Poisson and Gaussian approximation of weighted local empirical processes
Einmahl, J.H.J.

Published: 01/01/1995

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 author's 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

Citation for published version (APA):

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 ?

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Poisson and Gaussian approximation of weighted local empirical processes

J.H.J. Einmahl

Eindhoven, September 1995
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: Main Building 9.10
Telephone: 040-47 3130
E-mail: wsbscaro@win.tue.nl or wscosor@win.tue.nl

ISSN 0926 4493
POISSON AND GAUSSIAN APPROXIMATION OF 
WEIGHTED LOCAL EMPIRICAL PROCESSES 

John H.J. Einmahl*
Department of Mathematics and Computing Science
Eindhoven University of Technology
P.O. Box 513, 5600 MB Eindhoven, The Netherlands
e-mail: einmahl@win.tue.nl

We consider the local empirical process indexed by sets, a greatly generalized version of the well-studied uniform tail empirical process. We show that the weak limit of weighted versions of this process is Poisson under certain conditions, whereas it is Gaussian in other situations. Our main theorems provide a unified approach to a number of asymptotic distributional results for weighted empirical processes, which up to now appeared to be isolated facts. Our results are likely to have applications in local statistical procedures, e.g., in the study of multivariate extreme values.

Mathematics Subject Classification (1991): 60F05, 60F17, 62E20, 62G30.
Running head: Local empirical processes.

* Research partially supported by European Union HCM grant ERB CHRX-CT 940693.
1 Introduction

Let \( X_1, X_2, \cdots \), be i.i.d. real-valued random variables with common distribution function \( F \). Choose a sequence \( \{a_n\}_{n=1}^{\infty} \) of positive constants converging to 0 and \( x \in \mathbb{R} \). Introduce the local empirical function at \( x \)

\[
L_{n,x}(t) = \frac{1}{n} \sum_{i=1}^{n} 1_{[x-ta_n,x+ta_n]}(X_i), \quad t \in [0,T], \tag{1.1}
\]

for some positive \( T \). Observe that \( L_{n,x}(1)/(2a_n) \) is the naive density estimator of the density \( f \) of \( F \) at \( x \), assuming that \( f \) exists and is continuous at \( x \). If it is also assumed that \( na_n \rightarrow \infty \), classical weak convergence theory shows that the local empirical process

\[
v_n(t) = \left( \frac{a_n}{n} \right)^{1/2} (L_{n,x}(t) - EL_{n,x}(t)), \quad t \in [0,T], \tag{1.2}
\]

converges weakly in \( D[0,T] \) to \( (2f(x))^{1/2}W \), where \( W \) is a standard Wiener process. In Deheuvels and Mason (1994) a detailed study of functional laws of the iterated logarithm for a generalized version of this local empirical process as well as examples of its applicability, e.g., to density estimation, are presented.

Now consider the tail empirical function

\[
T_n(t) = \frac{1}{n} \sum_{i=1}^{n} 1_{[0,ta_n]}(U_i), \quad t \in [0,T], \quad T > 0, \tag{1.3}
\]

where \( U_1, U_2, \cdots \), are i.i.d. uniform-(0,1) random variables. It can be easily verified that the tail empirical process

\[
w_n(t) = \left( \frac{a_n}{n} \right)^{1/2} (T_n(t) - ta_n), \quad t \in [0,T], \tag{1.4}
\]

converges weakly in \( D[0,T] \) to \( W \), whenever \( a_n \rightarrow 0 \) and \( na_n \rightarrow \infty \). The notion of the tail empirical process was introduced in Mason (1988). In Einmahl and Mason (1988), Deheuvels and Mason (1990) and Einmahl (1992) its almost sure limiting behaviour is exhaustively studied and in Deheuvels and Mason (1991) numerous applications of its strong limiting properties in probability theory and statistics are detailed.

Much is known about the asymptotic distribution of weighted versions of the tail empirical process. It is shown in Mason (1983) that when \( a_n = 1/n, n \in \mathbb{N} \), then as \( n \rightarrow \infty \)

\[
\sup_{0 \leq t \leq n} \left| w_n(t) \right|/t^\nu \to_d \sup_{t > 0} \left| N(t) - t \right|/t^\nu, \tag{1.5}
\]

for all \( \frac{1}{2} < \nu \leq 1 \), where \( N \) is a standard rate 1 Poisson process on \([0,\infty)\). Under the assumption that \( a_n \rightarrow 0 \) and \( na_n \rightarrow \infty \), it is proved in Csörgő and Mason (1985) that as \( n \rightarrow \infty \)
for all $0 \leq \nu < \frac{1}{2}$, whereas it is established in Mason (1985) that as $n \to \infty$

\[
\sup_{0 < t \leq t_n} \frac{|w_n(t)|/t^\nu}{d} \to_d \sup_{0 < t \leq 1} \frac{|W(t)|/t^\nu}{d}, \tag{1.7}
\]

for all $\frac{1}{2} < \nu \leq 1$. Notice that the original results (1.5)-(1.7) are stated in terms of the (global) empirical process and not in terms of the tail empirical process. However, (1.5)-(1.7) and the remainder of this paper show that it is natural and sensible to consider these and related results from a tail or local empirical process viewpoint.

We will show that it is much more effective to study local and tail empirical processes within a greatly generalized setup. Towards this end let $S$ be a complete separable metric space (Polish space) with Borel $\sigma$-field $\mathcal{S}$. For each $n \in \mathbb{N}$, let $\{X_{n,i} : i = 1, \ldots, n\}$ be i.i.d. random elements taking values in $(S, \mathcal{S})$. Denote the empirical measure at stage $n$ by

\[
P_n(B) = \frac{1}{n} \sum_{i=1}^{n} 1_B(X_{n,i}), \quad B \in \mathcal{S}, \tag{1.8}
\]

and denote the true probability measure with $P_n$, i.e., $P_n(B) = P(X_{n,1} \in B), \quad B \in \mathcal{S}$. Further, let $\mathcal{A}$ be a subset of $\mathcal{S}$, which satisfies properties to be specified later on. We shall study the weighted approximation by Poisson processes or Gaussian processes of suitably normed versions of the centered empirical measure indexed by $\mathcal{A}$

\[
P_n(A) - P_n(A), \quad A \in \mathcal{A}. \tag{1.9}
\]

We shall call such a process, when appropriately normed, a local empirical process, whenever $P_n(A)$ converges to 0 for all $A \in \mathcal{A}$.

Here are some important special cases of local empirical processes.

(I) Setting, with $x \in \mathbb{R}^d$, $d \in \mathbb{N}$,

\[
X_{n,i} = (X_i - x)/a_n, \quad \text{for } i = 1, \ldots, n, \quad n \in \mathbb{N},
\]

where the $X_i$ are i.i.d. random vectors in $\mathbb{R}^d$, results, after norming in a generalization of the local empirical process in (1.2).

(II) Set

\[
X_{n,i} = U_i/a_n, \quad \text{for } i = 1, \ldots, n, \quad n \in \mathbb{N},
\]

where the $U_i$ are as in (1.3) and choose \( \mathcal{A} = \{[0, t] : t \in (0, T]\} \). With these choices we obtain the tail empirical process in (1.4).
Set

\[ X_{n,i} = (X_i - b_n)/a_n, \text{ for } i = 1, \ldots, n, \ n \in \mathbb{N}, \]

where the \( X_i \) are as in (I), the \( a_n, b_n \in \mathbb{R}^d \) are norming vectors and, e.g., \( \mathcal{A} = \{(\mathbb{R}_{-\infty} \times \cdots \times (-\infty, t_d])^c: t_i \geq T_i \text{ for all } 1 \leq i \leq d\} \). This results in a point process, important in (multivariate) extreme value theory; see, e.g., Reiss (1993, chapter 6).

In sections 2 and 3 we will formulate weighted approximations to the local empirical process indexed by \( \mathcal{A} \), which constitute our main results. It will be seen that whether an approximation by a Poisson process (section 2) or by a Gaussian process (section 3) is appropriate depends on whether \( nP(n)(A), A \in \mathcal{A} \), tends to a finite positive constant or to infinity, respectively. Among other results, the asymptotic distributions described in (1.5)–(1.7), along with their bivariate extensions (Csörgő and Horváth (1990), Horváth (1991)) follow readily from our approximations. We shall also present applications of our results to (multivariate) extreme value theory. Finally, all the proofs of our main results are given in section 4.

### 2 Approximation with Poisson processes

Before presenting the approximation of the local empirical processes by Poisson processes, we first give a sharp result which guarantees the finiteness of the supremum of the relevant weighted Poisson process, hence making our approximation useful. Recall the notation of section 1, see e.g. (1.8) and (1.9). To state our results, we need more notation and assumptions.

We will use that

(A.1) \( \mathcal{A} \) is a VC (Vapnik-Chervonenkis) class with index \( v \) (for the definition see, e.g., Alexander (1984)),

and, to avoid measurability problems, that

(A.2) there exists a countable subset \( \mathcal{D} \) of \( \mathcal{A} \) such that for every \( A \in \mathcal{A} \) there exists a

sequence \( \{D_m\}_{m=1}^{\infty} \) in \( \mathcal{D} \) such that for every \( x \in S \), \( 1_{D_m}(x) \rightarrow 1_A(x) \), as \( m \rightarrow \infty \).

Let \( q \), the weight function, be a positive real valued function defined on \( \mathcal{A} \). As in Alexander (1987), set for \( u > 0 \)

\[ E(u) = \bigcup \{ A \in \mathcal{A} : q(A) \leq u \} . \] (2.1)

We shall assume that
for each \( u > 0 \), \( E(u) \) is closed.

Let \( \mu \) be a \( \sigma \)-finite measure on \((S, S)\) and set

\[
\lambda(u) = \mu(E(u)) \quad \text{and} \quad h(u) = \inf\{v : \lambda(v) \geq u\}.
\]

We will also need that

\[
\lambda \text{ is continuous.}
\]

Now let \( N \) be a Poisson process on \((S, S)\) with mean measure \( \mu \). For \( u \geq 0 \), set

\[
M_q(u) = \sup\{|N(A) - \mu(A)|/q(A) : A \in \mathcal{A}, q(A) > u\} \vee 0.
\]

PROPOSITION 2.1. Assume that \( \mathcal{A} \) satisfies (A.1) and (A.2) and that \( q \) satisfies (A.3), (A.4) and

\[
\int_1^{\infty} \frac{1}{u} \exp \left( -\frac{ch^2(u)}{u} \right) du < \infty, \text{ for all } c > 0.
\]

Then

\[
\lim_{u \to \infty} M_q(u) = 0 \text{ a.s.}
\]

and

\[
M_q(1) < \infty \text{ a.s.}
\]

Furthermore, if in addition,

\[
\lim_{u \to 0} \lambda(u) = 0
\]

and for some \( b > 0 \)

\[
q(A) \geq b\mu(A) \text{ for all } A \in \mathcal{A} \text{ with } q(A) \leq 1,
\]

then

\[
M_q(0) < \infty \text{ a.s.}
\]

Now we present the main result of this section along with a number of interesting corollaries (and their proofs), which demonstrate how powerful the theorem is.

THEOREM 2.1. Let \( P_n, n \in \mathbb{N} \), and \( \mu \) be as before and write \( \mu_n = nP_n \). Assume that we have, in addition to (A.1)-(A.5),
that for some $C > 0$, for all large $n$

(A.9) \[ \mu_n(E(u)) \leq C \lambda(u), \text{ for all } u > 0, \]

and that

(A.10) \[ \sup_{A \in \mathcal{A}} \frac{|\mu_n(A) - \mu(A)|}{q(A)} \to 0 \text{ (n \to \infty), for all } u > 0. \]

Then the probability space on which the $\{X_{n,i}, i = 1, \cdots, n\}, n \in \mathbb{N}$, are defined, can be enlarged to include a sequence of Poisson processes $\{N_n\}_{n=1}^{\infty}$ on $(\mathcal{S}, \mathcal{S})$, all with mean measure $\mu$, such that, as $n \to \infty$,

\[ \sup_{A \in \mathcal{A}} \frac{|(n P_n(A) - \mu_n(A)) - (N_n(A) - \mu(A))|}{q(A)} \to p 0. \quad (2.7) \]

**COROLLARY 2.1.** Under the assumptions of Theorem 2.1 and (A.6)-(A.7) we have, as $n \to \infty$,

\[ \sup_{A \in \mathcal{A}} \frac{|n P_n(A) - \mu_n(A)|}{q(A)} \to_d \sup_{A \in \mathcal{A}} \frac{|N(A) - \mu(A)|}{q(A)} <_{\text{a.s.}} \infty. \quad (2.8) \]

**COROLLARY 2.2.** (cf. special case (I), section 1). Let $X_1, \cdots, X_n$ be i.i.d. random vectors in $\mathbb{R}^d$, $d \in \mathbb{N}$, with fixed common probability measure $P$, let $x \in \mathbb{R}^d$ and $a_n = (a_{n,1}, \cdots, a_{n,d})$, $a_{n,j} > 0$, $j = 1, \cdots, d$. Define

\[ X_{n,i} = \frac{X_i - x}{a_n} = \left( \frac{X_{i,1} - x_1}{a_{n,1}}, \ldots, \frac{X_{i,d} - x_d}{a_{n,d}} \right). \quad (2.9) \]

Let $A \subset \mathcal{B}$, the Borel sets on $\mathbb{R}^d$, and let $\tilde{P}_n$ be the empirical measure of the $X_i$, $i = 1, \cdots, n$. Then under the assumptions of Theorem 2.1

\[ \sup_{A \in \mathcal{A}} \frac{|n(\tilde{P}_n(x + a_n A) - \tilde{P}(x + a_n A)) - (N_n(A) - \mu(A))|}{q(A)} \to_p 0, \quad (2.10) \]

where $a_n A = \{(a_{n,1} y_1, \cdots, a_{n,d} y_d) : (y_1, \cdots, y_d) \in A\}$. If, in addition, (A.6) and (A.7) hold, then

\[ \sup_{A \in \mathcal{A}} \frac{n|\tilde{P}_n(x + a_n A) - \tilde{P}(x + a_n A)|}{q(A)} \to_d \sup_{A \in \mathcal{A}} \frac{|N(A) - \mu(A)|}{q(A)} <_{\text{a.s.}} \infty. \quad (2.11) \]

**COROLLARY 2.3** (cf. Theorem 3 in Csõrgõ and Horváth (1992)). Consider the situation of special case (II), section 1, i.e. the tail empirical process but replace $(0, T]$ by $(0, \infty)$. Write for simplicity $q(t)$ for $q([0, t])$ and assume that $q$ is strictly increasing and left-continuous. Then
\[
\int_{-\infty}^{\infty} \frac{1}{u} \exp \left(-e^{q^2(u)} \right) du < \infty, \text{ for all } c > 0,
\]
implies
\[
\sup_{t > 0} \left| nF_n(t/n) - (t \wedge n) - (N_n(t) - t)/q(t) \right| \to_{\mathcal{D}} 0,
\]
where \( F_n \) is the uniform empirical distribution function \( (F_n(t) = \frac{1}{n} \sum_{i=1}^{n} I_{[0,t]}(X_i), \ 0 \leq t \leq 1) \) and the \( N_n \) are standard rate 1 Poisson processes on \([0, \infty)\). In particular, this result in conjunction with Proposition 2.1 yields (1.5).

The next corollary is a generalization with respect to the dimension of (1.5) and a further specialization of Corollary 2.2.

**COROLLARY 2.4** (cf. Remark 2.2 in Csörgő and Horváth (1990)). Consider the situation of Corollary 2.2 and let now \( \bar{P} \) be the uniform distribution on \((0,1)^d\) and take \( x = (0, \ldots, 0) \). Denote with \( F \) and \( F_n \) the distribution functions, corresponding with \( P \) and \( \bar{P}_n \), respectively. Set \( A = \{(0, t) = \prod_{j=1}^{d} [0, t_j]; t = (t_1, \ldots, t_d), t_j \geq c_j, j = 1, \ldots, d\} \) for \((c_1, \ldots, c_d) \in (0, \infty)^d\). Let \( N \) be a standard rate 1 Poisson process on \([0, \infty)^d\). Assume \( a_{n,j} \to 0 \) \((n \to \infty)\), for all \( j = 1, \ldots, d \), and \( \prod_{j=1}^{d} a_{n,j} = \frac{1}{n} \) for all \( n \in \mathbb{N} \). Then, for all \( \nu_j > \frac{1}{2}, j = 1, \ldots, d \),
\[
\sup_{t \in \prod_{j=1}^{d} [c_j, \infty)} \frac{n|F_n(a_n t) - F(a_n t)|}{\prod_{j=1}^{d} t_{\nu_j}^j} \to_{\mathcal{D}} \sup_{t \in \prod_{j=1}^{d} [c_j, \infty)} \frac{|N(t) - \prod_{j=1}^{d} t_{\nu_j}^j|}{\prod_{j=1}^{d} t_{\nu_j}^j} <_{\text{a.s.}} \infty. \tag{2.14}
\]

**COROLLARY 2.5** (cf. Theorems 1 and 2 in Horváth (1991)). Consider the situation of Corollary 2.2 with \( d = 2 \). Let \( \bar{P} \) be the uniform distribution on \((0,1)^2\) and take \( x = (0, 0) \). Denote with \( F \) and \( F_n \) the distribution functions corresponding with \( P \) and \( \bar{P}_n \), respectively. Set \( A = \{(0, t_1) \times [0, t_2]; 0 < t_1 \leq 1, 0 < t_2 < \infty\} \). Let \( N \) be a standard rate 1 Poisson processes on \([0, \infty)^2\) and let \( \{b_n\}_{n=1}^{\infty} \) be a sequence of numbers in \((0,1] \) with \( nb_n \to \infty \) \((n \to \infty)\). Then for all \( \nu_1, \nu_2 \) satisfying \( \frac{1}{2} < \nu_1 < \nu_2 \leq 1 \)
\[
\sup_{0 < t_1 \leq 1, 0 < t_2 < \infty} \frac{n|F_n(b_n t_1, \frac{t_1}{b_n} t_2) - F(b_n t_1, \frac{t_1}{b_n} t_2)|}{t_1^\nu_1 t_2^\nu_2} \to_{\mathcal{D}} \sup_{0 < t_1 \leq 1, 0 < t_2 < \infty} \frac{|N(t_1, t_2) - t_1 t_2|}{t_1^\nu_1 t_2^\nu_2} <_{\text{a.s.}} \infty. \tag{2.15}
\]

**COROLLARY 2.6.** (cf. special case (III), section 1). Let \( X_1, \ldots, X_n \) be i.i.d. real-valued random variables with fixed common distribution \( F \), which is in the domain of max-attraction of an extreme value distribution \( G \). Denote the normalizing constants with \( a_n > 0 \) and \( b_n \), i.e., we have with \( M_n \) the maximum of the \( X_i, 1 \leq i \leq n \),
for all \( t \in \mathbb{R} \). Define
\[
X_{n,i} = \frac{X_i - b_n}{a_n}.
\]
Let \( \mathcal{A} = \{(t, \infty) : 0 < G(t) < 1\} \) and let \( F_n \) be the empirical distribution function of the \( X_i, 1 \leq i \leq n \). Define \( \mu \) by \( \mu((t, \infty)) = -\log G(t), G(t) > 0 \). Assume (A.3)--(A.5), and (A.8)--(A.10) hold, then
\[
\sup_{t: 0 < G(t) < 1} \left| n(F_n(a_n t + b_n) - F(a_n t + b_n)) + (N_n((t, \infty)) - \mu((t, \infty))) \right| / q((t, \infty)) \rightarrow P: 0. \tag{2.17}
\]

**Proof of Corollary 2.1.** Corollary 2.1 follows by combining Theorem 2.1 and Proposition 2.1. \( \Box \)

**Proof of Corollary 2.2.** Corollary 2.2 is a special case of Theorem 2.1 and Corollary 2.1. \( \Box \)

**Proof of Corollary 2.3.** The situation in Corollary 2.3 is a further specialization of Corollary 2.2. For (2.13) we have to check the conditions of Theorem 2.1. This however is rather trivial and will be omitted. (Note that \( h \equiv q \) in this situation, hence (2.12) trivially implies (A.5).) To see (1.5) observe that for \( \nu > \frac{1}{2} \).
\[
\int_1^{\infty} \frac{1}{u} \exp \left( -\frac{c \nu^2 u}{u} \right) du = \int_1^{\infty} \frac{1}{u} \exp(-c u^{2\nu-1}) du < \infty.
\]
The condition \( \nu \leq 1 \) is needed for (A.7), which in turn is needed for the a.s. finiteness of the right hand side of (1.5). \( \Box \)

**Proof of Corollary 2.4.** For convenience we take \( c_j = 1 \), for all \( j = 1, \ldots, d \); it is easy to adapt the proof for general positive \( c_j \). Again we have to check (A.1)--(A.10). It is straightforward to do so, except for (A.5), which we discuss briefly now. Set \( \nu = \min_{j=1, \ldots, d} \nu_j \). Observe that \( \lambda \leq \bar{\lambda} \), where \( \bar{\lambda} \) is defined as \( \lambda \) in (2.1)--(2.2), but with \( q([0, t]) = \prod_{j=1}^d t_j^{\nu_j} \) replaced by \( \bar{q}([0, t]) = (\prod_{j=1}^d t_j)^{\nu} \). From this observation it is not hard to show that \( \lambda(u) = O(u^{1/\nu} \log u)^{d-1} \), for \( u \to \infty \), hence \( u^\nu / (\log u)^{\nu(d-1)} = O(h(u)) \). But this yields (A.5), since \( \nu > \frac{1}{2} \). \( \Box \)

**Proof of Corollary 2.5.** We have to check (A.1)--(A.10). If \( \nu_1 \leq 0 \) everything is easy, so we can confine ourselves to the case \( \nu_1 > 0 \), i.e. we have \( \nu_1 > 0, \frac{1}{2} < \nu_2 \leq 1 \), and \( \nu_1 < \nu_2 \). Now still most of the conditions are trivially fulfilled, but in this corollary we have, in contrast with the previous two, a situation where \( E(u) \) is unbounded. Therefore we will check (A.5)--(A.8) and (A.10).

The function \( \lambda \) can easily be calculated explicitly:
\[
\lambda(u) = u^{1/\nu_2} \int_0^1 t_1^{-\nu_1/\nu_2} dt_1 = u^{1/\nu_2} / (1 - \nu_1/\nu_2).
\]
This yields (A.5) and (A.6). Now we prove (A.7) for \( b = 1 \). Let \( q(A) = u \leq 1 \). So
we have $t_1^2 t_2^2 = u$, which implies $t_2 = u^{1/2} t_1^{1-v_2/v_1}$. Hence $\mu(A) = t_1 t_2 = u^{1/2} t_1^{1-v_2/v_1} \leq u^{1/2} \leq u = q(A)$. For (A.8) observe that

$$\sup_{B \in S} \sup_{B \subseteq E(u)} |\mu_n(B) - \mu(B)| = \mu(E(u) \setminus [0, 1] \times [0, nb_n]) ,$$

which tends to zero ($n \to \infty$) since $\mu(E(u)) = \lambda(u) < \infty$. Remains to show that (A.10) holds. Take $u > 0$ and for $A \in \mathcal{A}$ with $q(A) \leq u$ write $\tilde{u} = q(A)$. Then we have

$$|\mu_n(A) - \mu(A)| \leq \mu(E(u) \setminus [0, 1] \times [0, nb_n]) = \frac{\tilde{u}^{1/v_1}}{(\nu_2/\nu_1 - 1)(nb_n)^{\nu_2/\nu_1 - 1}} .$$

Hence

$$\sup_{A \in \mathcal{A}} \sup_{q(A) \leq u} \frac{|\mu_n(A) - \mu(A)|}{q(A)} \leq \sup_{0 < \tilde{u} \leq u} \frac{\tilde{u}^{1/v_1 - 1}}{(\nu_2/\nu_1 - 1)(nb_n)^{\nu_2/\nu_1 - 1}}$$

$$= \frac{u^{1/v_1 - 1}}{(\nu_1/\nu_1 - 1)(nb_n)^{\nu_2/\nu_1 - 1}} = 0 \ (n \to \infty) .$$

This completes the proof. \qed

PROOF OF COROLLARY 2.6. This corollary follows immediately from Theorem 2.1. It suffices to note the well-known fact that (2.16) implies (take logarithms)

$$n(1 - F(a_n t + b_n)) = -\log G(t) \ (n \to \infty) ,$$

for all $t$ with $G(t) > 0$. \qed

DISCUSSION OF THE RESULTS. First of all let us emphasize that the local empirical process approach is the natural one for the results in this paper, since then the ‘index’ $A$ (or $t$) is the same for the empirical process and its limit. Presenting results like, e.g., (1.5) or (2.11) for global empirical processes is of course possible, but then these results look less obvious, since there is no corresponding approximation result (like (2.13) for (1.5)) then. Also note the importance of Proposition 2.1 since a statement like (2.8) is rather weak if there is no guarantee that the limiting random variable is finite almost surely. E.g., if we replace $\prod_{j=1}^{d} [c_j, \infty)$ by $(0, \infty)^d$, in Corollary 2.4 it is easy to see that the right hand side of (2.14) is equal to infinity almost surely (cf. Theorem 2.2 and Remark 2.2 in Csörgő and Horváth (1990)). For the proper result in that situation, see Einmahl (1995). All this explains why Corollary 2.4 is not a complete multivariate analogue of (1.5), i.e. why some truncation is needed. Finally, for more background about these results, the reader might wish to consult the books by Reiss (1993) or Resnick (1987) and some of the references therein.
3 Approximation with Gaussian processes

In this section we will approximate the local empirical processes by Gaussian processes. The structure of this section and, strikingly, also the results will be quite similar to section 2. We will also use the notation and conditions of the previous section(s).

Let $W$ be a Wiener process with ‘time’ $\mu$ on $\mathcal{A}$, i.e. a continuous mean zero Gaussian process with covariance structure

$$EW(A)W(A') = \mu(A \cap A'), \quad A, A' \in \mathcal{A}. \quad (3.1)$$

For all $u \geq 0$, set

$$\widetilde{M}_q(u) = \sup\{|W(A)|/q(A) : A \in \mathcal{A}, q(A) > u\} \vee 0. \quad (3.2)$$

**Proposition 3.1.** Assume that (A.1)-(A.5) hold, then

$$\lim_{u \to \infty} \widetilde{M}_q(u) = 0 \text{ a.s.} \quad (3.3)$$

and

$$\widetilde{M}_q(1) < \infty \text{ a.s.} \quad (3.4)$$

Throughout this section let $\{k_n\}_{n=1}^{\infty}$ be a sequence of positive numbers satisfying $k_n \to \infty$ and $k_n/n \to 0$ ($n \to \infty$). Write $\mu_n = nP_n/k_n$ and define the local empirical process by

$$v_n(A) = k_n^{1/2} \left( \frac{nP_n(A)}{k_n} - \mu_n(A) \right), \quad A \in \mathcal{A}. \quad (3.5)$$

Set $\mathcal{A}(u) = \{A \in \mathcal{A} : q(A) \leq u\}$, $\overline{\mathcal{A}}(u) = \mathcal{A}(u) \cup \{E(u)\}$, and $\mathcal{A}'(u) = \{A \cap A' : A, A' \in \overline{\mathcal{A}}(u)\}$.

**Theorem 3.1.** Assume that we have (A.1)-(A.5), (A.9) and either

(A.11) $q(A) > 1$ for all $A \in \mathcal{A}$, and

$$\sup_{A \in \mathcal{A}(u)} |\mu_n(A) - \mu(A)| \to 0 \ (n \to \infty) \text{ for all } u > 0,$$

or

(A.12) for all large $u : \mu_n(A) = \mu(A)$ for all $A \in \overline{\mathcal{A}}(u)$, for large $n$, $\mu(A)/q(A), \ A \in \mathcal{A}(u)$, is bounded, and

$$m^{1/2}(\nu_m - \nu)/q \to_d B/\nu \ (m \to \infty), \text{ on } \mathcal{A}(u),$$
where \( \nu = \mu/\lambda(u) \), \( \nu_m \) is the empirical measure based on a random sample of size \( m \) from \( \nu \), and \( B \) is a bounded, continuous mean zero Gaussian process with covariance structure

\[
EB(A)B(A') = \nu(A \cap A') - \nu(A)\nu(A'), \quad A, A' \in \mathcal{A}(\mu).
\]  

(3.6)

Then the probability space on which the \( \{X_{n,i}, i = 1, \ldots, n\} \), \( n \in \mathbb{N} \), are defined can be enlarged to include a sequence of Wiener processes \( \{W_n\}_{n=1}^{\infty} \), on \( \mathcal{A} \), all with 'time' \( \mu \), such that, as \( n \to \infty \),

\[
\sup_{A \in \mathcal{A}} \frac{|v_n(A) - W_n(A)|}{q(A)} \to_P 0.
\]  

(3.7)

COROLLARY 3.1. Under the assumptions of Theorem 3.1 we have, as \( n \to \infty \),

\[
\sup_{A \in \mathcal{A}} \frac{|v_n(A)|}{q(A)} \to_d \sup_{A \in \mathcal{A}} \frac{|W(A)|}{q(A)} <_{a.s.} \infty
\]  

(3.8)

COROLLARY 3.2 (cf. special case (1), section 1). In the setup of Corollary 2.2 we have under the assumptions of Theorem 3.1.

\[
\sup_{A \in \mathcal{A}} \frac{|n^{1/2}(-P_n(x + a_nA) - \tilde{P}(x + a_nA)) - W_n(A)|}{q(A)} \to_P 0
\]  

(3.9)

and

\[
\sup_{A \in \mathcal{A}} \frac{|n^{1/2}(-P_n(x + a_nA) - \tilde{P}(x + a_nA))|}{q(A)} \to_d \sup_{A \in \mathcal{A}} \frac{|W(A)|}{q(A)} <_{a.s.} \infty.
\]  

(3.10)

COROLLARY 3.3 (cf. Theorem 5 in Csörgö and Horváth (1992)). In the setup of Corollary 2.3 we have under (2.12)

\[
\sup_{t \geq 1} \frac{n}{k_n^{1/2}}\left(F_n\left(t\frac{k_n}{n}\right) - \left(t\frac{k_n}{n} \wedge 1\right) - W_n(t)\right)\to_P 0,
\]  

(3.11)

where the \( W_n \) are standard Wiener processes on \([0, \infty)\). In particular, this result in conjunction with Proposition 3.1 yields (1.7).

COROLLARY 3.4 (cf., e.g., Theorem 3 in Einmahl (1992) and Theorem 1 in Csörgö and Horváth (1992)). In the setup of Corollary 2.3 we have under

\[
\int_{0}^{1} \exp \left(-c\frac{q^2(u)}{u}\right) du < \infty, \text{ for all } c > 0
\]  

(3.12)

that
where $W_n$ is as in the previous corollary. In particular, this results yields (1.6).

**COROLLARY 3.5** (cf. Theorem 2.3 in Csörgő and Horvath (1990)). In the setup of Corollary 2.4, but with $\prod_{j=1}^{d} a_{n,j} = k_n/n$ for all $n \in N$, we have for all $\nu_j > \frac{1}{2}$, $j = 1, \cdots, d$,

$$
\sup_{t \in \Pi_{j=1}^{d} \in [c_j, \infty)} \frac{n}{k_n^{1/2}} \frac{n}{t_j^{\nu_j}} \frac{|F_n(a_n t) - F(a_n t)|}{\prod_{j=1}^{d} t_j^{\nu_j}} \to_d \sup_{t \in \Pi_{j=1}^{d} \in (0, \infty)} \frac{|W(t)|}{\prod_{j=1}^{d} t_j^{\nu_j}} <_{a.s.} \infty ,
$$

where $W$ is a standard Wiener process on $[0, \infty)^d$, i.e., in particular $EW(s)W(t) = s \wedge t$, where $\wedge$ denotes the coordinatewise minimum.

**COROLLARY 3.6.** (cf. Theorem 2.1 in Csörgő and Horvath (1990)). In the setup of Corollary 2.4, but with $0 < a_{n,j} \leq 1$ and $\prod_{j=1}^{d} a_{n,j} = k_n/n$ for all $n \in N$, we have for all $\nu_j < \frac{1}{2}, j = 1, \cdots, d$,

$$
\sup_{t \in \Pi_{j=1}^{d} \in (0, \infty)} \frac{n}{k_n^{1/2}} \frac{n}{t_j^{\nu_j}} \frac{|F_n(a_n t) - F(a_n t)|}{\prod_{j=1}^{d} t_j^{\nu_j}} \to_d \sup_{t \in \Pi_{j=1}^{d} \in (0, \infty)} \frac{|W(t)|}{\prod_{j=1}^{d} t_j^{\nu_j}} <_{a.s.} \infty ,
$$

where $W$ is as in the previous corollary.

**COROLLARY 3.7** (cf. Theorem 3 in Horvath (1991)). In the setup of Corollary 2.5 but with $\mathcal{A} = \{(0,t_1] \times [0,t_2] : 0 < t_1 \leq 1, t_2 \geq c\}$, $c \in (0, \infty)$, and with $\frac{k_n}{nb_n} \to 0$ (and $k_n \to \infty$), we have for all $\nu_1 < \frac{1}{2}$ and $\nu_2 > \frac{1}{2}$,

$$
\sup_{0 < t_1 \leq 1} \frac{n}{k_n^{1/2}} \frac{n}{t_1^{\nu_1}} \frac{n}{t_2^{\nu_2}} \frac{|F_n(b_n t_1, \frac{k_n}{nb_n} t_2) - F(b_n t_1, \frac{k_n}{nb_n} t_2)|}{t_1^{\nu_1} t_2^{\nu_2}} \to_d \sup_{0 < t_1 \leq 1} \frac{|W(t_1, t_2)|}{t_1^{\nu_1} t_2^{\nu_2}} <_{a.s.} \infty ,
$$

where $W$ is a standard Wiener process on $[0, \infty)^2$.

**COROLLARY 3.8** (cf. Lemma 1 (p. 30) and Remark 5 (p. 43) in Huang Xin (1992)). Let $F$ be a distribution function on $\mathbb{R}^d$, $d \in N$, with uniform-(0,1) marginals and assume $\lim_{u \to 0} F(ut_1, \cdots, ut_d)/u =: R(t_1, \cdots, t_d) > 0$, for all $(t_1, \cdots, t_d) \in [0, \infty)^d$. Now let $X_1, \cdots, X_n$ be i.i.d. random vectors with distribution function $F$. Then for all $c_1, \cdots, c_d > 0$

$$
\sup_{t \in \Pi_{j=1}^{d} \in (0, c_j]} \frac{n}{k_n^{1/2}} \frac{n}{t_j^{\nu_j}} \frac{|F_n(k_n t/n) - F(k_n t/n)|}{\prod_{j=1}^{d} t_j^{\nu_j}} \to_d \sup_{t \in \Pi_{j=1}^{d} \in (0, c_j]} \frac{|W(t)|}{\prod_{j=1}^{d} t_j^{\nu_j}} <_{a.s.} \infty ,
$$

where $W$ is a standard Wiener process on $[0, \infty)^2$. 

11
where $F_n$ is the empirical distribution function of the $X_i$ and $W$ is a Wiener process with as 'time' the measure induced by $R$.

The next corollary, just like the previous one, comes from multivariate extreme value theory. For its presentation we have to specify our setup and introduce more notation. For more background, see Resnick (1987, Chapter 5) or de Haan and Resnick (1993); cf. also special case (III), section 1. Let $X_1, \ldots, X_n$ be i.i.d. bivariate random vectors with fixed common distribution $F$, which is in the domain of max-attraction of a bivariate extreme value distribution $G$. Denote the normalizing vectors with $(a_1(n), a_2(n)), a_1(n), a_2(n) > 0$, and $(b_1(n), b_2(n))$, i.e. we have (cf. Corollary 2.6 and its proof)

$$F^n(a_1(n)t_1 + b_1(n), a_2(n)t_2 + b_2(n)) \rightarrow G(t_1, t_2) \quad (n \rightarrow \infty),$$

(3.18)

for all $(t_1, t_2) \in \mathbb{R}^2$. Note that (3.18) can be written as

$$\lim_{u \rightarrow -\infty} u(1 - F(a_1(u)t_1 + b_1(u), a_2(u)t_2 + b_2(u))) = -\log G(t_1, t_2),$$

for all $t_1, t_2$ with $G(t_1, t_2) > 0$, where $u$ now ranges through $\mathbb{R}$, instead of $\mathbf{N}$. Define our $X_{n,i} = (X_{n,i,1}, X_{n,i,2}), i = 1, \ldots, n$, by

$$X_{n,i,j} = \left(1 + \gamma_j \left(\frac{X_{i,j} - b_j(\frac{t_j}{\gamma_j})}{a_j(\frac{t_j}{\gamma_j})}\right)\right)^{1/\gamma_j}, \quad j = 1, 2,$$

where the $\gamma_1, \gamma_2 \in \mathbb{R}$ are the so-called extreme value indices (of the marginals). (For $1 + ba \leq 0$ redefine $(1 + ba)^{1/b}$ as 0, interpret $(1 + ba)^{1/b}$ as $e^a$ for $b = 0$. Note also that we have chosen the $a_j$ and $b_j$ such that the marginals of $G$ are of the form $\exp(-(1 + \gamma_j(u)^{-1/\gamma_j}))$.) Let $\theta \in [0, \pi/2]$ and write $C_\theta = \{(t_1, t_2) \in [0, \infty)^2 : t_1 \vee t_2 \geq 1, t_1 / t_2 \leq \tan \theta\}$. For $\alpha \in (0, \infty)^2, \beta, \eta \in \mathbb{R}^2$, write

$$D_\theta = D_\theta(\alpha, \beta, \eta) = \left(1, 1 + \gamma \left(\frac{C_\theta^n - (1, 1)}{\eta} + \beta\right)\right)^{1/\gamma},$$

with $\gamma = (\gamma_1, \gamma_2)$, where all the vector operations are means coordinatewise e.g., $C_\theta^n = \{(t_1^n, t_2^n) : (t_1, t_2) \in C_\theta\}$. (For $\eta = 0, (C_\theta^n - 1)/\eta$ has to be interpreted as $\log C_\theta = \{(\log t_1, \log t_2) : (t_1, t_2) \in C_\theta\}$.) Write $E = \{(t_1, t_2) \in [0, \infty)^2 : t_1 \vee t_2 \geq \frac{1}{2}\}$. Finally set $A = \{D_\theta : D_\theta \subset E, \theta \in [0, \pi/2], \alpha \in (0, \infty)^2, \beta, \eta \in \mathbb{R}^2 \} \cup E$, and $A' = \{A \cap A' : A, A' \in A\}$.

**COROLLARY 3.9** (cf. Proposition 4 in Einmahl, de Haan and Sinha (1995)). Assume that, as $n \rightarrow \infty$,

$$\sup_{A \in A'} \left|\frac{n}{k_n} P_{k_n}(A) - \mu(A)\right| \rightarrow 0,$$

where $\mu$ is a so-called exponent measure, defined by

$$\mu([0, t_1] \times [0, t_2]^{c}) = -\log G\left(\frac{t_1^n - 1}{\gamma_1}, \frac{t_2^n - 1}{\gamma_2}\right), \quad t_1, t_2 > 0.$$
Then
\[
\sup_{A \in \mathcal{A}} \left| \frac{n}{k_n^{1/2}} (P_n(A) - P_{(n)}(A)) - W_n(A) \right| \rightarrow_{\mathcal{P}} 0, \tag{3.19}
\]
with \(W_n, P_n\) and \(P_{(n)}\) as in Theorem 3.1.

The next corollary is not really a corollary to Theorem 3.1 but more a modification; the final corollary is a special case of it. Let \(\tilde{q}\) be a function on \([0, 1]\), with \(\tilde{q} > 0\) on \((0, 1]\).

Assume
\[(A.13) \quad I/\tilde{q} \text{ (I the identity function; } 0/0 := 0) \text{ is continuous on } [0, 1].\]

Let \(\mathcal{A}, \text{ etc.}, \) be as before but assume that
\[(A.14) \quad \text{there exists an } \overline{A} \in \mathcal{A} \text{ such that } A \subseteq \overline{A}, \text{ for all } A \in \mathcal{A}.\]

Set \(k_n = nP_{(n)}(\overline{A})\) and assume as before that \(k_n \to \infty\) and \(k_n/n \to 0.\)

**COROLLARY 3.10.** Assume that we have (A.2), (A.13), (A.14) and
\[(A.15) \quad \sup_{A \in \mathcal{A}} |\mu_n(A) - \mu(A)| \to 0 \quad (n \to \infty), \text{ and}
\[
m^{1/2}(\mu_{n,m} - \mu_n)/(\tilde{q} \circ \mu) \to_d B/(\tilde{q} \circ \mu) \quad (m \wedge n \to \infty), \text{ on } \mathcal{A},
\]
where \(\mu_{n,m}\) is the empirical measure based on a random sample of size \(m\) from \(\mu_n\) and \(B\) is as in (3.6), but with \(\nu\) replaced by \(\mu\). Then, as \(n \to \infty,
\[
\sup_{A \in \mathcal{A}} \left| \frac{v_n(A)}{\tilde{q}(\mu_n(A))} - \frac{W_n(A)}{\tilde{q}(\mu(A))} \right| \to_{\mathcal{P}} 0 \quad \text{.} \tag{3.20}
\]

**COROLLARY 3.11.** Let \(X_1, \ldots, X_n\) be i.i.d. positive random variables with common distribution function \(F\). Let \(\{a_n\}_{n=1}^\infty\) be a sequence of positive numbers such that \(F(a_n) \to 0\) and \(k_n := nF(a_n) \to \infty \quad (n \to \infty).\) Set \(G_n(t) = F(ta_n)/F(a_n), \ 0 \leq t \leq 1,\) and assume
\[
\sup_{0 \leq t \leq 1} |G_n(t) - G(t)| \to 0, \tag{3.21}
\]
for some distribution function \(G\) on \([0, 1]\). Then if \(\tilde{q}\) is continuous, non-decreasing and satisfies (3.12) we have, as \(n \to \infty,
\[
\sup_{0 < t \leq 1} \left| \frac{n}{k_n^{1/2}} (F_n(ta_n) - F(ta_n))}{\tilde{q}(G_n(t))} - \frac{W_n(G(t))}{\tilde{q}(G(t))} \right| \to_{\mathcal{P}} 0 \quad \text{,} \tag{3.22}
\]
where \(F_n\) is the empirical distribution function of the \(X_i\) and \(W_n\) is a standard Wiener process on \([0, 1].\) In particular, this yields again (3.13) and hence (1.6).
PROOF OF COROLLARY 3.1. If (A.11) holds, Corollary 3.1 follows by combining Theorem 3.1 and Proposition 3.1. If (A.12) holds, we also have to show that
\[
\sup_{A \in \mathcal{A}(1)} \frac{|W(A)|}{q(A)} < \infty \text{ a.s.,}
\]
but this follows from the third statement in (A.12); see, e.g., Lemma 3.1 in Alexander (1987).

PROOF OF COROLLARY 3.2. Corollary 3.2 is a special case of Theorem 3.1 and Corollary 3.1.

PROOF OF COROLLARY 3.3. The proof follows readily along the lines of the proof of Corollary 2.3.

PROOF OF COROLLARY 3.4. Immediate; condition (A.12) is satisfied, mainly because of the Chibisov-O'Reilly theorem.

PROOF OF COROLLARY 3.5. The proof is very much the same as that of Corollary 3.4.

PROOF OF COROLLARY 3.6. Immediate; the third statement in (A.12) is the weak convergence of the weighted multivariate uniform empirical process indexed by points, see, e.g., Theorem 3.1 in Einmahl (1987) or Corollary 3.9 in Alexander (1987).

PROOF OF COROLLARY 3.7. For the proof of Corollary 3.7 some work has to be done. Actually Corollary 3.7 is not a direct corollary to Theorem 3.1 and Proposition 3.1, since neither (A.11) nor (A.12) is satisfied. Note that the first statement of (A.12) is not fulfilled, since \( E(u) = \infty \) is unbounded. On the other hand, since \( \mu(E(u)) = 0 \) (\( n \to \infty \)), it is possible to adapt the proof of Theorem 3.1 such that it covers the situation of the present corollary. For the sake of brevity, we will omit this, we will suffice with checking the conditions of Theorem 3.1, apart from the just discussed first statement in (A.12), and proving the almost sure finiteness of the limiting random variable in (3.16).

We will assume \( \nu_1 > 0, \nu_2 \leq 1 \) and \( c = 1 \). The other cases can be easily deduced from this. Checking most of the conditions is trivial. Therefore we will only consider (A.5), the rest of (A.12) and the finiteness of the limiting random variable. Just like in the proof of Corollary 2.5 we have, for \( u > 1, \lambda(u) = u^{1/\nu_2}/(1 - \nu_1/\nu_2), \) hence (A.5). Now we show the second statement of (A.12). Set \( q(A) = t_1^{\nu_1} t_2^{\nu_2} \mapsto \bar{u} \leq u \). Then \( \mu(A)/q(A) = t_1 t_2/(t_1^{\nu_1} t_2^{\nu_2}) = t_1^{1-\nu_1/\nu_2} \bar{u}^{1/\nu_2-1} \leq u^{1/\nu_2-1}, \) so \( \sup_{A \in \mathcal{A}(1)} \mu(A)/q(A) \leq u^{1/\nu_2-1} < \infty \). For the third statement in (A.12), we apply Theorem 3.4 in Alexander (1987). We need to show \( \nu(E(\epsilon)) = o(\epsilon^2), \) as \( \epsilon \to 0, \) and \( B(A)/q(A) \to 0 \) a.s. as \( \nu(A) \to 0. \) For \( \epsilon < 1 \) we have \( \lambda(\epsilon/(1 - \nu_1/\nu_2)) = (1 - \nu_1/\nu_2)\epsilon^{1/\nu_1}, \) hence \( \nu(E(\epsilon)) = \lambda(\epsilon) = \lambda(\epsilon)/\lambda(u) = \epsilon^{1/\nu_1}/u^{1/\nu_2} = \epsilon(\epsilon^2) = (\epsilon^2) = (\epsilon - 0), \) since \( \nu_1 < 1/2. \) An easy way to see that \( B(A)/q(A) \to 0 \) a.s. as \( \nu(A) \to 0 \) and also the almost sure finiteness of the limiting random variable in (3.15), is to note that if \( W(t_1, t_2) \) is a standard Wiener process on \([0, \infty)^2\), then \( t_2 W(t_1, 1/t_2) \) is also a standard Wiener process on \([0, \infty) \times [0, 1] \) (and relating \( B \) to \( W \)). This transforms \((0, 1] \times [1, \infty) \) to \([0, 1]^2\), where we can apply well-known results (note that also the weights transform). We omit details.

PROOF OF COROLLARY 3.8. This is an easy corollary, since \( q \equiv 1. \) So since we only have to prove (4.34) below with \( u = 1 \) and \( q = 1 \), there is no need to check (A.4) and
(A.9). Obviously \( q(A) > \frac{1}{2} \) is also sufficient instead of \( q(A) > 1 \) in (A.11). The second part of (A.11) follows since \( R \) is continuous and increasing in each coordinate. □

PROOF OF COROLLARY 3.9. Similar remarks as in the proof of the previous corollary apply, but (A.1) and (A.2), especially (A.1), are not immediate here. Checking (A.1) and (A.2) is rather cumbersome, but it is carried out, for a slightly easier situation, in the appendix of Einmahl, de Haan and Sinha (1995). Therefore it is omitted here. □

PROOF OF COROLLARY 3.10. The proof is very much the same as a combination of the proof of (4.34) under (A.11) and (A.12) respectively (see (4.38) and (4.42)); note that \( I/\bar{q} \) is bounded and uniformly continuous. □

PROOF OF COROLLARY 3.11. We only consider the second statement in (A.15), the others are readily checked. But this follows rather easily from the Chibisov-O'Reilly theorem, the fact that, in the notation of (A.15), \( m^{1/2}(G_{n,m} - G_n) = m^{1/2}(\Gamma_m \circ G_n - G_n) \), where \( \Gamma_m \) is a uniform-(0,1) empirical distribution function, and the uniform continuity of the sample paths of \( B/\bar{q} \) in combination with (3.21). □

DISCUSSION OF THE RESULTS. First note that most of the discussion of section 2 holds also true for the results of this section. Condition (A.5), the main condition for the finiteness of the limiting random variables, see Propositions 2.1 and 3.1, is, for the uniform-(0,1) empirical process setup, sharp for Theorems 2.1 and 3.1 and ‘almost’ sharp for Corollaries 2.1 and 3.1, see Csörgő and Horváth (1992). Conditions (A.11) and (A.12), required to prove Theorem 3.1, are of different nature. Condition (A.11) is the main condition to prove weak convergence of \( m^{1/2}(\nu_{n,m} - \nu_m) \) (see (4.36)), whereas in (A.12) the weak convergence itself of the weighted empirical process is assumed. See Alexander (1987) for conditions for the latter weak convergence. Also, the second and third statement in (A.12) are very much related, see (3.2) in Theorem 3.4 of Alexander (1987). Corollaries 3.8 and 3.9 are extremely useful in the estimation of multivariate extreme value distributions, especially in estimating the dependence structure. A corollary, that is a kind of mixture of Corollaries 3.8 and 3.9, is also possible but is omitted for the sake of brevity. Such a corollary would play a similar role for the estimator in de Haan and Resnick (1993) as Corollaries 3.8 and 3.9 do for the estimators in the papers cited in these corollaries. Finally, for a paper related to the setup of Corollary 3.11, but in the more complicated random censorship model, the reader is referred to Einmahl and Ruymgaart (1995).

4 Proofs of the main results

PROOF OF PROPOSITION 2.1. It is not hard to see that (2.4) and (2.5) hold true if \( \lim_{t \to \infty} \lambda(t) < \infty \). Therefore in the proof of (2.4) and (2.5) we assume that \( \lim_{t \to \infty} \lambda(t) = \infty \). First consider (2.4). For any \( k \in \mathbb{N} \), set \( \mathcal{A}_k = \{ A \in \mathcal{A} : h(2^{k-1}) \leq q(A) < h(2^k) \} \). Obviously

\[
\sup\{|N(A) - \mu(A)|/q(A) : A \in \mathcal{A}_k\} \leq \sup\{|N(A) - \mu(A)|/h(2^{k-1}) : A \in \mathcal{A}_k\} \vee 0 =: N_k.
\]
Therefore to prove (2.4) it suffices to show that

$$\lim_{k \to \infty} N_k = 0 \text{ a.s.} \tag{4.1}$$

Let $X_1, X_2, \ldots$ be i.i.d. random elements taking values in $E(h(2^k))$ with distribution

$$\mu(B \cap E(h(2^k)))/2^k, \quad B \in \mathcal{S}.$$ 

Further, let $N(2^k)$ be a Poisson($2^k$) random variable, independent of the $X_i$. As in Gaenssler (1983, p. 7) we have

$$\{N(A) : A \in \mathcal{A}_k\} \overset{d}{=} \left\{ \sum_{i=1}^{N(2^k)} 1_A(X_i) : A \in \mathcal{A}_k \right\},$$

where the empty sum is defined to be zero. Hence for any $\varepsilon > 0$ and $k \in \mathbb{N}$

$$\mathbb{P}(N_k > 2\varepsilon) = \mathbb{P}(\sup\{|S(N(2^k), A) - \mu(A)|/h(2^{k-1}) : A \in \mathcal{A}_k\} > 2\varepsilon)$$

$$=: p_k(\varepsilon), \tag{4.2}$$

where for $m \in \mathbb{N}$

$$S(m, A) = \sum_{i=1}^{m} 1_A(X_i) \text{ and } S(0, A) = 0, A \in \mathcal{A}. \tag{4.3}$$

Write

$$p(A) = \mathbb{P}(X_1 \in A), A \in \mathcal{A},$$

and set

$$T(m, A) = S(m, A) - mp(A), A \in \mathcal{A}.$$ 

We have

$$p_k(\varepsilon) \leq \mathbb{P}(\sup\{|T(N(2^k), A)|/h^{(k-1)} : A \in \mathcal{A}_k\} > \varepsilon)$$

$$+ \mathbb{P}(\sup\{|N(2^k)p(A) - \mu(A)|/h^{(2k-1)} : A \in \mathcal{A}_k\} > \varepsilon)$$

$$=: p_{1, k}(\varepsilon) + p_{2, k}(\varepsilon). \tag{4.4}$$

Also
$$\sum_{|m-2^k| \leq 2^{k+1} \log k} \mathbb{P}(\sup |T(m, A)|/h(2^{k-1}) : A \in A_k > \varepsilon) \cdot \mathbb{P}(N(2^k) = m)$$

$$+ \mathbb{P}(|N(2^k) - 2^k| > 2^{k+1} \log k)^{1/2} =: p_{3,k}(\varepsilon) + p_{4,k} .$$

(4.5)

Observe that we have from (A.5), that for all $c > 0$

$$\int_1^{\infty} \frac{1}{u} \exp \left( \frac{-ch^2(u)}{u} \right) du = \sum_{k=1}^{\infty} \int_{2^{k-1}}^{2^k} \frac{1}{u} \exp \left( \frac{-ch^2(u)}{u} \right) du \geq \sum_{k=1}^{\infty} \frac{2^k - 2^{k-1}}{2^k} \exp \left( \frac{-ch^2(2^k)}{2^{2k-1}} \right) = \sum_{k=1}^{\infty} \frac{1}{2} \exp \left( \frac{-2ch^2(2^k)}{2^k} \right) .$$

Hence

$$\sum_{k=1}^{\infty} \exp \left( \frac{-ch^2(2^k)}{2^k} \right) < \infty \text{ for all } c > 0 .$$

(4.6)

To proceed we need the following two inequalities. The first one follows from Theorem 2.11 in Alexander (1984).

**FACT 4.1.** For all $m \in \mathcal{B}$ and $x \geq c$, where $c$ depends only on the index $v$ of the VC class $A$,

$$\mathbb{P}(\sup \{|T(m, A)| : A \in A\} > x) \leq \exp(-x^2/m) .$$

The second inequality is a probability bound for the tails of a centered Poisson($\tau$) random variable $Y$. Introduce the decreasing, continuous function $\psi$ defined on $[0, \infty)$ by

$$\psi(x) = 2x^{-2}((1 + x) \log(1 + x) - x), x > 0; \psi(0) = 1 .$$

**FACT 4.2.** For all $x \geq 0$

$$\mathbb{P}(|Y - \tau| \geq x) \leq 2 \exp \left( \frac{-x^2}{2\tau} \psi \left( \frac{x}{\tau} \right) \right) .$$

This follows from, e.g., (2.16) in Einmahl (1987).

By Fact 4.1 and an elementary argument we have for large $k$

$$p_{3,k}(\varepsilon) \leq \exp \left( \frac{-\varepsilon^2 h^2(2^{k-1})}{2^{2k+1}} \right),$$

(4.7)
which is summable in \( k \) by (4.6). We also have by Fact 4.2 for large \( k \)
\[
p_{4,k} \leq 2 \exp \left( -\frac{2^{k+2} \log k}{2^{k+1}} \psi \left( \frac{2^{k+1} \log k}{2^{k}} \right) \right) \leq \frac{2}{k^{3/2}},
\]
(4.8)

implying that \( p_{4,k} \) is also summable in \( k \). We also use Fact 4.2 to deal with \( p_{2,k}(\varepsilon) \). Note that for \( A \in \mathcal{A}_k, 2^k p(A) = \mu(A) \), hence
\[
p_{2,k}(\varepsilon) \leq P \left( \left| N(2^k) - 2^k \right| > \varepsilon h(2^{k-1}) \right)
\leq 2 \exp \left( -\frac{\varepsilon^2 h^2(2^{k-1})}{2^{k+1}} \psi \left( \frac{\varepsilon h(2^{k-1})}{2^k} \right) \right).
\]
(4.9)

This yields, using that \( I \psi \) (\( I \) the identity function) is increasing and applying (4.6)
\[
\sum_{k=1}^{\infty} p_{2,k}(\varepsilon)
\leq 2 \sum_{h(2^{k-1}) \leq 2^k} \exp \left( -\frac{\varepsilon^2 h^2(2^{k-1})}{2^{k+1}} \psi \left( \frac{\varepsilon h(2^{k-1})}{2^k} \right) \right)
\]
\[+ 2 \sum_{h(2^{k-1}) > 2^k} \exp \left( -\frac{\varepsilon^2 h^2(2^{k-1})}{2^{k+1}} \psi \left( \frac{\varepsilon h(2^{k-1})}{2^k} \right) \right)
\]
\[
\leq 2 \sum_{k=1}^{\infty} \exp \left( -\frac{1}{4} \varepsilon^2 \psi^2(h(2^{k-1})) \right)
\]
\[+ 2 \sum_{h(2^{k-1}) > 2^k} \exp \left( -\frac{1}{4} \varepsilon^2 \psi(\varepsilon) h(2^{k-1}) \right)
\]
\[
\leq 2 \sum_{k=1}^{\infty} \exp \left( -\frac{1}{4} \varepsilon^2 \psi(\varepsilon) h(2^{k-1}) \right)
\]
\[+ 2 \sum_{k=1}^{\infty} \exp(-\frac{1}{4} \varepsilon^2 \psi(\varepsilon) 2^k) < \infty.
\]
(4.10)

Now (4.7), (4.8) and (4.10) when combined with (4.5), (4.4) and (4.2), along with the Borel-Cantelli lemma give (4.1) by the arbitrary choice of \( \varepsilon > 0 \). This proves (2.4).

It is easy to see that
\[
\sup\{|N(A) - \mu(A)|/q(A) : A \in \mathcal{A}, 1 < q(A) \leq u\}
\leq \sup\{|N(A) - \mu(A)| : A \in \mathcal{A}, A \subset E(u)\}
\leq N(E(u)) \vee \lambda(u) < \infty \text{ a.s.,}
\]
which with (2.4) implies (2.5).

Towards a proof of (2.6) we observe that we have by (A.7)

\[
M_q(1) := \sup\{ |N(A) - \mu(A)|/(q(A) : A \in \mathcal{A}, q(A) \leq 1} \leq \sup \left\{ \frac{N(A)}{q(A)} : A \in \mathcal{A}, q(A) \leq 1 \right\} \vee \frac{1}{b}.
\]

This last bound is in turn on the event \( \{N(E(u)) = 0\}, 0 < u \leq 1, \) bounded from above by

\[
\frac{N(E(1))}{u} \vee \frac{1}{b}.
\]

This event has probability \( \exp(-\lambda(u)) \), which converges to 1 as \( u \downarrow 0 \) by (A.6). Therefore \( M_q(1) < \infty \) a.s., which when combined with (2.5) yields (2.6).

For the proof of Theorem 2.1 we need a result, comparable to (2.4), for the local empirical process. Towards this aim set for \( u > 0 \)

\[
M_{n,q}(u) = \sup\{ |n P_n(A) - \mu_n(A)|/(q(A) : A \in \mathcal{A}, q(A) > u}.
\]

**PROPOSITION 4.1.** Assume (A.1)-(A.5) and (A.9) are satisfied, then for any \( \varepsilon > 0 \) there exists a \( u > 0 \) such that

\[
\limsup_{n \to \infty} P(M_{n,q}(u) > \varepsilon) \leq \varepsilon.
\]

**PROOF.** We follow the lines of the proof of (2.4). As in that proof we assume w.l.o.g. that \( \lim_{t \to \infty} \lambda(t) = \infty \), also recall the definition of \( A_k \) there. Everywhere in this proof we tacitly assume that \( n \) is so large that (A.9) holds. Obviously

\[
\sup\{ |n P_n(A) - \mu_n(A)|/(q(A) : A \in A_k} \leq \sup\{ |n P_n(A) - \mu_n(A)|/h(2^{k-1}) : A \in A_k} =: N_{n,k}.
\]

It is clear that for every \( k_0 \in \mathbb{N} \) there exists a \( u > 0 \) such that

\[
P(M_{n,q}(u) > \varepsilon) \leq P \left( \sup_{k \geq k_0} N_{n,k} > \varepsilon \right) \leq \sum_{k=k_0}^{\infty} P(N_{n,k} > \varepsilon).
\]

Let \( Y_1, Y_2, \ldots \), be i.i.d. random elements taking values in \( E(h(2^k)) \) with distribution

\[
P_n(B \cap E(h(2^k)))/P_n(E(h(2^k)) = P_n(B), \ B \in \mathcal{S}.
\]

We also need a binomial random variable with parameters \( n \) and \( P_n(E(h(2^k))) \), independent of the \( Y_i \), which we will denote by \( N_n(2^k) \). Now we have that
\{nP_n(A), A \in \mathcal{A}_k\} \overset{d}{=} \left\{ \sum_{i=1}^{N_n(2^k)} 1_{A_i}, A \in \mathcal{A}_k \right\},

where the empty sum is defined to be zero. Hence for any \(\varepsilon > 0\) and \(k \in \mathbb{N}\)

\[
IP(N_{n,k} > \varepsilon) 
\leq IP(\sup\{|S_n(N_n(2^k), A) - \mu_n(A)|/h(2^{k-1}) : A \in \mathcal{A}_k\} > \varepsilon)
= p^{(n)}_k(\varepsilon),
\]

(4.14)

where \(S_n(m, A)\) is defined similarly as \(S(m, A)\) in (4.3). Also define \(p_n(A)\) and \(T_n(m, A)\) similarly as in the proof of Proposition 2.1.

We have, similar to (4.4),

\[
p^{(n)}_k(\varepsilon) \leq IP(\sup\{|T_n(N_n(2^k), A)|/h(2^{k-1}) : A \in \mathcal{A}_k\} > \frac{1}{2} \varepsilon)
+ IP(\sup\{|N_n(2^k)p_n(A) - \mu_n(A)|/h(2^{k-1}) : A \in \mathcal{A}_k\} > \frac{1}{2} \varepsilon)
= p^{(n)}_{1,k}(\varepsilon) + p^{(n)}_{2,k}(\varepsilon),
\]

(4.15)

and, as in (4.5)

\[
p^{(n)}_{1,k}(\varepsilon)
\leq \sum_{m-\mu_n(E(h(2^k))) \leq 2^{\frac{k}{2}+1}(C \log k)^{1/2}} IP(\sup\{|T_n(m, A)|/h(2^{k-1}) : A \in \mathcal{A}_k\} > \frac{1}{2} \varepsilon)
\cdot IP(N_n(2^k) = m)
+ IP(|N_n(2^k) - \mu_n(E(h(2^k)))| > 2^{\frac{k}{2}+1}(C \log k)^{1/2})
= p^{(n)}_{3,k}(\varepsilon) + p^{(n)}_{4,k}.
\]

(4.16)

Since \(\mu_n(E(h(2^k))) \leq C\mu(E(h(2^k))) = C2^k\) and hence for large \(k\), \(\mu_n(E(h(2^k))) + 2^{\frac{k}{2}+1}(C \log k)^{1/2} \leq C2^{k+1}\), we obtain from Fact 4.1 that for large \(k\)

\[
p^{(n)}_{3,k}(\varepsilon) \leq \exp\left(-\frac{\varepsilon h^2(2^{k-1})}{C2^{k+3}}\right).
\]

(4.17)

Now it follows from (A.6) and (4.6) that for some large \(k_0^{(1)}\)

\[
\sum_{k=k_0^{(1)}}^{\infty} p^{(n)}_{3,k}(\varepsilon) \leq \frac{1}{3} \varepsilon.
\]

(4.18)
We need a probability inequality for the tails of a binomial random variable, cf. Fact 4.2, which readily follows from Bennett's (1962) inequality using the fact that $I\psi$ is increasing.

**FACT 4.3.** Let $B(n,p)$ denote a binomial random variable with parameters $n$ and $p$. Then for all $x \geq 0$

$$BP(|B(n,p) - np| \geq x) \leq 2 \exp \left(\frac{-x^2}{2np} \psi \left(\frac{x}{np}\right)\right).$$

We use Fact 4.3 to deal with $p_{4,k}^{(n)}$ and $p_{2,k}^{(n)}(\varepsilon)$. We obtain

$$p_{4,k}^{(n)} \leq 2 \exp \left(\frac{-2^{k+2}C \log k}{2\mu_n(E(h(2^k)))} \psi \left(\frac{2^{\frac{k}{2}+1}(C \log k)^{1/2}}{\mu_n(E(h(2^k)))}\right)\right). \tag{4.19}$$

Using again that $I\psi$ is increasing and that $\lim_{x \to 0} \psi(x) = 1$, we obtain from (4.19) that for large $k$

$$p_{4,k}^{(n)} \leq 2 \exp \left(\frac{-2^{k+2}C \log k}{C2^{k+1}} \psi \left(\frac{2^{\frac{k}{2}+1}(C \log k)^{1/2}}{C2^k}\right)\right) \leq \frac{2}{k^{3/2}}. \tag{4.20}$$

Hence for some large $k_0^{(2)}$

$$\sum_{k=k_0^{(2)}}^{\infty} p_{4,k}^{(n)} \leq \frac{1}{3} \varepsilon. \tag{4.21}$$

Finally consider $p_{2,k}^{(n)}(\varepsilon)$. For $A \in \mathcal{A}_k$ we see that

$$\mu_n(E(h(2^k)))\mu_n(A) = \mu_n(A).$$

So

$$p_{2,k}^{(n)}(\varepsilon) \leq BP(|N_n(2^k) - \mu_n(E(h(2^k)))| > \frac{1}{2}h(2^{k-1}))$$

$$\leq 2 \exp \left(\frac{-\varepsilon^2 h^2(2^{k-1})}{C2^{k+3}} \psi \left(\frac{\varepsilon h(2^{k-1})}{C2^{k+1}}\right)\right).$$

Now using a similar argument as in (4.10) (cf. (4.9)) we have that for some large $k_0^{(3)}$

$$\sum_{k=k_0^{(3)}}^{\infty} p_{2,k}^{(n)}(\varepsilon) \leq \frac{1}{3} \varepsilon. \tag{4.22}$$

Hence, with $k_0 = \max(k_0^{(1)}, k_0^{(2)}, k_0^{(3)})$, we obtain from (4.17), (4.21) and (4.22), when combined with (4.14)–(4.16) that
\[
\sum_{k=k_0}^{\infty} \mathbb{P}(N_{n,k} > \varepsilon) \leq \varepsilon . \tag{4.23}
\]

Now (4.13) completes the proof. \(\square\)

**PROOF OF THEOREM 2.1.** We have to show that for all \(\varepsilon > 0\), there exists an \(n_0\) such that for \(n \geq n_0\)

\[
\mathbb{P} \left( \sup_{A \in \mathcal{A}} \frac{|(n\mu_n(A) - \mu_n(A)) - (N_n(A) - \mu(A))|/q(A) > 3\varepsilon}{3\varepsilon} \right) \leq 3\varepsilon .
\]

From (2.4) in Proposition 2.1 we have the existence of a \(u > 1\) such that for all \(n \in \mathbb{N}\)

\[
\mathbb{P} \left( \sup_{A \in \mathcal{A}, q(A) > u} \frac{|N_n(A) - \mu(A)|/q(A) > \varepsilon}{\varepsilon} \right) \leq \varepsilon . \tag{4.24}
\]

From the just proved Proposition 4.1 we have (possibly enlarging \(u\))

\[
\limsup_{n \to \infty} \mathbb{P} \left( \sup_{A \in \mathcal{A}, q(A) > u} \frac{|n\mu_n(A) - \mu_n(A)|/q(A) > \varepsilon}{\varepsilon} \right) \leq \varepsilon . \tag{4.25}
\]

From (4.24) and (4.25) it remains to show that for large \(n\)

\[
\mathbb{P} \left( \sup_{A \in \mathcal{A}, q(A) \leq u} \frac{|(n\mu_n(A) - \mu_n(A)) - (N_n(A) - \mu(A))|/q(A) > \varepsilon}{\varepsilon} \right) \leq \varepsilon . \tag{4.26}
\]

Because of (A.11) it is now sufficient to show for large \(n\)

\[
\mathbb{P} \left( \sup_{A \in \mathcal{A}, q(A) \leq u} \frac{|n\mu_n(A) - N_n(A)|/q(A) > 0}{\varepsilon} \right) \leq \varepsilon . \tag{4.27}
\]

or, more simple,

\[
\mathbb{P} \left( \sup_{A \in \mathcal{A}, q(A) \leq u} \frac{|n\mu_n(A) - N_n(A)| > 0}{\varepsilon} \right) \leq \varepsilon . \tag{4.28}
\]

This, however, follows readily from the literature. Let \(d_n\) be the total variation distance, on \(E(u)\), between the laws of \(nP_n\) and \(N_n\), respectively. Then according to, e.g., Dobrushin (1970, p.472) we can construct versions of \(nP_n\) and \(N_n\) such that
\[ \mathbb{P}(nP_n \neq N_n, \text{on } E(u)) = d_n. \quad (4.29) \]

But from (A.8) and Theorem 3.2.3 in Reiss (1993) we have \( d_n \to 0 \), as \( n \to \infty \). This essentially completes the proof.

There are two remarks needed to make the proof mathematically rigorous. First, the processes \( N_n \) depend, unwanted, on \( \varepsilon \). This, however, can be overcome by a routine diagonal selection argument. Second, to get the processes of (4.24), (4.25), and (4.29) on one probability space, (the one of the \( X_{n,i} \)) we make repeated use of Lemma A.1 in Berkes and Philipp (1979), cf. also Csörgő and Révész (1981, p. 140) and Csörgő (1983, p. 21-23).

PROOF OF PROPOSITION 3.1. The proof of this proposition is similar to, but easier than the proof of Proposition 3.1. Therefore we only sketch it, with emphasis on the differences. We again assume \( \lim_{t \to \infty} \lambda(t) = \infty \). First consider (3.3). Recall the definition of \( \mathcal{A}_k \) and define

\[ \tilde{N}_k = \sup\{|W(A)|/h(2^{k-1}) : A \in \mathcal{A}_k \} \vee 0. \quad (4.30) \]

To prove (3.3) it suffices to show that

\[ \lim_{k \to \infty} \tilde{N}_k = 0 \text{ a.s.} \quad (4.31) \]

For (4.31) we need the following inequality.

**Lemma 4.1.** For all \( x \geq \varepsilon \), where \( \varepsilon \) depends only on the index \( v \) of the VC class \( \mathcal{A} \),

\[ \mathbb{P}(\sup\{|W(A)| : A \in \mathcal{A}_k \} > x) \leq 2 \exp(-x^2/2^{k+3}). \quad (4.32) \]

**Proof.** Write \( \nu = \mu/2^k \) and let \( B \) be as in (3.6). Now it follows from Fact 4.1 and a weak convergence argument that

\[ \mathbb{P}(\sup\{|B(A)| : A \in \mathcal{A}_k \} > x) \leq \exp(-x^2). \quad (4.33) \]

But \( W \overset{d}{=} 2^{k/2}(B + Z\nu) \), on \( \mathcal{A}_k \), where \( Z \) is a standard normal random variable, independent of \( B \). Hence by (4.33)

\[ \mathbb{P}(\sup\{|W(A)| : A \in \mathcal{A}_k \} > x) \]

\[ \leq (\sup\{|B(A)| : A \in \mathcal{A}_k \} > x/2^{k+1}) + \mathbb{P}(|Z| > x/2^{k+1}) \]

\[ \leq \exp(-x^2/2^{k+2}) + \exp(-x^2/2^{k+3}) \leq 2 \exp(-x^2/2^{k+3}), \]

where for the next to last inequality Mill's ratio is used.

Now we have by Lemma 4.1 for arbitrary \( \varepsilon > 0 \), for large \( k \),

23
Hence (4.6) in combination with the Borel-Cantelli lemma completes the proof of (3.3).

Obviously, since \( \mu(A) \leq \mu(E(u)) = \lambda(u) < \infty \) for \( A \in \mathcal{A} \) with \( q(A) \leq u \)

\[
\sup\{ |W(A)|/q(A) : A \in \mathcal{A}, 1 < q(A) \leq u \} 
\leq \sup\{ |W(A)| : A \in \mathcal{A}(u) \} < \infty \text{ a.s.,}
\]

which with (3.3) implies (3.4).

For the proof of Theorem 3.1 we need an analogue of Proposition 4.1. Its proof is very much the same as that of Proposition 4.1 and hence will be omitted. Set for \( u \geq 0 \)

\[
\mathcal{M}_{n,q}(u) = \sup\{ |v_n(A)|/q(A) : A \in \mathcal{A}, q(A) > u \}.
\]

**PROPOSITION 4.2.** Assume (A.1)-(A.5) and (A.9) are satisfied, then for any \( \varepsilon > 0 \) there exists a \( u > 0 \) such that

\[
\limsup_{n \to \infty} P(\mathcal{M}_{n,q}(u) > \varepsilon) \leq \varepsilon.
\]

**PROOF OF THEOREM 3.1.** The first part of this proof resembles very much that of Theorem 2.1 with the applications of Proposition 2.1 and 4.1 replaced by applications of Propositions 3.1 and 4.2. Then we arrive at a statement similar to (4.26). More precisely, it is sufficient to show that for large \( n \)

\[
P\left( \sup_{A \in \mathcal{A}(u)} |v_n(A) - W_n(A)|/q(A) > \varepsilon \right) \leq \varepsilon.
\]

In fact we will show that there exist versions of \( v_n \) (which we still denote with \( v_n \)) and one single Wiener process, with time \( \mu, W \), such that, as \( n \to \infty \),

\[
\sup_{A \in \mathcal{A}(u)} |v_n(A) - W(A)|/q(A) \to 0 \text{ a.s.} \quad (4.34)
\]

We have to prove (4.34) under (A.11) and under (A.12), respectively. First assume (A.11) holds. Then we can omit \( q(A) \) in (4.34). Write again \( \nu = \mu/\lambda(u) \) (\( u \) so large that \( \lambda(u) > 0 \)) and set \( \nu_{(n)} = P_{(n)}/P_{(n)}(E(u)) \). First we show that for all large \( u \)

\[
\sup_{A \in \mathcal{A}(u)} |\nu_{(n)}(A) - \nu(A)| \to 0 \quad (n \to \infty). \quad (4.35)
\]

We have for \( A \in \mathcal{A}'(u) \)
\[\nu_{(n)}(A) - \nu(A) = \frac{P_{(n)}(A)}{P_{(u)}(E(u))} - \frac{\mu(A)}{\mu(E(u))} = \frac{\mu_n(A)}{\mu_n(E(u))} - \frac{\mu(A)}{\mu(E(u))}\]

\[= \frac{\mu_n(A) - \mu(A)}{\mu_n(E(u))} + \frac{\mu(A)}{\mu(E(u))} \frac{\mu(E(u)) - \mu_n(E(u))}{\mu(E(u))\mu_n(E(u))}.\]

Applying the second statement in (A.11) immediately yields (4.35).

Now we will prove (4.34). Set \(r_n = nP_{(n)}(E(u))\) and observe that, as \(n \to \infty\), \(r_n/k_n = nP_{(n)}(E(u))/k_n \to \lambda(u)\) and hence \(r_n \to \infty\) and \(r_n/n \to 0\). Using (4.35) it follows from Corollary 3.1 in Sheehy and Wellner (1992) in conjunction with Dudley (1987, p. 1310) that, as \(m \wedge n \to \infty\),

\[\alpha_{n,m} := m^{1/2}(\nu_{n,m} - \nu_{(n)}) \to_d B, \text{ on } A_{(u)}.\]  \hspace{1cm} (4.36)

where \(\nu_{n,m}\) is the empirical measure based on a random sample of size \(m\) from \(\nu_{(n)}\) and \(B\) is again as in (3.6). Applying the Skorohod-Dudley-Wichura construction (see, e.g., Gaenssler (1983, p. 82)), we can find \(\alpha_{n,m} \overset{d}{=} \alpha_{n,m}\) and \(B \overset{d}{=} B\) such that

\[\lim_{n \to \infty} \sup_{\frac{1}{2}r_n \leq m \leq 2r_n} \sup_{A \in A_{(u)}} |\alpha_{n,m}(A) - B(A)| = 0 \text{ a.s.} \hspace{1cm} (4.37)\]

Define \(R_n = nP_{(n)}(E(u))\) and observe that \(R_n\) is binomially distributed with parameters \(n\) and \(r_n/n\). Using the central limit theorem and again applying the Skorohod construction gives us \(R_n \overset{d}{=} R_n\) and a standard normal random variable \(Z\) such that, as \(n \to \infty\),

\[Z_n := \frac{\overline{R}_n - r_n}{r_n^{1/2}} \to Z \text{ a.s.} \]

Moreover \(\{\alpha_{n,m}\}\) and \(\{Z_n\}\) are independent and \(\overline{B}\) and \(Z\) are independent.

We have

\[\frac{n}{r_n^{1/2}}(P_n(A) - P_{(u)}(A))\]

\[= \left(\frac{R_n}{r_n}\right)^{1/2} R_n^{1/2} \left(\frac{n}{R_n} P_n(A) - \nu_{(n)}(A)\right) + \frac{R_n - r_n}{r_n^{1/2}} \nu_{(n)}(A), \ A \in A_{(u)}.\]

Now the following crucial observation readily follows

\[\frac{n}{r_n^{1/2}}(P_n - P_{(u)}) \overset{d}{=} \left(\frac{\overline{R}_n}{r_n}\right)^{1/2} \alpha_{n,\overline{R}_n} + Z_n \nu_{(n)} =: v'_{(n)}, \text{ on } A_{(u)}.\]

Write \(W'(A) = B(A) + Z \nu(A), \ A \in A_{(u)}\). Then \(W'\) is a Wiener process with 'time' \(\nu\). So we have
Now we show that the right hand side of (4.38) converges to zero almost surely as \( n \to \infty \). The first term can be dealt with using \( \overline{R}_n/r_n \to 1, R_n \to \infty \) a.s. and (4.37). For the second term we use again \( \overline{R}_n/r_n \to 1 \) a.s. and the a.s. boundedness of \( B \) on \( \mathcal{A} \). The third term vanishes since \( Z_n \to Z \) a.s. in combination with (4.35). Finally the fourth term tends to zero a.a. since \( Z_n \to Z \) a.s. and \( \sup_{\mathcal{A}(u)} \nu(A) \leq \nu(E(u)) = 1 \). So we have from (4.38)

\[
\sup_{\mathcal{A}(u)} |v'_n(A) - W'(A)| \to 0 \text{ a.s.} \tag{4.39}
\]

Observe that

\[
v_n = k_n^{1/2} \left( n P_n - f_n \right) = \frac{n}{k_n^{1/2}} \left( P_n - P_{n(t)} \right) \overset{d}{\to} \left( \frac{r_n}{k_n} \right)^{1/2} \nu', \text{ on } \mathcal{A}(u). \tag{4.40}
\]

Set \( W'(A) = \lambda(u))^{1/2} \nu'(A), \ A \in \mathcal{A}(u). \) Then we have

\[
\left( \frac{r_n}{k_n} \right)^{1/2} v'_n(A) - W'(A) = \left( \frac{r_n}{k_n} \right)^{1/2} v'_n(A) - (\lambda(u))^{1/2} W'(A)
\]

\[
= \left( \frac{r_n}{k_n} \right)^{1/2} (v'_n(A) - W'(A)) + \left( \frac{r_n}{k_n} \right)^{1/2} (\lambda(u))^{1/2} W'(A), \ A \in \mathcal{A}(u). \tag{4.41}
\]

But now (4.34) is immediate from (4.39), \( r_n/k_n \to \lambda(u) \) and the a.s. boundedness of \( B \) hence of \( W' \) on \( \mathcal{A}(u) \). This completes the proof of (4.34) under (A.11).

Now assume (A.12) holds. The proof of (4.34) is similar to the proof under (A.11). We sketch it briefly. From the third statement in (A.12) we have by the Skorohod-Dudley-Wichura construction, similar to (4.37),

\[
\lim_{n \to \infty} \sup_{ \frac{1}{2} r_n \leq m \leq 2 r_n } \sup_{\mathcal{A}(u)} |\overline{\alpha}_m(A) - \overline{B}(A)|/q(A) = 0 \text{ a.s.,}
\]

where \( \overline{\alpha}_m \) is similar to \( \overline{\alpha}_{n,m} \), but with \( \nu_{(n)} \) replaced by \( \nu \). Now we can proceed as below (4.37), arriving at an expression like (4.38), namely

\[
\sup_{\mathcal{A}(u)} |v'_n(A) - W'(A)|/q(A)
\]
\[
\left( \frac{R_n}{T_n} \right)^{1/2} \sup_{A \in \mathcal{A}(w)} |(\overline{a}_n(A) - \overline{B}(A))/q(A)|
+ \left( \frac{R_n}{T_n} \right)^{1/2} - 1 \sup_{A \in \mathcal{A}(w)} |\overline{B}(A)/q(A) + |Z_n - Z| \sup_{A \in \mathcal{A}(w)} \nu(A)/q(A). \tag{4.42}
\]

The right hand side of (4.42) converges to zero almost surely as \( n \to \infty \), for similar reasons as those given below (4.38). As in (4.40) and (4.41) this yields (4.34).

After noting that similar remarks apply as at the very end of the proof of Theorem 2.1, the proof is complete. \( \square \)

ACKNOWLEDGEMENT

This paper started as joint work with David M. Mason, who later decided to withdraw from the project. As a consequence, part of the mathematics in this paper is due to him. I wish to thank him for this contribution.

REFERENCES


