EINDHOVEN UNIVERSITY OF TECHNOLOGY

Maximal type test statistics based on conditional processes

Citation for published version (APA):
Beirlant, J., & Einmahl, J. H. J. (1993). *Maximal type test statistics based on conditional processes*. (Memorandum COSOR; Vol. 9317). Technische Universiteit Eindhoven.

Document status and date:
Published: 01/01/1993

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Download date: 08. Feb. 2024

EINDHOVEN UNIVERSITY OF TECHNOLOGY
Department of Mathematics and Computing Science

Memorandum COSOR 93-17

Maximal type test statistics based on conditional processes

J. Beirlant
J.H.J. Einmahl

Eindhoven, June 1993
The Netherlands

Eindhoven University of Technology
Department of Mathematics and Computing Science
Probability theory, statistics, operations research and systems theory
P.O. Box 513
5600 MB Eindhoven - The Netherlands

Secretariat: Dommel building 0.03
Telephone: 040-47 3130

ISSN 0926 4493

MAXIMAL TYPE TEST STATISTICS BASED ON CONDITIONAL PROCESSES

Jan Beirlant [1]
Catholic University Leuven

and

John H.J. Einmahl
Eindhoven University of Technology

Abstract: A general methodology is presented for non-parametric testing of independence, location and dispersion in multiple regression. The proposed testing procedures are based on the concepts of conditional distribution function, conditional quantile, and conditional shortest t-fraction. Techniques involved come from empirical process and extreme-value theory. The asymptotic distributions are standard Gumbel.

AMS 1991 Subject classifications. 62G07, 62G10, 62G20.

Key words and phrases. Non-parametric regression, empirical processes, extreme-value theory.

[1] Research performed while the author was research fellow at the Eindhoven University of Technology.

I. INTRODUCTION AND MAIN RESULTS

Let $(\mathbf{X},Y)$, $(\mathbf{X}_1,Y_1)$, $\cdots$, $(\mathbf{X}_n,Y_n)$ be i.i.d. random vectors from a distribution $\tilde{\mu}$ on $\mathbb{R}^{d+1}$, $\mathbf{X}_i \in \mathbb{R}^d$, $Y_i \in \mathbb{R}$ $(i = 1, \cdots, n)$. The marginal distribution of the $\mathbf{X}$'s is denoted by μ; let S be the support of μ.
In this paper we are concerned with the conditional distribution of Y given $\mathbf{X} = \mathbf{x}$, determined by (a version of) the conditional distribution function (*df*) $F_{\mathbf{x}}$. The corresponding conditional quantiles

$$Q_{\mathbf{x}}(p) = \inf \{y : F_{\mathbf{x}}(y) \geq p\},\ p \in (0,1)\ ,$$

can be used to describe the location of Y given $\mathbf{X} = \mathbf{x}$, as employed in median regression. Dispersion characteristics will be measured by means of lengths of shortest t-fractions (shortt); see e.g. Rousseeuw and Leroy (1988), Grübel (1988), and Einmahl and Mason (1992). For any *df* G and any interval $[c,d] \subset \mathbb{R}$ we use the notation $G([c,d])$ for $G(d) - G(c-)$.
The conditional length of a shortt is now defined by

$$U_{\mathbf{x}}(t) = \inf \{b - a : F_{\mathbf{x}}([a,b]) \geq t\},\ t \in (0,1)\ .$$

It is our aim to provide new tests for independence, constant location, and homoscedasticity through $F_{\mathbf{x}}$, $Q_{\mathbf{x}}(p)$ and $U_{\mathbf{x}}(t)$ respectively. More precisely, the following hypotheses will be considered for $0 < p,t < 1$ fixed:

$$H_0^{(1)} : F_{\mathbf{x}} \text{ is independent of } \mathbf{x} \in S\ (\mu \text{ a.e.})\ ;$$

$$H_0^{(2)} : Q_{\mathbf{x}}(p) \text{ is independent of } \mathbf{x} \in S\ (\mu \text{ a.e.})\ ;$$

$$H_0^{(3)} : U_{\mathbf{x}}(t) \text{ is independent of } \mathbf{x} \in S\ (\mu \text{ a.e.})\ .$$

Our statistical test procedures will be based on an appropriately chosen partition $\{A_{j,n} : j = 1,\cdots,m_n\}$ of S, with for convenience,

$$\mu_j := \mu(A_{j,n}) \geq \mu(A_{j+1,n}) =: \mu_{j+1},\ \text{for all } 1 \leq j \leq m_n - 1\ .$$

Empirical estimates of

$$F_j(y) := P(Y \leq y \mid \mathbf{X} \in A_{j,n})\ ,$$

$$Q_j(p) := \inf \{y : F_j(y) \geq p\}\ ,$$

and

$$U_j(t) := \inf \{b - a : F_j([a,b]) \geq t\}$$

are given by

$$F_{j,n}(y) := \frac{\sum_{i=1}^{n} I_{A_{j,n}\times(-\infty,y]}(\mathbf{X}_i, Y_i)}{\sum_{i=1}^{n} I_{A_{j,n}}(\mathbf{X}_i)} ,$$

$$Q_{j,n}(p) := \inf \{y : F_{j,n}(y) \geq p\} ,$$

and

$$U_{j,n}(t) := \inf \{b - a : F_{j,n}([a,b]) \geq t\} .$$

Throughout we assume F_j ($j = 1, \cdots, m_n$) to be continuous on $\mathbb{R}$. Let μ_n denote the empirical measure based on $\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_n$, and set

$$\mu_{j,n} = \mu_n(A_{j,n}),\ 1 \leq j \leq m_n .$$

Note that the common values of $F_{\mathbf{x}}, Q_{\mathbf{x}}(p)$ under $H_0^{(1)}, H_0^{(2)}$ respectively are equal to $F, Q(p)$, the marginal *df* and p-th quantile of the Y-distribution. Hence they are appropriately estimated by F_n and $Q_n(p)$, with

$$F_n(y) = n^{-1} \sum_{i=1}^{n} I_{(-\infty,y]}(Y_i) \ ,\ y \in \mathbb{R} ,$$

$$Q_n(p) = \inf \{y : F_n(y) \geq p\} .$$

Concerning the hypothesis $H_0^{(3)}$, observe that the common value of $U_{\mathbf{x}}(t)$, denoted by $U_{\cdot}(t)$, is not necessarily equal to the length of the marginal shortt of the Y-distribution. We will estimate $U_{\cdot}(t)$ by

$$U_{\cdot n}(t) = \sum_{j=1}^{m_n} \mu_{j,n} U_{j,n}(t) .$$

Now we are ready to state our main results.

Let

$$\Lambda(x) = \exp(-e^{-x}), \qquad x \in \mathbb{R} ,$$

be the standard Gumbel *df*, Γ a rv with *df* Λ, and write

$$I_n = \sup_{y \in \mathbb{R}} \max_{1 \leq j \leq m_n} \sqrt{n\mu_{j,n}} |F_{j,n}(y) - F_n(y)| .$$

THEOREM 1. If $n\mu_{m_n}/((\log n)^2 \log m_n) \to \infty$ and $\mu_1 \log m_n \to 0$ as $n \to \infty$, then we have under $H_0^{(1)}$ that

$$\sqrt{8 \log m_n}(I_n - \sqrt{\frac{1}{2}\log(2m_n)}) \xrightarrow{d} \Gamma .$$

Let c_α be such that $1 - \Lambda(c_\alpha) = \alpha$, $\alpha \in (0,1)$. Our asymptotic test for independence can now be specified.

COROLLARY 1. The test which rejects $H_0^{(1)}$ when

$$I_n \geq \sqrt{\frac{1}{2}\log(2m_n)} + c_\alpha/\sqrt{8\log m_n}$$

has asymptotic significance level α if the assumptions of Theorem 1 are satisfied.

The following corollary can be applied when the **X**-distribution is known and continuous.

COROLLARY 2. If $m_n \to \infty$, $\mu_1 = \mu_{m_n}$, and $n\mu_1/(\log n)^3 \to \infty$, then

$$\sqrt{8 \log m_n}(I_n - \sqrt{\frac{1}{2}\log(2m_n)}) \xrightarrow{d} \Gamma .$$

In the statement of our next result we make use of the following conditions:

(C.1) for some constant $c_1 > 0$,

$$\limsup_{n\to\infty} \max_{1\leq j\leq m_n} \sup_{y\in R} f_j(y) < c_1$$

where f_j denotes the derivative of F_j;

(C.2) the derivative f of F exists at $Q(p)$ and satisfies $f(Q(p)) > 0$.

Furthermore, let

$$c_{\alpha,n} = \sqrt{2\log m_n} + (c_\alpha - \frac{1}{2}(\log\log m_n + \log\pi))/\sqrt{2\log m_n} .$$

THEOREM 2. Let $p \in (0,1)$ be fixed. The test which rejects $H_0^{(2)}$ when for some $j \in \{1, 2, \cdots, m_n\}$

$$Q_n(p) \notin \left[Q_{j,n}(p - c_{\alpha,n}\sqrt{\frac{p(1-p)}{n\mu_{j,n}}}),\ Q_{j,n}(p + c_{\alpha,n}\sqrt{\frac{p(1-p)}{n\mu_{j,n}}})\right)$$

has asymptotic significance level α if (C1) and (C2) are satisfied and if $n\mu_{m_n}/((\log n)^2 \log m_n) \to \infty$ and $\mu_1 \log m_n \to 0$.

In order to establish our last result some additional regularity conditions are required. The first one reads as follows:

(C.3) for large n, every F_j $(1 \le j \le m_n)$ has a density f_j which is continuous on $I\!R$ and has support $(\beta_j, \gamma_j), -\infty \le \beta_j < \gamma_j \le \infty$, is strictly increasing on $(\beta_j, y_{0,j}]$ and strictly decreasing on $[y_{0,j}, \gamma_j)$ for some $y_{0,j} \in (\beta_j, \gamma_j)$.
Moreover, every $f_{\mathbf{x}}, \mathbf{x} \in S$, satisfies this unimodality assumption.

Let $t \in (0,1)$ be fixed. Under (C.3) we have for large n that there exists a unique interval $[a_{j,t}, b_{j,t}]$ (the shortt) such that $F_j([a_{j,t}, b_{j,t}]) = t,\ f_j(a_{j,t}) = f_j(b_{j,t})$, and $f_j(y) > f_j(a_{j,t})$ for every $y \in (a_{j,t}, b_{j,t})$ $(1 \le j \le m_n)$.

We also need that

(C.4) there exist constants $c_2, \delta_2 > 0$ such that the derivatives f_j' of f_j satisfy

$$\liminf_{n\to\infty} \min_{1\le j\le m_n} \inf_{y\in[a_{j,t},b_{j,t}]\setminus[a_{j,t}+\delta_2,\ b_{j,t}-\delta_2]} |f_j'(y)| > c_2 \ .$$

Introducing the derivative u_j of U_j $(1 \le j \le m_n)$ we assume

(C.5) there exist constants $c_3, c_4 > 0$ such that for every $s \in (0,1)$

$$\limsup_{n\to\infty} \max_{1\le j\le m_n} |u_j(s) - u_j(t)| \le c_3|s-t| \ ,$$

$$\limsup_{n\to\infty} \max_{1\le j\le m_n} u_j(t) < c_4 \ .$$

Finally we will assume

(C.6) $$\sqrt{n\mu_1 \log m_n} \max_{1\le j\le m_n} (U_j(t) - \sup_{\mathbf{x}\in A_{j,n}} U_{\mathbf{x}}(t))^+ \to 0 \ .$$

THEOREM 3. Let $t \in (0,1)$ be fixed. The test which rejects $H_0^{(3)}$ when for some $j \in \{1, 2, \cdots, m_n\}$

$$U_{\cdot n}(t) \notin \left[U_{j,n}(t - c_{\alpha,n}\sqrt{\frac{t(1-t)}{n\mu_{j,n}}}),\ U_{j,n}(t + c_{\alpha,n}\sqrt{\frac{t(1-t)}{n\mu_{j,n}}})\right)$$

has asymptotic significance level α if (C.1), (C.3) - (C.6) are satisfied and if $\mu_1 \log m_n \to 0$, $\mu_1^4(\sum_{j=1}^{m_n} \mu_j^{\frac{1}{2}})^8 (\log n)^4 (\log m_n)^5/(n\mu_{m_n}) \to 0$.

For any $\mathbf{x} \in S$, let $m_t(\mathbf{x})$ be defined as the midpoint of the interval pertaining to $U_{\mathbf{x}}(t)$. This robust regression curve is strongly related to the least median of squares regression estimator introduced in Rousseeuw (1984) (see also Rousseeuw and Leroy (1988)). The following smoothness conditions on m_t and $F_{\mathbf{x}}$ ($\mathbf{x} \in S$) can be used instead of assumption (C.6), as shown by the following corollaries:

(C.7) for some constant $c_5 > 0$,

$$|m_t(\mathbf{x}_1) - m_t(\mathbf{x}_2)| \leq c_5 ||\mathbf{x}_1 - \mathbf{x}_2||$$

for any $\mathbf{x}_1, \mathbf{x}_2 \in S$;

(C.8) the second order derivatives $f'_{\mathbf{x}}$ of $F_{\mathbf{x}}$ exist, and for some $c_6 > 0$,

$$\sup_{\mathbf{x}\in S} \sup_{y \in \mathbb{R}} |f'_{\mathbf{x}}(y)| < c_6 .$$

Let $\text{diam}(A) := \sup\{||\mathbf{x}_1 - \mathbf{x}_2|| : \mathbf{x}_1, \mathbf{x}_2 \in A\}$, where $||\mathbf{x}_1 - \mathbf{x}_2||$ denotes the Euclidian distance between $\mathbf{x}_1$ and $\mathbf{x}_2$.

COROLLARY 3. The test which rejects $H_0^{(3)}$ when for some $j \in \{1, 2, \cdots, m_n\}$

$$U_{\cdot n}(t) \notin \left[U_{j,n}\left(t - c_{\alpha,n}\sqrt{\frac{t(1-t)}{n\mu_{j,n}}}\right),\ U_{j,n}\left(t + c_{\alpha,n}\sqrt{\frac{t(1-t)}{n\mu_{j,n}}}\right)\right)$$

has asymptotic significance level α if (C.1), (C.3) - (C.5), (C.7) and (C.8) are satisfied, and if $n\mu_1 \log m_n (\max_{1\leq j\leq m_n} \text{diam}\,(A_{j,n}))^4 \to 0$, $\mu_1 \log m_n \to 0$, and $\mu_1^4 (\sum_{j=1}^{m_n} \mu_j^{\frac{1}{2}})^8 (\log n)^4 (\log m_n)^5 / (n\mu_{m_n}) \to 0$.

COROLLARY 4. If $m_n \to \infty$, $\mu_1 = \mu_{m_n}$, $n\mu_1/(\log n)^9 \to \infty$, and $n\mu_1 \log m_n (\max_{1\leq j\leq m_n} \text{diam}\,(A_{j,n}))^4 \to 0$, then it follows under (C.1), (C.3) - (C.5), (C.7) and (C.8) that the test which rejects $H_0^{(3)}$ when for some $j \in \{1, 2, \cdots, m_n\}$

$$U_{\cdot n}(t) \notin \left[U_{j,n}\left(t - c_{\alpha,n}\sqrt{\frac{t(1-t)}{n\mu_{j,n}}}\right),\ U_{j,n}\left(t + c_{\alpha,n}\sqrt{\frac{t(1-t)}{n\mu_{j,n}}}\right)\right)$$

has asymptotic significance level α.

REMARKS

1. The choice of $Q_{\mathbf{x}}(p)$, resp. $U_{\mathbf{x}}(t)$, rather than $m_p(\mathbf{x})$, resp. the interquartile range $Q_{\mathbf{x}}(\frac{1+t}{2}) - Q_{\mathbf{x}}(\frac{1-t}{2})$, to produce tests for $H_0^{(2)}$, resp. $H_0^{(3)}$, was motivated in part by

considerations of statistical relevance. Indeed, $m_p(\mathbf{x})$ ($x \in A_{j,n}$) can only be estimated at a rate of $(n\mu_j)^{-\frac{1}{3}}$ (see e.g. Kim and Pollard (1990)), whereas interquartile ranges have a lower breakdown point than the corresponding shortt measures when $t > \frac{1}{3}$ (see Rousseeuw and Leroy (1988) for the case $t = \frac{1}{2}$).
The techniques we use to derive our results however, can also be applied to other testing procedures, e.g. those based on $m_p(\mathbf{x})$ and interquartile ranges.

2. In the cases considered in Theorems 2 and 3, similar results on sup-norm statistics where t, p vary over non-degenerate intervals can be obtained with the technique of proof introduced in the next section.

3. Our statistic I_n discussed in Theorem 1 is somewhat similar to the V- quantities in Kiefer (1959) to test equality of distributions in a one-way layout of several populations. (See also the references in that paper.) The situation considered here provides a generalization of Kiefer's result to the case where the number of groups increases with the sample size.

4. In a non-regression setting an analogue of our type of test statistics is the goodness-of-fit test statistic in Dijkstra, Rietjens and Steutel (1984). In case S is compact, these authors propose to reject uniformity on S when $P_n = \max_{1 \le j \le m_n} \mu_{j,n}$ becomes too large, where the partition is taken to be such that under the null hypothesis the μ_j are all equal. Their simulation study shows that the power of this test is at least comparable to the power of the classical χ^2-test for uniformity against peaked alternatives.
A 'continuous' version of this 'peak-test' is given by the scan statistic (see e.g. Naus (1966, 1982) and Cressie (1980, 1987)) which uses a maximal type statistic obtained from continuous scanning of S with a fixed window. In case $d = 1$, Deheuvels and Révész (1987) derived asymptotics for the scan statistic using a similar condition as in Corollary 2; i.e. $(na_n)/(\log n)^3 \to \infty$, where a_n is the window length.
When $\mu_1 = \mu_{m_n}$ one can also derive the following result for P_n:
if $(n\mu_1)/(\log n)^3 \to \infty$ and $\mu_1 \to 0$, we have under the hypothesis of uniformity that
$$\sqrt{2\log m_n}\left\{\sqrt{\frac{n}{\mu_1}}(P_n - \mu_1) - \sqrt{2\log m_n} + \frac{1}{2}(\log\log m_n + \log 4\pi)/(2\log m_n)^{\frac{1}{2}}\right\} \xrightarrow{d} \Gamma \ .$$

5. The condition $n\mu_1(\log m_n)(\max_{1 \le j \le m_n} \text{diam } (A_{j,n}))^4 \to 0$ specifies to $nh_n^{4+d} \log \frac{1}{h_n} \to 0$ in case $\mathbf{X}$ possesses a uniform distribution on $[0,1]^d$, say, and the partition is taken to be cubic with diam $(A_{j,n}) \sim h_n$ ($j = 1, \cdots, m_n$). This rate condition of h_n lies close to the optimal rate of the window size in kernel density estimation when minimizing the mean squared error.

6. If one wants to restrict attention to a subset of the support S of $\mathbf{X}$, all of our results can still be used by translating them in terms of conditional distributions given $\mathbf{X}$ belongs to that subset.

II. PROOFS

The proofs of our main results rely on the following proposition which states that jointly over all elements $A_{j,n}$ of the partition of S, we can approximate the different empirical processes

$$\alpha_{j,n} = \sqrt{n\mu_{j,n}}(F_{j,n} - F_j) \quad , \ j = 1, \cdots, m_n,$$

by independent Gaussian processes, and this, per j, at a rate which is comparable to the one attained by the Komlós-Major-Tusnády (1975) approximation of the (one-dimensional) uniform empirical process.

Denoting the joint distribution of $(\mathbf{X}, Y)$ by $\tilde{\mu}$, and the empirical measure based on $(\mathbf{X}_1, Y_1), \cdots, (\mathbf{X}_n, Y_n)$ by $\tilde{\mu}_n$, we will use the following quantities:

$$\tilde{\mu}_j(y) = P(\mathbf{X} \in A_{j,n} \text{ and } Y \leq y) = \tilde{\mu}(A_{j,n} \times (-\infty, y]) ,$$

$$\tilde{\mu}_{j,n}(y) = \tilde{\mu}_n(A_{j,n} \times (-\infty, y]) ,$$

so that

$$F_{j,n}(y) = \frac{\tilde{\mu}_{j,n}(y)}{\mu_{j,n}} .$$

PROPOSITION. If $m_n \to \infty$ and $(n\mu_{m_n})/(\log n)^2 \to \infty$, then there exists a triangular scheme of rowwise independent Brownian bridges $\{B_{j,n}(t),\ 0 \leq t \leq 1\}$ $(1 \leq j \leq m_n,\ n \geq 1)$ such that

$$\sup_{y \in \mathbb{R}} \max_{1 \leq j \leq m_n} |\alpha_{j,n}(y) - B_{j,n}(F_j(y))| = O_P(\frac{\log n}{\sqrt{n\mu_{m_n}}}).$$

Proof. We consider the transformation from $S \times \mathbb{R}$ to $[0,1]$

$$(\mathbf{x}, y) \to T(\mathbf{x}, y) = \sum_{j=1}^{m_n} 1_{A_{j,n}}(\mathbf{x})[\sum_{k=1}^{j-1} \mu_k + \mu_j F_j(y)] ,$$

and the transformed rv's

$$Z_i = T(\mathbf{X}_i, Y_i) \qquad , i = 1, 2, \cdots, n .$$

On easily checks that $Z_1, Z_2, \cdots, Z_n$ are independent uniformly (0,1) distributed rv's. Let $\{e_n(t), 0 \leq t \leq 1\}$ denote the empirical process based on $Z_1, Z_2, \cdots, Z_n$. The approximation theorem of Komlós, Major and Tusnády (1975) entails then the existence of a sequence of Brownian bridges $\{\tilde{B}_n(t), 0 \leq t \leq 1\}$ such that as $n \to \infty$

$$\sup_{0 \leq t \leq 1} |e_n(t) - \tilde{B}_n(t)| = O_P(\frac{\log n}{\sqrt{n}}) .$$

It follows that (in the obvious notation)

$$\max_{1 \leq j \leq m_n} |\sqrt{n}(\mu_{j,n} - \mu_j) - \tilde{B}_n(T(A_{j,n} \times (-\infty, \infty]))| = O_P(\frac{\log n}{\sqrt{n}})$$

and that

$$\max_{1\le j\le m_n}\ \sup_{y\in\mathbb{R}} |\sqrt{n}(\tilde\mu_{j,n}(y)-\tilde\mu_j(y)) - \tilde B_n(T(A_{j,n}\times(-\infty,y]))| = O_P(\frac{\log n}{\sqrt{n}})\ .$$

Now uniformly in $j\in\{1,\cdots,m_n\}$ and $y\in\mathbb{R}$ we have

$$\begin{aligned}\alpha_{j,n}(y) &= \sqrt{n\mu_{j,n}}\left(\tfrac{\tilde\mu_{j,n}(y)}{\mu_{j,n}} - \tfrac{\tilde\mu_j(y)}{\mu_j}\right)\\ &= \sqrt{n}(\tilde\mu_{j,n}(y)-\tilde\mu_j(y))/\sqrt{\mu_{j,n}} - \sqrt{n}(\mu_{j,n}-\mu_j)(\tilde\mu_j(y)/(\mu_j\sqrt{\mu}_{j,n}))\\ &= \{\{\tilde B_n(T(A_{j,n}\times(-\infty,y])) + O_P(\tfrac{\log n}{\sqrt{n}})\}\mu_j^{-\frac12} - \{\tilde B_n(T(A_{j,n}\times(-\infty,\infty]))\\ &\quad +O_P(\tfrac{\log n}{\sqrt{n}})\}F_j(y)\mu_j^{-\frac12}\}\tau_{j,n}^{-\frac12}\end{aligned}$$

where

$$\tau_{j,n} = \mu_{j,n}/\mu_j = 1 + n^{-\frac12}\mu_j^{-1}\tilde B_n(T(A_{j,n}\times(-\infty,\infty])) + \mu_j^{-1}O_P(\tfrac{\log n}{n}). \tag{2.1}$$

We can define a sequence of Wiener processes $\{W_n(t), 0\le t\le 1\}$ such that $\tilde B_n = W_n - IW_n(1)$, where I denotes the identity function. Hence, as (with λ denoting Lebesgue measure) $\lambda(T(A_{j,n}\times(-\infty,y])) = \mu_j F_j(y)$, we find that

$$\begin{aligned}\alpha_{j,n}(y) = &\ \{\{W_n(T(A_{j,n}\times(-\infty,y])) - F_j(y)W_n(T(A_{j,n}\times(-\infty,\infty]))\}\mu_j^{-\frac12}\\ &+\mu_j^{-\frac12}O_P(\tfrac{\log n}{\sqrt{n}})\}\tau_{j,n}^{-\frac12}\qquad (n\to\infty)\ .\end{aligned}$$

We now set

$$B_{j,n}(F_j(y)) := \mu_j^{-\frac12}\{W_n(T(A_{j,n}\times(-\infty,y])) - F_j(y)W_n(T(A_{j,n}\times(-\infty,\infty]))\}\ .$$

One easily checks that the $B_{j,n}$ are indeed independent in $j\in\{1,2,\cdots,m_n\}$ and distributed as Brownian bridges.

Now as $n\to\infty$

$$|\alpha_{j,n}(y) - B_{j,n}(F_j(y))| \le |B_{j,n}(F_j(y))|\,(\tau_{j,n}^{-\frac12}-1) + \mu_j^{-\frac12}\tau_{j,n}^{-\frac12}O_P(\frac{\log n}{\sqrt{n}})\ .$$

For a function φ on [0,1], write $\|\varphi\| = \sup_{0\le t\le 1}|\varphi(t)|$. First remark that as the F_j are assumed to be continuous

$$\max_{1\le j\le m_n}\ \sup_{y\in\mathbb{R}} |B_{j,n}(F_j(y))| = \max_{1\le j\le m_n}\ \|B_{j,n}\| = O_P(\sqrt{\log m_n})\ ,$$

as, because of the independence of the rv's $\|B_{j,n}\|$ $(1 \le j \le m_n)$, we have for any $M > 0$, that

$$P(\max_{1\le j\le m_n} \|B_{j,n}\| > M\sqrt{\log m_n}) \le 2m_n\, e^{-2M^2 \log m_n} = 2m_n^{1-2M^2} ,$$

which tends to zero as $m_n \to \infty$ when $M > 2^{-\frac{1}{2}}$. (Here we also used the fact that for a Brownian bridge B we have $P(\|B\| > u) \le 2e^{-2u^2}$.)

Furthermore,

$$n^{-\frac{1}{2}} \max_{1\le j\le m_n} \mu_j^{-1} |\tilde{B}_n((\sum_{k=1}^{j-1} \mu_k, \sum_{k=1}^{j} \mu_k])|$$

$$\le n^{-\frac{1}{2}} \mu_{m_n}^{-\frac{1}{2}} \max_{1\le j\le m_n} (\mu_j^{-\frac{1}{2}}\, |W_n\, ((\sum_{k=1}^{j-1} \mu_k,\ \sum_{k=1}^{j} \mu_k])| + \mu_{m_n}^{\frac{1}{2}}\, |W_n(1)|)$$

$$= (n\mu_{m_n})^{-\frac{1}{2}} O_P(\sqrt{\log m_n}) ,$$

since $W_n((\sum_{k=1}^{j-1} \mu_k,\ \sum_{k=1}^{j} \mu_k])/\sqrt{\mu_j}$ $(1 \le j \le m_n)$ are m_n independent standard normal rv's whose maximum is well known to be of order $O_P(\sqrt{\log m_n})$ as $n \to \infty$.

Hence,

$$\max_{1\le j\le m_n} |\tau_{j,n} - 1| = O_P\left(\sqrt{\frac{\log n}{n\mu_{m_n}}} + \frac{\log n}{n\mu_{m_n}}\right), \tag{2.2}$$

and

$$\max_{1\le j\le m_n} \sup_{y\in \boldsymbol{R}} |B_{j,n}(F_j(y))|\, |\tau_{j,n}^{-\frac{1}{2}} - 1| = O_P(\frac{\log n}{\sqrt{n\mu_{m_n}}} + \frac{(\log n)^{3/2}}{n\mu_{m_n}}) \quad (n \to \infty) .$$

Finally, with

$$(\max_{1\le j\le m_n} (\mu_j \tau_{j,n})^{-\frac{1}{2}}) O_P(\frac{\log n}{\sqrt{n}}) = O_P(\frac{\log n}{\sqrt{n\mu_{m_n}}}) ,$$

the result follows. □

Proof of Theorem 1. First remark that by the well known fact that

$$\sqrt{n} \sup_{y\in \boldsymbol{R}} |F_n(y) - F(y)| = O_P(1) \quad (n \to \infty) ,$$

we have,

$$\sqrt{\log m_n}\sup_{y\in\mathbf{R}}\max_{1\le j\le m_n}\sqrt{n\mu_{j,n}}\,|F_n(y)-F(y)| \;=(\mu_1\log m_n)^{\frac{1}{2}}O_P(1)\qquad(n\to\infty)$$

since, as in the proof of the Proposition we find that uniformly in $j\in\{1,\cdots,m_n\}$

$$\mu_{j,n}^{\frac{1}{2}}=\mu_j^{\frac{1}{2}}(1+O_P(\sqrt{\frac{\log m_n}{n\mu_{m_n}}}))\ .$$

Hence since $\mu_1\log m_n\to 0$, it suffices to show that, under $H_0^{(1)}$,

$$\sqrt{8\log m_n}\left(\sup_{y\in\mathbf{R}}\max_{1\le j\le m_n}\sqrt{n\mu_{j,n}}\,|F_{j,n}(y)-F(y)|\;-\sqrt{\frac{1}{2}\log(2m_n)}\right)\xrightarrow{d}\Gamma\qquad(n\to\infty)\ .$$

Under $H_0^{(1)}$ it now follows from the Proposition that

$$\sqrt{8\log m_n}\;|\sup_{y\in\mathbf{R}}\max_{1\le j\le m_n}\sqrt{n\mu_{j,n}}\,|F_{j,n}(y)-F(y)|\;-\sup_{y\in\mathbf{R}}\max_{1\le j\le m_n}\;|B_{j,n}(F_j(y))|\;|$$
$$=O_P(\tfrac{\log n\sqrt{\log m_n}}{\sqrt{n\mu_{m_n}}})=o_p(1)$$

if $n\to\infty$ and $n\mu_{m_n}/((\log n)^2\log m_n)\to\infty$.

Finally, remark that by the independence of the $\|B_{j,n}\|$ $(1\le j\le m_n)$ we can apply standard extreme value theory to show that

$$\text{(2.3)}\qquad \sqrt{8\log m_n}\left\{\max_{1\le j\le m_n}\|B_{j,n}\|-\sqrt{\frac{1}{2}\log(2m_n)}\right\}\xrightarrow{d}\Gamma$$

since $P(\|B_{j,n}\|>u)\sim 2e^{-2u^2}$ (see Proposition 1.19 in Resnick (1987)). □

Proof of Corollary 2. If $\mu_1=\mu_2=\cdots=\mu_{m_n}$, then $m_n=\mu_1^{-1}$.

The condition $\mu_1\log m_n\to 0$ is then automatically satisfied when $m_n\to\infty$. □

Proof of Theorem 2. Observe that, under $H_0^{(2)}$,

$$P(Q_n(p)\notin[Q_{j,n}(p-c_{\alpha,n}\sqrt{\frac{p(1-p)}{n\mu_{j,n}}}),\;Q_{j,n}(p+c_{\alpha,n}\sqrt{\frac{p(1-p)}{n\mu_{j,n}}})),\text{for some } j\in\{1,\cdots,m_n\})$$

$$\to\alpha\qquad\text{as } n\to\infty$$

if

$$\text{(2.4)}\qquad \sqrt{2\log m_n}\{\max_{1\le j\le m_n}\left(\sqrt{n\mu_{j,n}}\,|F_{j,n}(Q_n(p))-p|\,/\sqrt{p(1-p)}\right)-\sqrt{2\log m_n}$$
$$+\tfrac{1}{2}(\log\log m_n+\log\pi)(2\log m_n)^{-\frac{1}{2}}\}\xrightarrow{d}\Gamma\ .$$

Indeed, for any *df* G on the real line and any $p\in(0,1)$ we have

$G(x)\ge p$ if and only if $G^{-1}(p)\le x$

and hence

$G(x)<p$ if and only if $G^{-1}(p)>x$.

We first show that under (C.1), (C.2), $n\mu_{m_n}/((\log n)^2\log m_n)\to\infty$ and $\mu_1\log m_n\to 0$

$$\text{(2.5)}\qquad \sqrt{2\log m_n}\left\{\max_{1\le j\le m_n}\sqrt{n\mu_{j,n}}\,|F_{j,n}(Q_n(p))-p|\ -\max_{1\le j\le m_n}|B_{j,n}(F_j(Q_n(p)))|\right\}\xrightarrow{P}0.$$

where $\{B_{j,n}\}$ $(1\le j\le m_n, n\ge 1)$ is the sequence of Brownian bridges described in the Proposition. Now (2.5) follows from the Proposition if we can show that under our assumptions

$$\text{(2.6)}\qquad \sqrt{\log m_n}\max_{1\le j\le m_n}\sqrt{n\mu_{j,n}}\,|F_j(Q_n(p))-p|\xrightarrow{P}0.$$

The well-known central limit theorem for quantiles yields that under $H_0^{(2)}$ and (C.2)

$$Q_n(p)=Q(p)+O_P(n^{-\frac{1}{2}})=Q_j(p)+O_p(n^{-\frac{1}{2}})$$

when $n\to\infty$. Hence by the mean value theorem we have under $H_0^{(2)}$ that

$$\begin{aligned}F_j(Q_n(p)) &= F_j(Q_j(p)+O_P(n^{-\frac{1}{2}}))\\ &= p+O_P(n^{-\frac{1}{2}})f_j(\tilde{Q}_{j,n}(p))\end{aligned}$$

with $\tilde{Q}_{j,n}(p)\in(Q_n(p)\wedge Q_j(p),\ Q_n(p)\vee Q_j(p))$ $(1\le j\le m_n)$. Hence by (C.1) and under $H_0^{(2)}$,

$$\text{(2.7)}\qquad \sqrt{n}\max_{1\le j\le m_n}|F_j(Q_n(p))-p|=O_P(1)\ \ (n\to\infty),$$

so that it remains to check that $(\log m_n)(\max_{1\le j\le m_n}\mu_{j,n})\xrightarrow{P}0$ $(n\to\infty)$ for (2.6) (and hence (2.5)) to hold.
However, using $\tau_{j,n}$ in (2.1) again, we get that

$$\log m_n(\max_{1\le j\le m_n}\mu_{j,n})\le\mu_1(\log m_n)(\max_{1\le j\le m_n}\tau_{j,n})\ ,$$

which tends to zero in probability as $n \to \infty$ and $\mu_1 \log m_n \to 0$ because of (2.2).
Next, it follows from (2.7), and the modulus of continuity behaviour of Brownian bridges (see e.g. Lemma 1.1.1 in Csörgő and Révész (1981)) that

$$(2.8) \qquad \sqrt{\log m_n} \max_{1 \le j \le m_n} \big| \, |B_{j,n}(F_j(Q_n(p)))| - |B_{j,n}(p)| \, \big| = O_P(n^{-\frac{1}{4}}((\log n)(\log m_n))^{\frac{1}{2}})$$
$$= o_P(1) \quad (n \to \infty) .$$

As $B_{j,n}(p)$ $(1 \le j \le m_n)$ are independent $\mathcal{N}(0, p(1-p))$ rv's, standard techniques form extreme value theory yield that

$$(2.9) \qquad \sqrt{2 \log m_n} \, \{(p(1-p))^{-\frac{1}{2}} \max_{1 \le j \le m_n} |B_{j,n}(p)| - \sqrt{2 \log m_n} + \frac{1}{2}(\log \log m_n + \log \pi)$$
$$\cdot (2 \log m_n)^{-\frac{1}{2}}\} \xrightarrow{d} \Gamma \quad (m_n \to \infty) .$$

Limit statement (2.4) now follows from (2.5), (2.8) and (2.9). □

Proof of Theorem 3.

We introduce the functions

$$H_j(z) = \sup\{F_j([a,b]) : b - a \le z\} .$$

Note that H_j is the inverse of U_j (for n large enough). The derivative of H_j is denoted by h_j. Remark that condition (C.1) implies

$$(2.10) \qquad \limsup_{n \to \infty} \max_{1 \le j \le m_n} \sup_{z \ge 0} h_j(z) < \infty,$$

as for each $j \in \{1, \cdots, m_n\}$ we find that h_j is non-increasing and $h_j(0) = \max_{y \in \boldsymbol{R}} f_j(y)$.

Analogously we define the inverse function $H_{j,n}$ of $U_{j,n}$ by

$$H_{j,n}(z) = \inf\{t : U_{j,n}(t) \ge z\}$$

and note that

$$H_{j,n}(z) = \sup\{F_{j,n}([a,b]) : b - a \le z\}, \ 1 \le j \le m_n .$$

To prove Theorem 3 it now suffices to show that under $H_0^{(3)}$

$$(2.11) \qquad \sup_{t \in (0,1)} \max_{1 \le j \le m_n} |\sqrt{n \mu_{j,n}} (H_{j,n}(U_{\cdot n}(t)) - t) - \tilde{B}_{j,n}(t)| = O_P((n\mu_{m_n})^{-\frac{1}{8}} (\log m_n)^{\frac{1}{8}} (\log n)^{\frac{1}{2}})$$

for some triangular scheme of rowwise independent Brownian bridges $\{\tilde{B}_{j,n}\}$ $(1 \le j \le m_n, n \ge 1)$; cf. the proof of Theorem 2. We derive (2.11) in three steps by showing that under the given conditions

(2.12) $$\sup_{t\in(0,1)} \max_{1\le j\le m_n} |\sqrt{n\mu_{j,n}}\ (H_{j,n}(U_j(t)) - t) - \tilde{B}_{j,n}(t)| \ = O_P((n\mu_{m_n})^{-\frac{1}{8}}(\log m_n)^{\frac{1}{8}}(\log n)^{\frac{1}{2}})\ ,$$

(2.13) $$\sqrt{\log m_n}\ \max_{1\le j\le m_n} \sqrt{n\mu_{j,n}}\ |H_j(U_{\cdot n}(t)) - t| \ \xrightarrow{P}\ 0,$$

and

(2.14) $$\sqrt{\log m_n}\ \max_{1\le j\le m_n} |\tilde{B}_{j,n}\ (H_j(U_{\cdot n}(t))) - \tilde{B}_{j,n}(t)| \ \xrightarrow{P} 0 \qquad (n\to\infty).$$

First, we prove the existence of a sequence $\{\tilde{B}_{j,n}\}$ of Brownian bridges for which (2.12) holds. Remark that from the Proposition it follows that

(2.15) $$\sup_{[a,b]} \max_{1\le j\le m_n} |\alpha_{j,n}([a,b]) - (B_{j,n}(F_j(b)) - B_{j,n}(F_j(a)))| \ = O_P(\frac{\log n}{\sqrt{n\mu_{m_n}}})\ ,$$

as $B_{j,n}([F_j(a), F_j(b)]) = B_{j,n}(F_{j,n}(b)) - B_{j,n}(F_{j,n}(a))$. To derive (2.12) from (2.15) we apply and refine the method of proof of Proposition 3.1 in Einmahl and Mason (1992). We define

$$\tilde{B}_{j,n}(t) = B_{j,n}(F_j(b_{j,t})) - B_{j,n}(F_j(a_{j,t})),\ 1\le j\le m_n\ .$$

As the intervals $[a_{j,t}, b_{j,t}]$ are nested for different values of t, one easily checks that the $\tilde{B}_{j,n}$ are distributed as Brownian bridges for every $j\in\{1,2,\cdots,m_n\}$ and large n; moreover, $\tilde{B}_{1,n},\cdots,\tilde{B}_{m_n,n}$ are clearly independent.
Notice that for any $j\in\{1,\cdots,m_n\}$ and $0<t<1$

(2.16) $$\tilde{B}_{j,n}(t) - \sqrt{n\mu_{j,n}}(H_{j,n}(U_j(t)) - t) \le (B_{j,n}(F_j(b_{j,t})) - B_{j,n}(F_j(a_{j,t}))) \\ -\alpha_{j,n}([a_{j,t}, b_{j,t}])$$

which, by (2.15), is seen to be $O_P(\frac{\log n}{\sqrt{n\mu_{m_n}}})$, uniformly in $j\in\{1,\cdots,m_n\}$.

Next, we also have for any $j\in\{1,\cdots,m_n\}$ and any sequence $\varepsilon_n\downarrow 0$

(2.17) $$\sqrt{n\mu_{j,n}}(H_{j,n}(U_j(t)) - t) - \tilde{B}_{j,n}(t) \\ \le \left\{\sqrt{n\mu_{j,n}} \sup_{\substack{b-a\le U_j(t)\\ t-\varepsilon_n<F_j([a,b])\le t}} (F_{j,n}([a,b]) - t) - \tilde{B}_{j,n}(t)\right\} \\ \vee\left\{\sqrt{n\mu_{j,n}} \sup_{F_j([a,b])\le t-\varepsilon_n} (F_{j,n}([a,b]) - t) - \tilde{B}_{j,n}(t)\right\}\ .$$

The second term on the right hand side of (2.17) is

$$\leq \sqrt{n\mu_{j,n}} \sup_{F_j([a,b])\leq t} (F_{j,n}([a,b]) - F_j([a,b])) + |\tilde{B}_{j,n}(t)| - \varepsilon_n\sqrt{n\mu_{j,n}}$$
$$\leq 2 \max_{1\leq j\leq m_n} \sup_{[c,d]} |B_{j,n}([c,d])|$$
$$+ \max_{1\leq j\leq m_n} \sup_{[a,b]} |\alpha_{j,n}([a,b]) - (B_{j,n}(F_j(b)) - B_{j,n}(F_j(a)))| - \varepsilon_n \min_{1\leq j\leq m_n} \sqrt{n\mu_{j,n}} \, .$$

From (2.3), (2.15) and (2.2) it now follows that the second term on the right hand side of (2.17) can be asymptotically bounded from above by 0 in probability, by making the appropriate choice

$$\varepsilon_n = M(\log m_n/(n\mu_{m_n}))^{\frac{1}{2}}$$

with M a large enough positive constant.

The first term on the right hand side of (2.17) is

$$(2.18) \qquad \leq \sqrt{n\mu_{j,n}} \sup_{\substack{b-a\leq U_j(t)\\ t-\varepsilon_n<F_j([a,b])\leq t}} ((F_{j,n}(b) - F_{j,n}(a)) - (F_j(b) - F_j(a))) - \tilde{B}_{j,n}(t)$$
$$\leq \sup_{\substack{b-a\leq U_j(t)\\ t-\varepsilon_n<F_j([a,b])\leq t}} |\alpha_{j,n}([a,b]) - (B_{j,n}(F_j(b)) - B_{j,n}(F_j(a)))|$$
$$+ \sup_{\substack{b-a\leq U_j(t)\\ t-\varepsilon_n<F_j([a,b])\leq t}} (B_{j,n}(F_j(b)) - B_{j,n}(F_j(a))) - \tilde{B}_{j,n}(t) \, .$$

The first term on the right hand side of (2.18) is of order $O_P(\frac{\log n}{\sqrt{n\mu_{m_n}}})$, uniformly in $j \in \{1, 2, \cdots, m_n\}$, by (2.15).

Finally observe that for any $j \in \{1, \cdots, m_n\}$

$$(2.19) \qquad \sup_{\substack{b-a\leq U_j(t)\\ t-\varepsilon_n<F_j([a,b])\leq t}} \{B_{j,n}(F_j(b)) - B_{j,n}(F_j(a))\} - \tilde{B}_{j,n}(t) \leq$$
$$\max_{1\leq j\leq m_n} \sup_{\substack{b-a\leq U_j(t)\\ t-\varepsilon_n<F_j([a,b])\leq t}} \{|B_{j,n}(F_j(b)) - B_{j,n}(F_j(b_{j,t}))|$$
$$+ |B_{j,n}(F_j(a)) - B_{j,n}(F_j(a_{j,t}))| \} \, .$$

For any interval $[a, b]$ with $b - a = U_j(t)$ and $t - \varepsilon_n < F_j([a, b]) \leq t$, we find by (C.4) that (uniformly in j) $|a - a_{j,t}|$ and $|b - b_{j,t}|$ become arbitrarily small as $n \to \infty$. We then find, with δ_2 as in (C.4), that eventually as $n \to \infty$ whether $a \in [a_{j,t}, a_{j,t} + \delta_2)$ or $b \in [b_{j,t} - \delta_2, b_{j,t}]$. In case $a < a_{j,t} < y_{0,j} < b < b_{j,t}$, we have that

$$F_j([b, b_{j,t}]) \leq |b_{j,t} - b| \, f_j(y_{0,j}) \leq c_1(b_{j,t} - b) \, .$$

On the other hand, if $\varepsilon_n \geq F_j([a_{j,t}, b_{j,t}]) - F_j([a, b]) \geq 0$, then

$$\varepsilon_n \geq F_j([b, b_{j,t}]) - (b_{j,t} - b)f_j(b_{j,t}) = -((b_{j,t} - b)^2/2)f_j'(\tilde{b}_{j,t})$$

with $\tilde{b}_{j,t} \in (b_{j,t} \wedge b, b_{j,t} \vee b)$, so that (C.4) implies that for n large enough $F_j(b_{j,t}) - F_j(b) \leq C\varepsilon_n^{\frac{1}{2}}$, for some $C > 0$. Also in the other possible cases we can obtain this same bound for $|F_j(b_{j,t}) - F_j(b)| \vee |F_j(a_{j,t}) - F_j(a)|$. Hence the expression on the right hand side of (2.19) can be bounded by

$$(2.20) \qquad \omega(n, \varepsilon_n) := 2 \max_{1 \leq j \leq m_n} \sup_{0 \leq s \leq 1 - C\sqrt{\varepsilon_n}} \sup_{0 \leq t \leq C\sqrt{\varepsilon_n}} |B_{j,n}(s+t) - B_{j,n}(s)| \, .$$

By Lemma 1.1.1 in Csörgő and Révész (1981), the representation of Brownian bridges in terms of Wiener processes, and the independence of the Brownian bridges $B_{j,n}$ $(1 \leq j \leq m_n)$, we obtain that for any $K > 0$ there exist constants $K_1, K_2 > 0$ such that

$$P(\omega(n, \varepsilon_n) > K\gamma_n) \leq K_1 m_n \varepsilon_n^{-\frac{1}{2}} \exp(-K_2 K^2 \gamma_n^2 \varepsilon_n^{-\frac{1}{2}}) \, .$$

Choosing

$$\gamma_n = \varepsilon_n^{\frac{1}{4}} (\log n)^{\frac{1}{2}} = M^{\frac{1}{4}} (\log m_n)^{\frac{1}{8}} (n\mu_{m_n})^{-\frac{1}{8}} (\log n)^{\frac{1}{2}} \, ,$$

one easily checks that

$$P(\omega(n, \varepsilon_n) > K\gamma_n) \to 0 \qquad (n \to \infty)$$

choosing $K > 0$ large enough. This together with (2.16) - (2.20) implies (2.12).

To derive (2.13), note that under $H_0^{(3)}$ we have for any $j \in \{1, \cdots, m_n\}$

$$\begin{aligned} (2.21) \qquad & |H_j(U_{\cdot n}(t)) - t| \\ & \leq |H_j(U_{\cdot n}(t)) - H_j(U_{\cdot}(t))| + |H_j(U_{\cdot}(t)) - H_j(U_j(t))| \\ & \leq |U_{\cdot n}(t)) - U_{\cdot}(t)| \, h_j(\tilde{U}_n(t)) + |U_{\cdot}(t) - U_j(t)| \, h_j(\tilde{U}_j(t)) \end{aligned}$$

where $\tilde{U}_n(t) \in (U_{\cdot n}(t) \wedge U_{\cdot}(t), \, U_{\cdot n}(t) \vee U_{\cdot}(t))$ and $\tilde{U}_j(t) \in (U_j(t) \wedge U_{\cdot}(t), \, U_j(t) \vee U_{\cdot}(t))$, $1 \leq j \leq m_n$.

Now using (2.2) and (2.10), and the fact that under $H_0^{(3)}$ from (C.6) and $U_j(t) \geq U_{\cdot}(t)$ it follows that

$$(2.22) \qquad \sqrt{n\mu_1 \log m_n} \max_{1 \leq j \leq m_n} |U_j(t) - U_{\cdot}(t)| \to 0 \qquad (n \to \infty) \, ,$$

we now find that as $n \to \infty$

$$\begin{aligned} (2.23) \qquad & \sqrt{\log m_n} \max_{1 \leq j \leq m_n} (\sqrt{n\mu_{j,n}} \, |U_{\cdot}(t) - U_j(t)| \,) h_j(\tilde{U}_j(t)) \\ & = O_P(\sqrt{n\mu_1 \log m_n} \max_{1 \leq j \leq m_n} |U_j(t) - U_{\cdot}(t)| \,) = o_P(1). \end{aligned}$$

On the other hand, by (2.2) and (2.10), as $n \to \infty$

$$(2.24)\qquad \sqrt{\log m_n}\max_{1\le j\le m_n}(\sqrt{n\mu_{j,n}}h_j(\tilde U_n(t)))\,|U_{\cdot n}(t)-U_{\cdot}(t)|$$
$$= O_P(1)\sqrt{n\mu_1\log m_n}\,|U_{\cdot n}(t)-U_{\cdot}(t)|\ .$$

Furthermore,

$$(2.25)\qquad |U_{\cdot n}(t)-U_{\cdot}(t)|\ \le \sum_{j=1}^{m_n}\mu_{j,n}|\,U_j(t)-U_{\cdot}(t)|\ +|\sum_{j=1}^{m_n}\mu_{j,n}(U_{j,n}(t)-U_j(t))|.$$

Now

$$(2.26)\qquad \sqrt{n\mu_1\log m_n}\sum_{j=1}^{m_n}\mu_{j,n}\ |U_j(t)-U_{\cdot}(t)|\ \le \sqrt{n\mu_1\log m_n}\max_{1\le j\le m_n}\ |U_j(t)-U_{\cdot}(t)|$$

which tends to zero by (2.22).

The mean value theorem yields that for some $\tilde t_{j,n}\in(H_j(U_{j,n}(t))\wedge t,\ H_j(U_{j,n}(t))\vee t)$

$$(2.27)\qquad \sum_{j=1}^{m_n}\mu_{j,n}\ (U_{j,n}(t)-U_j(t))\ =\ \sum_{j=1}^{m_n}\mu_{j,n}\ (U_j(H_j(U_{j,n}(t)))-U_j(t))$$
$$=\ \sum_{j=1}^{m_n}\mu_{j,n}\ u_j(\tilde t_{j,n})(H_j(U_{j,n}(t))-t)\ .$$

We now show that in this last expression we can replace $\mu_{j,n}u_j(\tilde t_{j,n})$ by $\sqrt{\mu_j}u_j(t)\sqrt{\mu_{j,n}}$. To this end we first remark that using (2.2) and (2.12) we have as $n\to\infty$ that

$$(2.28)\qquad \max_{1\le j\le m_n}\ \sup_{t\in(0,1)}\ |H_j(U_{j,n}(t))-t|\ =\ \max_{1\le j\le m_n}\ \sup_{t\in(0,1)}\ |H_{j,n}(U_j(t))-t|$$
$$=\ (n\mu_{m_n})^{-\frac12}\max_{1\le j\le m_n}\ \|\tilde B_{j,n}\|\ +O_P((n\mu_{m_n})^{-\frac58}(\log m_n)^{\frac18}(\log n)^{\frac12})$$
$$= O_P((n\mu_{m_n})^{-\frac12}\ (\log m_n)^{\frac12}\ +\ (n\mu_{m_n})^{-\frac58}\ (\log m_n)^{\frac18}\ (\log n)^{\frac12})$$
$$= O_P((n\mu_{m_n})^{-\frac12}(\log m_n)^{\frac12})\ .$$

A similar argument yields that

$$(2.29)\qquad \max_{1\le j\le m_n}\sqrt{n\mu_{j,n}}\ \sup_{t\in(0,1)}\ |H_j(U_{j,n}(t))-t|\ = O_P((\log m_n)^{\frac12})\ \ (n\to\infty).$$

Using (C.5) we obtain that

$$|u_j(\tilde t_{j,n})-u_j(t)|\ \le c_3\ |\tilde t_{j,n}-t|\ \le c_3\max_{1\le j\le m_n}\ |H_j(U_{j,n}(t))-t|$$
$$=\ O_P((n\mu_{m_n})^{-\frac12}(\log m_n)^{\frac12})\ \ (n\to\infty).$$

Hence with (2.29) and the rate condition in the statement of the theorem we have that

$$(2.30)\quad \sqrt{n\mu_1 \log m_n}\sum_{j=1}^{m_n}\mu_{j,n}\ |u_j(\tilde{t}_{j,n})-u_j(t)|\ |H_j(U_{j,n}(t))-t|$$
$$= O_P(\mu_1^{\frac{1}{2}}(n\mu_{m_n})^{-\frac{1}{2}}\log m_n)\sum_{j=1}^{m_n}\mu_{j,n}^{\frac{1}{2}}(\sqrt{n\mu_{j,n}}\ |H_j(U_{j,n}(t))-t|\ )$$
$$= O_P((\log m_n)^{\frac{3}{2}}(n\mu_{m_n})^{-\frac{1}{2}}\mu_1^{\frac{1}{2}}(\sum_{j=1}^{m_n}\mu_{j,n}^{\frac{1}{2}}))$$
$$= o_P(1)\quad (n\to\infty)\ .$$

Next, using (C.5), (2.29), and $\max_{1\le j\le m_n}|\mu_{j,n}^{\frac{1}{2}}-\mu_j^{\frac{1}{2}}| = O_P((\log n)^{\frac{1}{2}}n^{-\frac{1}{2}})$ $(n\to\infty)$ we find

$$(2.31)\quad \sqrt{n\mu_1 \log m_n}\ |\sum_{j=1}^{m_n}\mu_{j,n}^{\frac{1}{2}}(\mu_{j,n}^{\frac{1}{2}}-\mu_j^{\frac{1}{2}})u_j(t)(H_j(U_{j,n}(t))-t)|$$
$$= O_P(\sqrt{\tfrac{\mu_1 \log m_n \log n}{n}})\ \sum_{j=1}^{m_n}u_j(t)\ |\sqrt{n\mu_{j,n}}(H_j(U_{j,n}(t))-t)|$$
$$= O_P(\log m_n\sqrt{\tfrac{\mu_1\log n}{n}}\ m_n)$$
$$= O_P(\log m_n\sqrt{\tfrac{\mu_1\log n}{n\mu_{m_n}}}\ \sum_{j=1}^{m_n}\mu_j^{\frac{1}{2}})\quad (n\to\infty)\ ,$$

which is $o_P(1)$ as $n\to\infty$ because of the rate conditions in the statement of the theorem.

From (2.27), (2.30) and (2.31) it now remains to show that

$$(2.32)\quad \sqrt{\mu_1 \log m_n}\ |\sum_{j=1}^{m_n}\mu_j^{\frac{1}{2}}u_j(t)\sqrt{n\mu_{j,n}}(H_j(U_{j,n}(t))-t)|\ \xrightarrow{P}\ 0$$

as $n\to\infty$ in order to verify (2.13).
To this end, as $|H_{j,n}(U_{j,n}(t))-t|\ \le (n\mu_{j,n})^{-1}$ a.s., the expression in the left hand side of (2.32) is equal to

$$(2.33)\quad \sqrt{\mu_1 \log m_n}\ |\sum_{j=1}^{m_n}\mu_j^{\frac{1}{2}}u_j(t)\sqrt{n\mu_{j,n}}(H_j(U_{j,n}(t))-H_{j,n}(U_{j,n}(t))|$$
$$+O_P((\sum_{j=1}^{m_n}\mu_j^{\frac{1}{2}})\ \sqrt{\frac{\mu_1\log m_n}{n\mu_{m_n}}})\quad (n\to\infty)\ .$$

Now, by (2.11),

$$(2.34)\quad \sqrt{\mu_1 \log m_n}\ |\sum_{j=1}^{m_n}\mu_j^{\frac{1}{2}}u_j(t)\sqrt{n\mu_{j,n}}(H_j(U_{j,n}(t))-H_{j,n}(U_{j,n}(t)))|$$
$$= \sqrt{\mu_1 \log m_n}\ |\sum_{j=1}^{m_n}\mu_j^{\frac{1}{2}}u_j(t)\tilde{B}_{j,n}(H_j(U_{j,n}(t)))|$$
$$+O_P(\mu_1^{\frac{1}{2}}(\sum_{j=1}^{m_n}\mu_j^{\frac{1}{2}})\ (n\mu_{m_n})^{-\frac{1}{8}}(\log m_n)^{\frac{5}{8}}(\log n)^{\frac{1}{2}})\ .$$

Using the modulus of continuity behaviour of Brownian bridges together with (2.27), we get

$$(2.35)\quad \sqrt{\mu_1 \log m_n}|\sum_{j=1}^{m_n} \mu_j^{\frac{1}{2}} u_j(t)\tilde{B}_{j,n}(H_j(U_{j,n}(t)))|$$

$$= \sqrt{\mu_1 \log m_n}|\sum_{j=1}^{m_n} \mu_j^{\frac{1}{2}} u_j(t)\tilde{B}_{j,n}(t)|$$

$$+O_P(\mu_1^{\frac{1}{2}}(\sum_{j=1}^{m_n} \mu_j^{\frac{1}{2}})\ (n\mu_{m_n})^{-\frac{1}{4}}(\log n)^{\frac{1}{2}}(\log m_n)^{\frac{3}{4}})\ .$$

Observe that because of the independence of the $\tilde{B}_{j,n}$ we have that

$$\sum_{j=1}^{m_n} \mu_j^{\frac{1}{2}} u_j(t)\tilde{B}_{j,n}(t) \sim \mathcal{N}(0, t(1-t)\sum_{j=1}^{m_n} \mu_j u_j^2(t))\ .$$

With (C.5)

$$t(1-t)\sum_{j=1}^{m_n} \mu_j u_j^2(t) = O(1) \quad (n \to \infty)$$

and hence

$$(2.36)\quad \sqrt{\mu_1 \log m_n}\ \sum_{j=1}^{m_n} \mu_j^{\frac{1}{2}} u_j(t)\tilde{B}_{j,n}(t) \xrightarrow{P} 0 \quad (n \to \infty).$$

Statements (2.33) - (2.36) yield (2.32), and (2.13) follows from (2.21) - (2.27) and (2.30) - (2.32).

Finally, statement (2.14) follows by (2.13), the behaviour of the modulus of continuity of Brownian bridges, and the independence of the $\tilde{B}_{j,n}$ $(j = 1, \cdots, m_n)$. This concludes the proof of Theorem 3. □

Proof of Corollary 3.

It suffices to show that, under $H_0^{(3)}$, $\sqrt{n\mu_1 \log m_n} \max_{1 \le j \le m_n} (U_j(t) - U_{\cdot}(t)) \to 0 \quad (n \to \infty)$ is implied by (C.7), (C.8) and the rate $n\mu_1 \log m_n(\max_{1 \le j \le m_n} \text{diam}\ (A_{j,n}))^4 \to 0 \quad (n \to \infty)$.

Let $K_{\mathbf{x}} = [a_{\mathbf{x}}, b_{\mathbf{x}}]$ denote the shortt pertaining to $F_{\mathbf{x}}$, let $\alpha_j = \inf_{\mathbf{x} \in A_{j,n}} a_{\mathbf{x}}$, $\tilde{\beta}_j = \sup_{\mathbf{x} \in A_{j,n}} b_{\mathbf{x}}$, and set

$$\beta_j = \alpha_j + U_{\cdot}(t),\ \tilde{\alpha}_j = \tilde{\beta}_j - U_{\cdot}(t)\ .$$

Let a be such that $\alpha_j \le a < a + U_\cdot(t) \le \tilde{\beta}_j$. A Taylor expansion, using $f_{\mathbf{x}}(a_{\mathbf{x}}) = f_{\mathbf{x}}(b_{\mathbf{x}})$ and (C.8), yields that for some $\tilde{a}_{\mathbf{x}} \in (a_{\mathbf{x}} \wedge a,\ a_{\mathbf{x}} \vee a)$ and $\tilde{b}_{\mathbf{x}} \in (b_{\mathbf{x}} \wedge (a + U_\cdot(t)),\ b_{\mathbf{x}} \vee (a + U_\cdot(t)))$ we have

$$\begin{aligned} t - F_{\mathbf{x}}([a, a + U_\cdot(t)]) &= (F_{\mathbf{x}}(a) - F_{\mathbf{x}}(a_{\mathbf{x}})) - (F_{\mathbf{x}}(a + U_\cdot(t)) - F_{\mathbf{x}}(b_{\mathbf{x}})) \\ &= \tfrac{1}{2}(a - a_{\mathbf{x}})^2 f'_{\mathbf{x}}(\tilde{a}_{\mathbf{x}}) - \tfrac{1}{2}(a + U_\cdot(t) - b_{\mathbf{x}})^2 f'_{\mathbf{x}}(\tilde{b}_{\mathbf{x}}) \\ &\le c_6(\alpha_j - \tilde{\alpha}_j)^2 , \end{aligned}$$

and hence,

$$t - F_j([a, a + U_\cdot(t)]) \le \max_{1 \le j \le m_n} (\alpha_j - \tilde{\alpha}_j)^2 c_6 =: \nu_n .$$

Set $\eta = (U_j(t) - U_\cdot(t))/\nu_n$. Since $U_j(t) \ge U_\cdot(t)$, we have $\eta \ge 0$. Observe that for $y_1 \in [\tilde{\alpha}_j, \beta_j]$ and $y_2 \le \alpha_j$ or $y_2 \ge \tilde{\beta}_j$ we have $f_j(y_1) \ge f_j(y_2)$. Hence it readily follows that $[\tilde{\alpha}_j, \beta_j] \subset K_j$. This means that we can find an a as above such that $K_j = [a - \eta\nu_n, a + U_\cdot(t)]$ or such that $K_j = [a, a + U_\cdot(t) + \eta\nu_n]$.

Without loss of generality assume the first equality holds. Observe that the second condition in (C.5) implies that

$$\liminf_{n \to \infty} \ \min_{1 \le j \le m_n} \ \inf_{y \in [a_{j,t}, b_{j,t}]} f_j(y) > 1/c_4 .$$

Hence

$$0 = t - F_j([a - \eta\nu_n,\ a + U_\cdot(t)]) \le \nu_n - F_j([a - \eta\nu_n, a]) \le \nu_n(1 - \eta/c_4) ,$$

which (when $\nu_n > 0$) implies $\eta \le c_4$. This, in combination with $\nu_n\sqrt{n\mu_1 \log m_n} \to 0 \ (n \to \infty)$, completes the proof. □

Acknowledgement. We thank Arthur van Soest for a useful conversation.

REFERENCES.

Cressie, N. (1980). The asymptotic distribution of the scan statistic under uniformity. *Ann. Probab.* **18**, 828-840.

Cressie, N. (1987). Using the scan statistic to test for uniformity. In: *Goodness-of-Fit. Colloquia Mathematica Societatis Janós Bolyai* **45**, 87-100. North Holland, Amsterdam.

Csörgő, M. and Révész, P. (1981). *Strong Approximations in Probability and Statistics.* Academic Press, New York.

Deheuvels, P. and Révész, P. (1987). Weak laws for the increments of Wiener processes, Brownian bridges, empirical processes and partial sums of i.i.d. rv's. In: *Mathematical Statistics and Probability Theory, Vol. A* (M.L. Puri et al. eds.) 69-88. Reidel, Dordrecht.

Dijkstra, J.B., Rietjens, T.J.M. and Steutel, F.W. (1984). A simple test for uniformity. *Statist. Neerlandica* **38**, 33-44.

Einmahl, J.H.J. and Mason, D.M. (1992). Generalized quantile processes. *Ann. Statist.* **20**, 1062-1078.

Grübel, R. (1988). The length of the shorth. *Ann. Statist.* **16**, 619-628.

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

Kim, J. and Pollard, D. (1990). Cube root asymptotics. *Ann. Statist.* **18**, 191-219.

Komlós, J., Major, P. and Tusnády, G. (1975). An approximation of partial sums of independent RV's and the sample DF. I. *Z. Wahrsch. Verw. Gebiete* **32**, 111-131.

Naus, J.I. (1966). A power comparison of two tests of nonrandom clustering. *Technometrics* **8**, 493-517.

Naus, J.I. (1982). Approximations for distributions of scan statistics. *J. Amer. Statist. Assoc.* **77**, 177-183.

Resnick, S.I. (1987). *Extreme Values, Regular Variation, and Point Processes.* Springer, New York.

Rousseeuw, P. (1984). Least median of squares regression. *J. Amer. Statist. Assoc.* **79**, 871-880.

Rousseeuw, P. and Leroy, A. (1988). A robust scale estimator based on the shortest half. *Statist. Neerlandica* **42**, 103-116.

Dept. of Mathematics
Catholic University Leuven
Celestijnenlaan 200 B
3001 Heverlee
Belgium

Dept. of Mathematics and Computing Science
Eindhoven University of Technology
P.O. Box 513
5600 MB Eindhoven
The Netherlands

List of COSOR-memoranda - 1993

Number	Month	Author	Title
93-01	January	P. v.d. Laan C. v. Eeden	Subset selection for the best of two populations: Tables of the expected subset size
93-02	January	R.J.G. Wilms J.G.F. Thiemann	Characterizations of shift-invariant distributions based on summation modulo one.
93-03	February	Jan Beirlant John H.J. Einmahl	Asymptotic confidence intervals for the length of the shortt under random censoring.
93-04	February	E. Balas J. K. Lenstra A. Vazacopoulos	One machine scheduling with delayed precedence constraints
93-05	March	A.A. Stoorvogel J.H.A. Ludlage	The discrete time minimum entropy H_∞ control problem
93-06	March	H.J.C. Huijberts C.H. Moog	Controlled invariance of nonlinear systems: nonexact forms speak louder than exact forms
93-07	March	Marinus Veldhorst	A linear time algorithm to schedule trees with communication delays optimally on two machines
93-08	March	Stan van Hoesel Antoon Kolen	A class of strong valid inequalities for the discrete lot-sizing and scheduling problem
93-09	March	F.P.A. Coolen	Bayesian decision theory with imprecise prior probabilities applied to replacement problems
93-10	March	A.W.J. Kolen A.H.G. Rinnooy Kan C.P.M. van Hoesel A.P.M. Wagelmans	Sensitivity analysis of list scheduling heuristics
93-11	March	A.A. Stoorvogel J.H.A. Ludlage	Squaring-down and the problems of almost-zeros for continuous-time systems
93-12	April	Paul van der Laan	The efficiency of subset selection of an ε-best uniform population relative to selection of the best one
93-13	April	R.J.G. Wilms	On the limiting distribution of fractional parts of extreme order statistics

Number	Month	Author	Title
93-14	May	L.C.G.J.M. Habets	On the Genericity of Stabilizability for Time-Day Systems
93-15	June	P. van der Laan C. van Eeden	Subset selection with a generalized selection goal based on a loss function
93-16	June	A.A. Stoorvogel A. Saberi B.M. Chen	The Discrete-time H_∞ Control Problem with Strictly Proper Measurement Feedback
93-17	June	J. Beirlant J.H.J. Einmahl	Maximal type test statistics based on conditional processes