stationary associated sequence with continuous marginal distribution F. Assume that, for $n \in \mathbb{N}^*$,*

$$\operatorname{Cov}\big(F(X_1), F(X_n)\big) = \mathcal{O}\big(n^{-b}\big), \quad \textit{for } b > 4. \tag{10}$$

Then

$$G_n(.) \to G(.) \quad \textit{in } D[-\infty, +\infty],$$

This argument in the weak dependent case goes very similar (almost the same) as in the case of iid.

The only difference would be

(1) assumptions on the power, and

(2) asymptotic variance.

(3) Need to examine the probability distribution of rank vectors if it is well asymptotically uniform over (0,1),or distributed to satisfy that sum of rank score to be zero:

$$S_n(r) \triangleq \frac{1}{n} \sum_{i:Z_i(r)>0} J\left(\left(1+\frac{R_i^+(r)}{n+1}\right)/2\right) + \frac{1}{n} \sum_{i:Z_i(r)<0} J\left(\left(1-\frac{R_i^+(r)}{n+1}\right)/2\right). \quad (2.1)$$

(2). θ and μ . Simultaneously.

(2). $X_1,\ldots,X_n$ distributed with $G(x)=h(F(x-\mu)\colon \theta)$.

$X_i = \mu + \varepsilon_i$,$i=1,2,\ldots,n$. ε_i follows $G(x)=h(F(x)\colon \theta)$

$S_{1,n}(r, q)$ and $S_{2,n}(r, q)$

$$G_n(x) \triangleq n^{-1} \sum_{i=1}^{n} I_{[X_i \le x]},$$

$$\tilde{G}_n(x) \triangleq \frac{x + iX_{(i+1)} - (i+1)X_{(i)}}{(n+1)(X_{(i+1)} - X_{(i)})}, \quad x \in \left[X_{(i)}, X_{(i+1)}\right]$$

$$Z_i(r) \triangleq \tilde{G}_n^{-1}\left(h\left(\frac{i}{n+1}; r\right)\right),$$ $Z_i(r)$ are aproximately distributed arround μ with F(• -μ).

3. Simultaneous Estimation of μ and θ

In this section, let $X_1, X_2, \ldots, X_n$ be i.i.d. with d.f. $G(x) = h(F(x - \mu_0); \theta_0)$. The parameters μ_0 and θ_0 are both unknown and are to be estimated simultaneously.

Let $Z_i(r)$ be as in Section 2 and define

$$R_i^+(r, q) = \left(\text{the number of } \{j : |Z_j(r) - q| \le |Z_i(r) - q|\}\right).$$

Note. For a value of r close to the true parameter value, $Z_i(r), i = 1, 2, \ldots, n$ are approximately symmetrically distributed arround zero. Hence this q is a generalization of Hodge Lehmann estimate of center of symmetry. (H-L from a linear score function.)

In this section assume that F has a bounded continuous density f. Let $J_1(\cdot)$ and $J_2(\cdot)$ be the score function used for estimation of θ and μ respectively. $J_1(\cdot)$ and $J_2(\cdot)$ satisfy the conditions for the score functions in Section 2. In addition, $J_1(\cdot)$ and $J_2(\cdot)$ are assumed different enough to satisfy

$$\frac{\int_0^1 \left[\frac{h_2(t;\theta_0)}{h_1(t;\theta_0)} + \frac{h_2(1-t;\theta_0)}{h_1(1-t;\theta_0)}\right] dJ_1(t)}{\int_0^1 \left[\frac{h_2(t;\theta_0)}{h_1(t;\theta_0)} + \frac{h_2(1-t;\theta_0)}{h_1(1-t;\theta_0)}\right] dJ_2(t)} \neq \frac{\int_0^1 f(F^{-1}(t))dJ_1(t)}{\int_0^1 f(F^{-1}(t))dJ_2(t)}. \tag{3.1}$$

The rank statistics for the simultaneous inference of μ and θ are defined as follows:

$$S_{1n}(r,q) \triangleq \frac{1}{n} \sum_{i:Z_i(r)>q} J_1\left(\left(1+\frac{R_i^+(r,q)}{n+1}\right)/2\right) + \frac{1}{n} \sum_{i:Z_i(r)\leq q} J_1\left(\left(1-\frac{R_i^+(r,q)}{n+1}\right)/2\right), \tag{3.2}$$

and

$$S_{2n}(r,q) \triangleq \frac{1}{n} \sum_{i:Z_i(r)>q} J_2\left(\left(1+\frac{R_i^+(r,q)}{n+1}\right)/2\right) + \frac{1}{n} \sum_{i:Z_i(r)\leq q} J_2\left(\left(1-\frac{R_i^+(r,q)}{n+1}\right)/2\right). \tag{3.3}$$

Our estimators of μ and θ are derived from the simultaneous equations $S_{1n}(r,q) \approx 0$ and $S_{2n}(r,q) \approx 0$. Define

$$D_n \triangleq \left\{ (r,q) : \sum_{k=1}^{2} |S_{kn}(r,q)| = \min \right\}.$$

$D_n \subset \Theta \times \mathbf{R}$ is not empty for all $X_1, X_2, \ldots, X_n$ since $S_{kn}(r,q)$, as a function of r and q with fixed $X_1, X_2, \ldots, X_n$, takes on a finite number of different values. $S_{kn}(r,q)$, $(k = 1, 2)$ are nonincreasing in each coordinate r and q separately, but it does not ensure the convexity of D_n, which may be used to determine the estimators uniquely. Our estimator $(\hat{\theta}_n, \hat{\mu}_n)$ is thus defined to be any point of D_n. Since $(\hat{\theta}_n, \hat{\mu}_n)$ may not be unique, there may be some arbitrariness in this definition. But, as will turn out in Theorem 3.2 below, all points in D_n are asymptotically equivalent; so, for large n, it will not matter much how $(\hat{\theta}_n, \hat{\mu}_n)$ chosen.

To investigate the asymptotic behavior of S_{kn}, we assume, in addition to (A.1) with J replaced by J_k and (A.2)-(A.4),

$$|J_k'(t)| \leq M[u(t)]^{-1+\delta}, \quad \delta > 0. \tag{A.5}$$

We also introduce the following notation: let $r = \theta_0 + b_1/\sqrt{n}$, $q = \mu_0 + b_2/\sqrt{n}$ and

$$\boldsymbol{S}_n(r,q) \triangleq \Big(S_{1n}(r,q), S_{2n}(r,q)\Big)', \quad \boldsymbol{b} \triangleq (b_1, b_2)'.$$

Furthermore, for $k = 1, 2$

$$T_k \triangleq \int_0^1 \left\{ \frac{U(h(t;\theta_0))}{h_1(t;\theta_0)} + \frac{U(h(1-t;\theta_0))}{h_1(1-t;\theta_0)} \right\} dJ_k(t),$$

and set $\boldsymbol{T} \triangleq (T_1, T_2)'$. Let $D = [d_{kl}]$ denote a 2×2 matrix, where

$$\begin{aligned}
d_{k1} &\triangleq \int_0^1 \left\{ \frac{h_2(t;\theta_0)}{h_1(t;\theta_0)} + \frac{h_2(1-t;\theta_0)}{h_1(1-t;\theta_0)} \right\} dJ_k(t), \\
d_{k2} &\triangleq -2\int_0^1 f(F^{-1}(t))dJ_k(t) \quad (k = 1, 2).
\end{aligned}$$

Note that D is nonsingular because of (3.1). Then we have the following asymptotic linearity result.

Theorem 3.1. *Suppose that F has a bounded continuous density f and that (A.1) with J replaced by J_k and (A.2)-(A.5) all hold. Then*

$$\max_{k=1,2} \sup_{|b_k| \le B} \left| \sqrt{n} S_{kn}(r, q) + \frac{1}{2} T_k - \frac{1}{2}(d_{k1} b_1 + d_{k2} b_2) \right| \xrightarrow{P} 0, \quad n \to \infty, \qquad (3.4)$$

for each $0 < B < \infty$.

Using matrix notation, express the relation (3.4) as

$$\sup_{|b_k| \le B} \left| \sqrt{n} \boldsymbol{S}_n(r, q) + \frac{1}{2} \boldsymbol{T} - \frac{1}{2} D\boldsymbol{b} \right| \xrightarrow{P} 0, \quad n \to \infty. \qquad (3.5)$$

Theorem 3.2. *Suppose that all the conditions of Theorem 3.1 are satisfied. Then each point of D_n is asymptotically normal $N(0, D^{-1}\Sigma(D^{-1})')$, that is,*

$$\sqrt{n} \begin{pmatrix} \hat{\theta}_n - \theta_0 \\ \hat{\mu}_n - \mu_0 \end{pmatrix} \xrightarrow{d} N\Big(0, D^{-1}\Sigma(D^{-1})'\Big),$$

as $n \to \infty$.

A Simple Linear Regression Model.

(On-going trial)

$Y=\alpha+\beta x+\varepsilon$. Where $\varepsilon \approx G(.)=h(F(.):\theta)$

$F(.)$ symmetric around zero.

$Y_i=\alpha+\beta x_i+\varepsilon_i \ ,i=1,2,\ldots,n$

$\varepsilon_i \ ,i=1,2,\ldots,n$:weakly dependent(associated),

And distributed with $G(.)=h(F(.):\theta)$

注意:最小二乗法との違い。

順位統計量から導かれる回帰係数βの推定は、回帰モデルの偶然誤差項が対称性を持つことなど分布形の仮定を置かない。また、αの推定は、βを推定した後に行う。

図4　回　帰　直　線

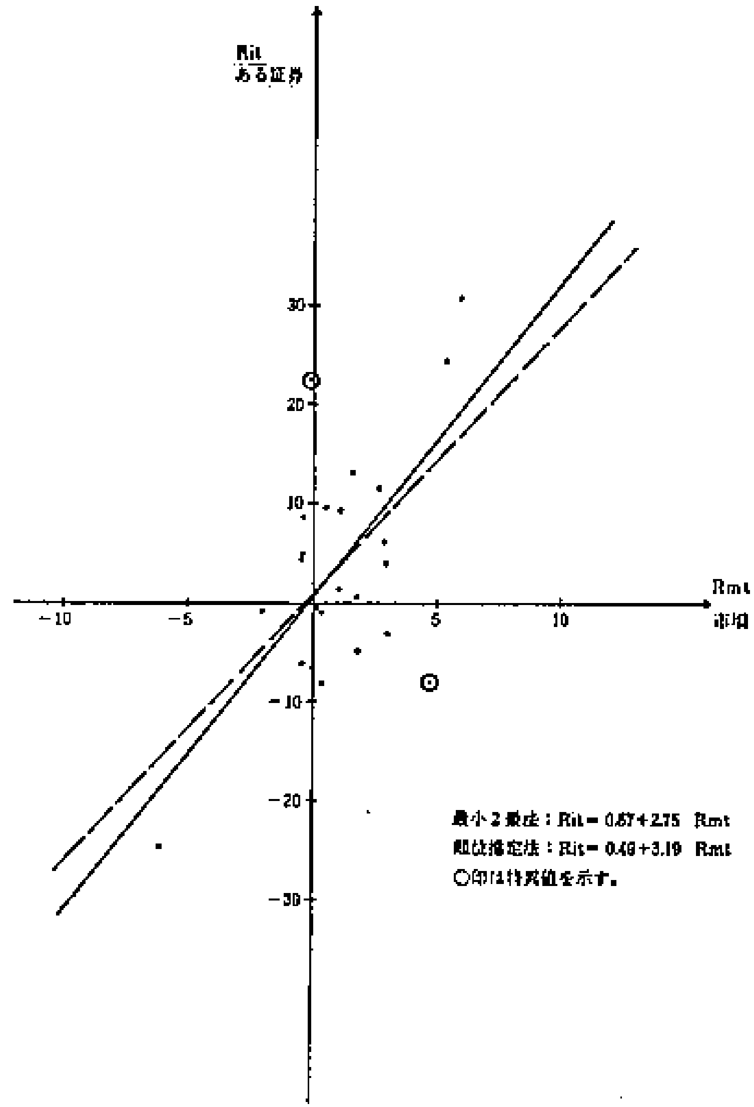

Rit
ある証券
30
20
10
−10
−20
−30
−10
−5
5
10
Rmt
市場
最小2乗法：Rit＝0.67＋2.75　Rmt
頑健推定法：Rit＝0.40＋3.19　Rmt
○印は特異値を示す。

図 9

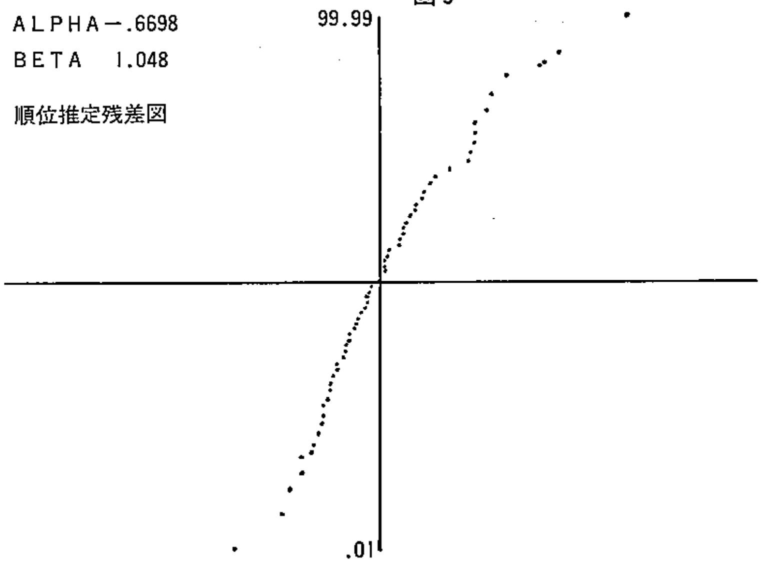


図 8

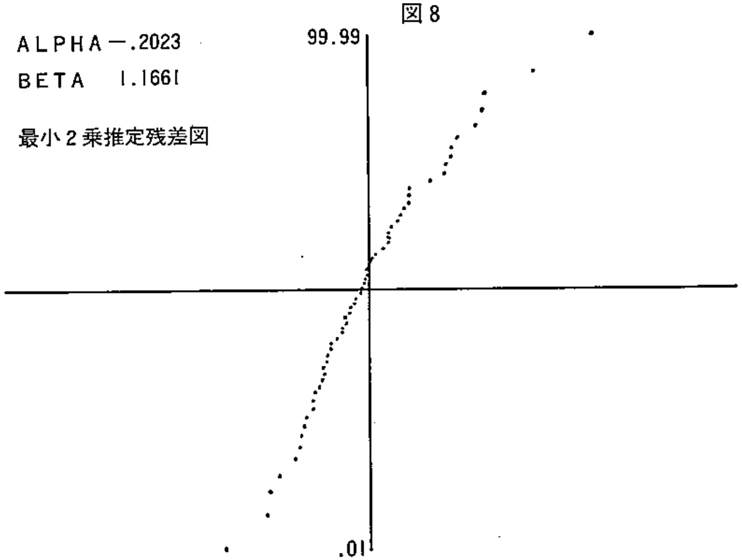
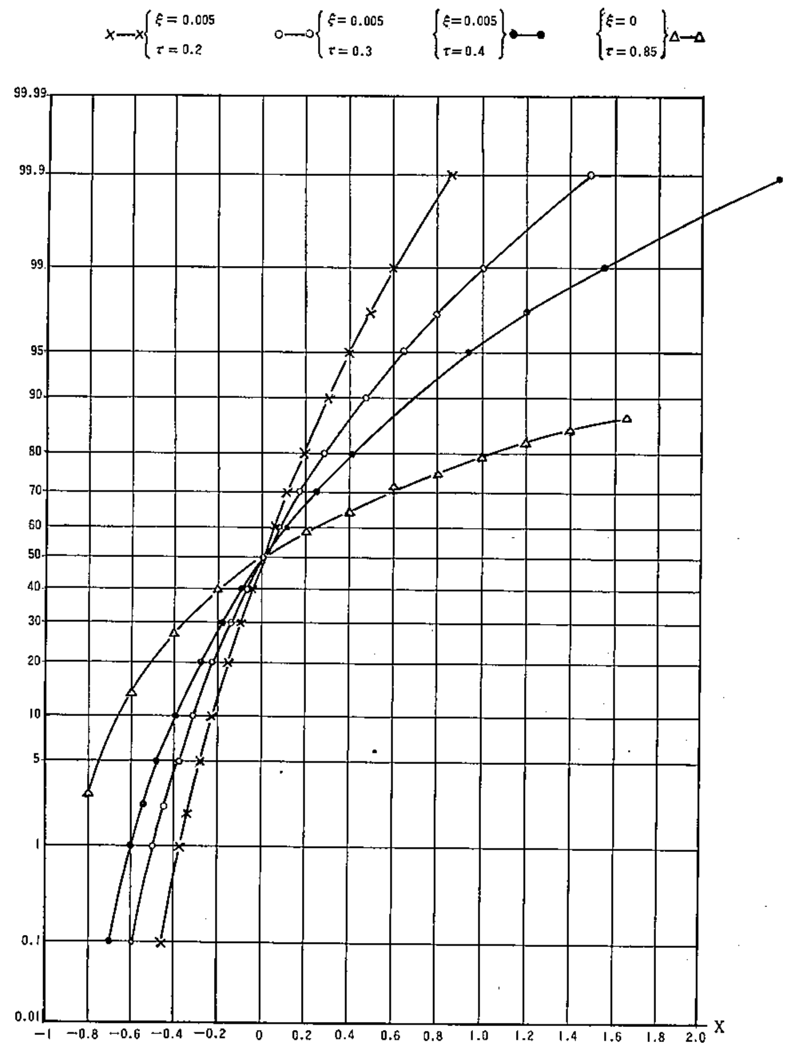


図12 $G(X)=\Phi\left(\frac{\log(X+1)-\xi}{\tau}\right)$ の正規確率のプロット



Motivations.

Data in Financial Engineering:

Time series data of Hedge Fund return .

: In a usual regression where returns of portfolios or of stocks are regressed on indices (stock index, for example), the residuals are rather implicitly regarded as iid. They may be almost so, but they may not be so.

In practice, they do not care about estimation error? (In econometrics, or in academics, they surely care about it.)

:*

: When hedge fund returns are regressed on market indices as in the following, the residuals do not seem iid, but may show some dependence, such as autoregressive.

: This kind of regression is often done in a search for possibilities of Hedge fund replication.

Estimate β , θ and α simultaneously.
$S_{1,n}$ (b, r) for β, or $S_{1,n}$ (b)
$S_{2,n}$ (b, r, a) for θ, and
$S_{3,n}$ (b, r, a) for α.

Mathematics.

(1) use the convergence of weighted empirical distribution functions, {Y- βx} empirical distribution function that is described in Wellner & Shorack(1986) and others.

(2) follow Hajek& Sidak(1967) and Jureckova (1971) with Jaeckel (1972), Also Sen& Jureckova(199?) and Wellner &Shorack (1986) with Shao&Yu (1996) and Louhichi(2000), in order to prove asymptotic normality of linear rank statistics in our weakly dependent case. We are on the way to prove fundamental theorems for distributions for rank vectors.

Can we do the same for a regression model as for a location only ?

Can we base everything on symmetry of F with empirical distribution functions ?

Let $e_i(\beta) = Y_i - \beta x_i$, i=1,2,..., n.

(1).Note that $e_i(\beta)$ is distributed with G(x-α)=h(F(x-α):θ_0).

(2).Note also that $\mathrm{E}[e_i(\beta)] = \alpha + \int_{-\infty}^{\infty} x d\mathrm{h}(\mathrm{F(x)}:\theta_0)$

$$=\alpha + \int_0^1 F^{-1}(t) d\mathrm{h}(\mathrm{t}:\theta_0)$$

$$=\alpha + \mathrm{m}(\theta_0)$$

where the second term m(θ_0) is non-zero unless $\theta_0 = \theta^*$.

(3).Note that one can estimate β based on rank statistics

with ranks of $e_i(b), i = 1, 2, ..., n$ without simultaneously estimate

α and θ_0. This is an important feature of rank statistics in

linear regression models and estimates derived from rank statistics.

Here, we will estimate α and θ_0.

Let $\mathrm{E}_n(x:\mathrm{b}) = \dfrac{1}{n+1}\sum_{i=1}^{n} I\{e_i(\mathrm{b}) \le x\}$.

Let $\mathrm{E}_{\mathrm{n}}(.)$ be a piecewise linear version of $\mathrm{E}_n(.)$.

(4).Note that $\mathrm{E}^{-1}(\mathrm{t}) = \alpha + \mathrm{F}^{-1}(\mathrm{h}^{-1}(\mathrm{t}:\theta_0))$ or $\mathrm{E}^{-1}(\mathrm{U}) \triangleq \alpha + \mathrm{F}^{-1}(\mathrm{h}^{-1}(\mathrm{U}:\theta_0))$

and that

$\mathrm{E}^{-1}(\mathrm{h}(\mathrm{t}:\theta_0)) = \alpha + \mathrm{F}^{-1}(\mathrm{t})$ and

$\mathrm{E}^{-1}(\mathrm{h}(\mathrm{U}:\theta_0))$ has a distribution $\mathrm{F}(\mathrm{x}-\alpha)$,

for $\mathrm{E}(\mathrm{x}) = \mathrm{P}\{e_i(\beta) < x\}$ and an uniform random variable U on [0,1].

Now, let $\mathrm{Z}_i(\mathrm{b}, \mathrm{r}) = \widetilde{\mathrm{E}_{\mathrm{n}}^{-1}}(\mathrm{h}(\dfrac{\mathrm{i}}{\mathrm{n}+1}:\mathrm{r}))$, i=1,2,..., n.

Then, these are expected to distribute symmetrically arround α; approximately with $\mathrm{F}(\mathrm{x}-\alpha)$.

(5).Note that expectation $\mathrm{E}[\mathrm{U}_{(i)}] = \dfrac{i}{n+1}$ and $\mathrm{U}_{(i)} - \dfrac{i}{n+1}$ converges to zero.

(6).Note that $\mathrm{m}(\theta)$ can be estimated by $\widehat{\mathrm{m}(\theta)}$.

and let $R_i(b, r)$ = rank of $Z_i(b, r)$ among $\{Z_j(b, r), j=1,2,...,n\}$

$$= \sum_{j=1}^{n} I\{Z_j(b, r) \le Z_i(b, r)\}.$$

let $R_i^+(b, r, a)$ = rank of $|Z_i(b, r) - a|$ among $\{|Z_j(b, r) - a|, j=1,2,...,n\}$

$$= \sum_{j=1}^{n} I\{|Z_j(b, r) - a| \le |Z_i(b, r) - a|\}.$$

$(\beta, \theta, \alpha) \leftarrow (b,r,a)$

$$\beta: \quad S_{1,n}(b,r) = \frac{1}{n}\sum_{i=1}^{n} J_1(R_i(b, r))(x_i - \bar{x})$$

(7).Note that this kind of rank statistics was given

in Jureckova(1969,1971 and others) and Jaeckel(1972).

(8).Note also that there is another choice of using this without r.

$$\theta: \quad S_{2,n}(b,r,a) = \frac{1}{n}\sum_{i:\ zi(b,r)>a} J_2\left(\frac{1}{2} + \frac{R_i^+(b, r, a)}{2}\right) + \frac{1}{n}\sum_{i:\ zi(b,r)\le a} J_2\left(\frac{1}{2} + \frac{R_i^+(b, r, a)}{2}\right)$$

$$\alpha: \quad S_{1,n}(b,r,a) = \frac{1}{n}\sum_{i:\ zi(b,r)>a} J_3\left(\frac{1}{2} + \frac{R_i^+(b, r, a)}{2}\right) + \frac{1}{n}\sum_{i:\ zi(b,r)\le a} J_3\left(\frac{1}{2} + \frac{R_i^+(b, r, a)}{2}\right)$$

(9).Note that these two rank statistics are similarly defined to

Tsukahara and Miura(1993); i.e. a version for a simple linear regression model.

i.i.d. case.

Jureckova (1971,1969) and Jaeckel (1972).

Hajek & Sidak [Theory of Rank Test] 1967.

Estimation of β based on rank statistics, where the distribution of ε can be asymmetric.

Note that β can be estimated separately from μ and θ.

We take advantage of this nature .

Plan of a proof.

One can use their result for estimation for β.

Then, in order to estimate μ and θ, one may use the estimated β in $Y_i - \beta x_i$, and work on empirical distribution function with the above X_i replaced by $Y_i - \beta x_i$, where estimated β converges to a true β.

Note that for iid case this can be done by conbining the results in the Jureckova(1971) and Tsukahara&Miura(1993). It will be a new result.

Dispersion measure D(Y-βx).

Estimator β ; minimizes D.

similar to LSE, but different in replacing (…) with ranks.

D' is a non-decreasing step function of β;

Steps at $(Y_i - Y_j)/(x_i - x_j)$, $1 \le i<j \le n$.

D(Y-βx)=…..   D'=……

Note that ranks are shift-invariant, so that intercept α or μ does not matter in D'.

$$D(Y - \beta C) = \sum a_N(k)(Y_{i(k)} - \beta c^{i(k)}) \,,$$

$$D'(Y - \beta C) = \frac{dD}{d\beta} = -\sum a_N(k) c^{i(k)} \,.$$

For fixed $Y_1, Y_2, \cdots, Y_N$ and for any β, the residuals are

(2.2) $$z_i = Y_i - \beta c^i \qquad i = 1, 2, \cdots, N\,.$$

The ordered residuals are

$$z_{(k)} = Y_{i(k)} - \beta c^{i(k)} \qquad k = 1, 2, \cdots, N\,,$$

where $i(k)$ is the index of the observation giving rise to the kth ordered residual. (If two residuals are equal there is an ambiguity in $i(k)$ but not in $z_{(k)}$.) The

Non iid, weakly dependent cases.

Proceed in the same way as in iid case.

Numerical calculations for the estimates of β is the same as in iid case.

But, the probability distribution of the estimator or the estimation error are different in the two cases.

Approaches for proving asymptotic normality.

: (1). Approach in Jureckova and Jaeckel. This does not look using a convergence of weighted empirical distribution function (in iid case).

: (2). Can we use "a convergence of weighted empirical distribution function " framework for this problem?

$Y_i - \beta x_i$, i=1,2,…,n. Yes, we can. See Wellner & Shorack(1986) for iid cases.

References.

：Miura.三浦(1985). "Hodges-Lehmann 型推定とLehmann 対立仮説
Special Lecture. Annual Meeting of Japan Society of Mathematics.
：Tsukahara, H. & Miura, R.(1993).
"One sample estimation for generalized Lehmann's alternative models."
Statistica Sinica. Vol.3. 83-101.
===
：Tsukahara, H. (2011). " Estimation of Distortion Risk Measure."
Working Paper. Department of Economics Seijo University.
: Shao,Q.M.&Yu,H. (1996).
Weak convergence for weighted empirical processes of dependent sequences,
Ann. Probab. 24 (4) (1996) 2052–2078.
: Louhichi,S.(2000). Weak convergence for empirical processes of associated sequences.
Ann. Inst. Henri Poincaré, Probabilités et Statistiques 36, 5 (2000) 547–567
===
：Azzalini, A.(1985) .
"A Class of Distributions which includes the Normal Ones."
Scandinavian Journal of Statistics. Vol.12. 171-178.
：Azzalini, A.(2005).
"The skew normal distribution and related multivariate families."
Scandinavian Journal of Statistics. Vol.32. 159-188.
===
：Lehmann, E.L.(1953) . "The power of rank tests." Annals of Mathematical Statistics. Vol.24. 23-43.
: Hallin,M. Jureckova, J.and their colleagues: "Serial and Nonserial sign-and-rank statistics: asymptotic representation and asymptotic normality."(2006) Ann.Statist。
Jaeckel, L.A. (1972). Estimating regression coefficients by minimizing the dispersion of the residuals." Annals of Mathematical Statistics. Vol.43. 1449-1458.
Jureckova,J.(1971). "Nonparametric estimate of regression coefficients."Annals of Mathematical Statistics. Vol.42. 1328-1338.

Thank you for your attention.