AN ANSCOMBE-TYPE THEOREM

PATRIZIA BERTI, IRENE CRIMALDI, LUCA PRATELLI, AND PIETRO RIGO

Abstract. Let (X_n) be a sequence of random variables (with values in a separable metric space) and (N_n) a sequence of random indices. Conditions for X_{N_n} to converge stably (in particular, in distribution) are provided. Some examples, where such conditions work but those already existing fail, are given as well.

1. Introduction

Anscombe's theorem (AT) gives conditions for X_{N_n} to converge in distribution, where (X_n) is a sequence of random variables and (N_n) a sequence of random indices. Roughly speaking, such conditions are: (i) $N_n \to \infty$ in some sense; (ii) X_n converges in distribution; (iii) For large n, X_j is close to X_n provided j is close to n. (Precise definitions are given in Subsection 3.2).

In particular, in AT, condition (i) is realized as

(a) $N_n/k_n \xrightarrow{P} u$, where $k_n > 0$ and $u > 0$ are constants and $k_n \to \infty$.

Under (a), it is very hard to improve on AT. The only possibility is to look for some optimal form of condition (iii). See e.g. [6].

But condition (a) is often generalized into

(a*) $N_n/k_n \xrightarrow{P} U$, where $U > 0$ is a random variable.

For instance, condition (a*) suffices for X_{N_n} to converge in distribution in case $X_n = n^{-1/2}\sum_{i=1}^{n}\{Z_i - E(Z_1)\}$, where (Z_n) is an i.i.d. sequence with $E(Z_1^2) < \infty$. However, under (a*), convergence in distribution of X_n is not enough. To get converge in distribution of X_{N_n}, condition (ii) is to be strengthened.

One natural solution is to request *stable* convergence of X_n. This is made precise by a result of Zhang Bo [8] (Theorem 1 in the sequel). According to Theorem 1, X_{N_n} converges stably (in particular, in distribution) provided X_n converges stably, condition (a*) holds, and some form of (iii) is satisfied. The statement of (iii) depends on whether U is, or it is not, discrete.

In this paper, Theorem 1 is (strictly) improved. Our main result (Theorem 2 in the sequel) has two possible merits. It does not depend on whether U is discrete. And, more importantly, it requests a form of (iii) weaker than the corresponding one in Theorem 1. Indeed, in Theorem 1, the asked version of (iii) does not involve the N_n. As a consequence, it potentially works for *every* sequence (N_n) of random

2010 *Mathematics Subject Classification.* 60B10, 60F05, 60G09, 60G57.
Key words and phrases. Anscombe theorem, Exchangeability, Random indices, Random sums, Stable convergence.

times but it is also rather strong. Instead, in Theorem 2, we exploit a form of (iii) which is tailor-made on the particular sequence of random times at hand.

A few examples, where Theorem 2 works but Theorem 1 fails, are given as well. We mention Examples 6 and 7 concerning the exchangeable CLT and the exchangeable empirical process.

2. Stable convergence

Let $\mathcal{X}$ be a metric space and $(\Omega, \mathcal{A}, P)$ a probability space. A *kernel* (or a *random probability measure*) on $\mathcal{X}$ is a map K on Ω such that:

- $K(\omega)$ is a Borel probability measure on $\mathcal{X}$ for each $\omega \in \Omega$;
- $\omega \mapsto K(\omega)(B)$ is $\mathcal{A}$-measurable for each Borel set $B \subset \mathcal{X}$.

For every bounded Borel function $f : \mathcal{X} \to \mathbb{R}$, we let $K(f)$ denote the real random variable

$$K(\omega)(f) = \int f(x)\, K(\omega)(dx).$$

Let (X_n) be a sequence of $\mathcal{X}$-valued random variables on $(\Omega, \mathcal{A}, P)$. Given a Borel probability measure μ on $\mathcal{X}$, say that X_n converges in distribution to μ if $\mu(f) = \lim_n E\{f(X_n)\}$ for all bounded continuous functions $f : \mathcal{X} \to \mathbb{R}$. In this case, we also write $X_n \xrightarrow{d} X$ for any $\mathcal{X}$-valued random variable X with distribution μ. Next, let $\mathcal{G} \subset \mathcal{A}$ be a sub-σ-field and K a kernel on $\mathcal{X}$. Say that X_n *converges $\mathcal{G}$-stably to K* if

$$E\{K(f) \mid H\} = \lim_n E\{f(X_n) \mid H\}$$

for all $H \in \mathcal{G}$ with $P(H) > 0$ and all bounded continuous $f : \mathcal{X} \to \mathbb{R}$.

$\mathcal{G}$-stable convergence always implies convergence in distribution (just let $H = \Omega$). Further, it reduces to convergence in distribution for $\mathcal{G} = \{\emptyset, \Omega\}$ and is connected to convergence in probability for $\mathcal{G} = \mathcal{A}$. Suppose in fact $\mathcal{X}$ is separable and take an $\mathcal{X}$-valued random variable X on $(\Omega, \mathcal{A}, P)$. Then, $X_n \xrightarrow{P} X$ if and only if X_n converges $\mathcal{A}$-stably to the kernel $K = \delta_X$.

We refer to [3] and references therein for more on stable convergence.

3. Results

3.1. **Notation.** All random variables appearing in the sequel, unless otherwise stated, are defined on a fixed probability space $(\Omega, \mathcal{A}, P)$.

Let (S, d) be a separable metric space. The basic ingredients are three sequences

$$(X_n : n \geq 0), \quad (N_n : n \geq 0), \quad (k_n : n \geq 0),$$

where the X_n are S-valued random variables, the N_n are random times (i.e., random variables with values in $\{0, 1, 2, \ldots\}$) and the k_n are strictly positive constants such that $k_n \to \infty$. We let

$$M_n(\delta) = \max_{j : |n-j| \leq n\,\delta} d(X_j, X_n)$$

for all $n \geq 0$ and $\delta > 0$. Finally, K denotes a kernel on S.

3.2. **Classical Anscombe's theorem and one of its developments.** Let μ be a Borel probability measure on S. According to AT, for X_{N_n} to converge in distribution to μ, it suffices that

(a) $N_n/k_n \xrightarrow{P} u$, where $u > 0$ is a constant;
(b) X_n converges in distribution to μ;
(c) $\inf_{\delta>0} \limsup_n P\big(M_n(\delta) > \epsilon\big) = 0$ for all $\epsilon > 0$.

Soon after its appearance, AT has been investigated and developed in various ways. See e.g. [4], [5], [6], [7], [8] and references therein. To our knowledge, most results preserve the structure of the classical AT, for they lead to convergence of X_{N_n} (in distribution or stably) under suitable versions of conditions (a)-(b)-(c). In particular, much attention is paid to possible alternative versions of condition (c). Also, as remarked in Section 1, condition (a) is often generalized into

(a*) $N_n/k_n \xrightarrow{P} U$, where $U > 0$ is a random variable.

Replacing (a) with (a*) is not free but implies strengthening (b) and/or (c). A remarkable example is the following. In the sequel, U denotes a real random variable and $\mathcal{G}$ a sub-σ-field of $\mathcal{A}$ such that

$$U > 0 \quad \text{and} \quad \sigma(U) \subset \mathcal{G}.$$

Theorem 1. (Zhang Bo [8]**).** *Let U be strictly positive and $\mathcal{G}$-measurable. Suppose condition (a*) holds and*

(b*) *X_n converges $\mathcal{G}$-stably to K.*

Then, X_{N_n} converges $\mathcal{G}$-stably to K provided condition (c) holds and U is discrete. Or else, X_{N_n} converges $\mathcal{G}$-stably to K provided

(c*) *For each $\epsilon > 0$, there is $\delta > 0$ such that*

$$\limsup_n P\big(M_n(\delta) > \epsilon \mid H\big) < \epsilon \quad \text{for all } H \in \mathcal{G} \text{ with } P(H) > 0.$$

Theorem 1 is our starting point. Roughly speaking, it can be summarized as follows. Suppose (a*) and (c) hold but (a) fails. If U is discrete, X_{N_n} still converges in distribution (in fact, it converges stably) up to replacing (b) with (b*). If U is not discrete, instead, condition (c) should be strengthened as well.

3.3. **Improving Theorem 1.** Suppose conditions (a*)-(b*) hold but U is not necessarily discrete. As implicit in Theorem 1, it may be that (c) holds and yet X_{N_n} fails to converge $\mathcal{G}$-stably to K; see Example 4. Hence, to get $X_{N_n} \xrightarrow{\mathcal{G}-stably} K$, condition (c) is to be modified. Plainly, a number of conditions could serve to this purpose. We now investigate two of them.

One (crude) possibility is just replacing n with N_n in condition (c), that is,

(d) $\inf_{\delta>0} \limsup_n P\big(M_{N_n}(\delta) > \epsilon\big) = 0$ for all $\epsilon > 0$,

where

$$M_{N_n}(\delta)=\max_{j:|N_n-j|\le N_n\,\delta} d(X_j,X_{N_n}).$$

Unlike condition (c*) of Theorem 1, which works for *every* sequence N_n (as far as (a*) and (b*) are satisfied), condition (d) is tailor-made on the particular sequence of random times at hand.

In view of (a*), another option is replacing $M_n(\delta)$ with

$$M_{[k_n\,U]}(\delta)=\max_{j:|[k_n\,U]-j|\le[k_n\,U]\,\delta} d(X_j,X_{[k_n\,U]}).$$

The corresponding condition is

(e) $\inf_{\delta>0}\limsup_n P\big(M_{[k_n\,U]}(\delta)>\epsilon\big)=0$ for all $\epsilon>0$.

Conditions (d) and (e) are actually equivalent. More importantly, they lead to the desired conclusion.

Theorem 2. *Let U be strictly positive and $\mathcal{G}$-measurable. Conditions (d) and (e) are equivalent under (a*). Moreover,*

$$X_{N_n}\overset{\mathcal{G}-stably}{\longrightarrow}K\quad\text{and}\quad X_{[k_n\,U]}\overset{\mathcal{G}-stably}{\longrightarrow}K$$

under conditions (a)-(b*)-(d) (or equivalently (a*)-(b*)-(e)).*

Proof. Let $R_n=[k_n\,U]$. We first show that (d) and (e) are equivalent under (a*).

Suppose (a*) and (e) hold and fix $\delta\in(0,1]$. If $|R_n-N_n|\le\delta\,R_n$ and j is such that $|j-N_n|\le\delta\,N_n$, then

$$|j-R_n|\le|j-N_n|+\delta\,R_n\le\delta\,N_n+\delta\,R_n\le 2\,\delta\,R_n+\delta\,|R_n-N_n|\le 3\,\delta\,R_n.$$

Hence, $|R_n-N_n|\le\delta\,R_n$ implies

$$M_{N_n}(\delta)\le d(X_{R_n},X_{N_n})+\max_{j:|j-R_n|\le 3\delta\,R_n} d(X_j,X_{R_n})\le 2\,M_{R_n}(3\,\delta).$$

Given $\epsilon>0$, it follows that

$$P\big(M_{N_n}(\delta)>\epsilon\big)\le P\big(|R_n-N_n|>\delta\,R_n\big)+P\big(M_{R_n}(3\,\delta)>\epsilon/2\big).$$

By (a*), $N_n/R_n\overset{P}{\longrightarrow}1$ so that $\lim_n P\big(|R_n-N_n|>\delta\,R_n\big)=0$. Therefore,

$$\limsup_n P\big(M_{N_n}(\delta)>\epsilon\big)\le\limsup_n P\big(M_{R_n}(3\,\delta)>\epsilon/2\big)$$

and condition (d) follows from condition (e). By precisely the same argument, it can be shown that (a*) and (d) imply (e).

Next, assume conditions (a*)-(b*)-(e). Since

$$d(X_{R_n},X_{N_n})\le M_{R_n}(\delta)\quad\text{provided } |R_n-N_n|\le\delta\,R_n,$$

conditions (a*) and (e) yield $d(X_{R_n},X_{N_n})\overset{P}{\longrightarrow}0$. Thus, it suffices to prove that $X_{R_n}\overset{\mathcal{G}-stably}{\longrightarrow}K$. To this end, for each $\delta\in(0,1]$, define

$$U_\delta=\delta\,I_{\{0<U\le\delta\}}+\sum_{j=1}^\infty j\,\delta\,I_{\{j\,\delta<U\le(j+1)\,\delta\}}\quad\text{and}\quad R_n(\delta)=[k_n\,U_\delta].$$

Since U_δ is discrete, strictly positive and $\mathcal{G}$-measurable, condition (b*) yields $X_{R_n(\delta)} \overset{\mathcal{G}-stably}{\longrightarrow} K$. Fix in fact $H \in \mathcal{G}$ with $P(H) > 0$ and let $H_j = H \cap \{U_\delta = j\,\delta\}$ for all $j \geq 1$. Then, (b*) implies

$$\lim_n E\big\{f(X_{R_n(\delta)}) \mid H\big\} = \lim_n \sum_j E\big\{f(X_{[k_n\, j\, \delta]}) \mid H_j\big\}\, P(H_j \mid H)$$
$$= \sum_j E\big\{K(f) \mid H_j\big\}\, P(H_j \mid H) = E\big\{K(f) \mid H\big\}$$

for each bounded continuous f, where the sum is over those j such that $P(H_j) > 0$.

Note also that, on the set $\{U > \delta\}$, one obtains

$$|R_n - R_n(\delta^2)| = R_n - R_n(\delta^2) = R_n \, \frac{[k_n\, U] - [k_n\, U_{\delta^2}]}{[k_n\, U]}$$
$$< R_n \, \frac{k_n\,(U - U_{\delta^2}) + 1}{k_n\, U - 1} < R_n \, \frac{k_n\, \delta^2 + 1}{k_n\, \delta - 1} < 2\,\delta\, R_n \quad \text{for large } n.$$

Thus, for $\epsilon > 0$ and large n,

$$P\Big(d(X_{R_n}\,,\, X_{R_n(\delta^2)}\,) > \epsilon\Big) \leq P(U \leq \delta) + P\big(M_{R_n}(2\,\delta) > \epsilon\big).$$

By condition (e) and since $U > 0$, it follows that

$$\inf_{\delta > 0} \limsup_n P\Big(d(X_{R_n}\,,\, X_{R_n(\delta^2)}\,) > \epsilon\Big) = 0. \tag{1}$$

Finally, fix $\epsilon > 0$, $H \in \mathcal{G}$ with $P(H) > 0$, and a closed set $C \subset S$. Let $C_\epsilon = \{x \in S : d(x, C) \leq \epsilon\}$. By (1), there is $\delta \in (0, 1]$ such that

$$\limsup_n P\Big(d(X_{R_n}\,,\, X_{R_n(\delta^2)}\,) > \epsilon\Big) < \epsilon\, P(H).$$

With such a δ, since $X_{R_n(\delta^2)} \overset{\mathcal{G}-stably}{\longrightarrow} K$, one obtains

$$\limsup_n P\big(X_{R_n} \in C \mid H\big) \leq \limsup_n \Big\{P\Big(d(X_{R_n}\,,\, X_{R_n(\delta^2)}\,) > \epsilon \mid H\Big) + P\Big(X_{R_n(\delta^2)} \in C_\epsilon \mid H\Big)\Big\}$$
$$< \epsilon + \limsup_n P\Big(X_{R_n(\delta^2)} \in C_\epsilon \mid H\Big) \leq \epsilon + E\big\{K(C_\epsilon) \mid H\big\}.$$

As $\epsilon \to 0$, it follows that $\limsup_n P\big(X_{R_n} \in C \mid H\big) \leq E\big\{K(C) \mid H\big\}$. Therefore, $X_{R_n} \overset{\mathcal{G}-stably}{\longrightarrow} K$ and this concludes the proof.

□

Theorem 2 unifies the two parts of Theorem 1 (U discrete and U not discrete). In addition, Theorem 2 strictly improves Theorem 1. In fact, condition (c*) implies condition (e) but not conversely. Two (natural) examples where (e) holds and (c*) fails are given in the next section; see Examples 5 and 6. Here, we prove the direct implication.

Theorem 3. *Let U be strictly positive and $\mathcal{G}$-measurable. If condition (c) holds and U is discrete, or if condition (c*) holds, then condition (e) holds.*

Proof. Let $R_n = [k_n U]$. Suppose (c) holds and U is discrete. Then it suffices to note that, for each $\epsilon > 0$ and $u > 0$ such that $P(U = u) > 0$, one obtains

$$\limsup_n P\big(M_{R_n}(\delta) > \epsilon \mid U = u\big) = \limsup_n P\big(M_{[k_n u]}(\delta) > \epsilon \mid U = u\big)$$
$$\leq P(U = u)^{-1} \limsup_n P\big(M_n(\delta) > \epsilon\big) \longrightarrow 0 \quad \text{as } \delta \to 0.$$

Next, suppose (c*) holds. Given $\epsilon > 0$, take $\delta > 0$ such that

$$\limsup_n P\big(M_n(\delta) > \epsilon/2 \mid H\big) < \epsilon/2 \quad \text{for all } H \in \mathcal{G} \text{ with } P(H) > 0.$$

Fix u, $\gamma > 0$ and define $H = \{u - \gamma \leq U < u + \gamma\}$. Take j and n such that

$$|j - R_n| \leq (\delta/4)\, R_n, \quad k_n\,\gamma > 1, \quad k_n\,u < 2\,[k_n\,u].$$

On the set H, one obtains

$$|j - [k_n\,u]| \leq |j - R_n| + |R_n - [k_n\,u]| \leq (\delta/4)\,R_n + |[k_n\,U] - [k_n\,u]|$$
$$< (\delta/4)\,k_n\,(u + \gamma) + k_n\,\gamma + 1 < [k_n\,u]\,\frac{2}{u}\,\{(\delta/4)\,(u + \gamma) + 2\,\gamma\}.$$

Letting $\delta^* = (2/u)\,\big\{(\delta/4)\,(u + \gamma) + 2\,\gamma\big\}$, it follows that

$$M_{R_n}(\delta/4) \leq M_{[k_n\,u]}(\delta^*) + d\big(X_{R_n}\,,\, X_{[k_n\,u]}\big) \leq 2\,M_{[k_n\,u]}(\delta^*)$$

on H for large n. Since $H \in \mathcal{G}$,

$$\limsup_n P\big(M_{R_n}(\delta/4) > \epsilon \mid H\big) \leq \limsup_n P\big(M_{[k_n\,u]}(\delta^*) > \epsilon/2 \mid H\big)$$
$$\leq \limsup_n P\big(M_n(\delta^*) > \epsilon/2 \mid H\big) < \epsilon/2$$

provided $P(H) > 0$ and u, γ are such that $\delta^* \leq \delta$, or equivalently

$$\frac{\gamma}{u} \leq \frac{\delta}{8 + \delta}.$$

Finally, take $0 < a < b$ such that $P(a \leq U < b) > 1 - (\epsilon/2)$. The set $\{a \leq U < b\}$ can be partitioned into sets $H_i = \{u_i - \gamma \leq U < u_i + \gamma\}$ such that $(\gamma/a) \leq \delta/(8+\delta)$ and $u_1 = a + \gamma < u_2 < \ldots$. On noting that $(\gamma/u_i) \leq \delta/(8 + \delta)$ for all i,

$$\limsup_n P\big(M_{R_n}(\delta/4) > \epsilon\big) < \epsilon/2 + \limsup_n P\big(M_{R_n}(\delta/4) > \epsilon,\, a \leq U < b\big)$$
$$\leq \epsilon/2 + \sum_i \limsup_n P\big(M_{R_n}(\delta/4) > \epsilon \mid H_i\big)\, P(H_i) < \epsilon$$

where the sum is over those i with $P(H_i) > 0$. This concludes the proof. $\square$

4. Examples

It is implicit in Theorem 1 that, when U is not discrete, conditions (a*)-(b*)-(c) are not enough for $X_{N_n} \overset{\mathcal{G}-stably}{\longrightarrow} K$ (where K is the kernel involved in condition (b*)). However, we do not know of any explicit example. So, we begin with one such example.

Example 4. (Conditions (a*)-(b*)-(c) do not imply $X_{N_n} \overset{\mathcal{G}-stably}{\longrightarrow} K$). Let $\Omega = [0, 1)$, $\mathcal{A}$ the Borel σ-field and P the Lebesgue measure. For each $n \geq 1$, define

$$A_n = \big[\log n,\, \log(n + 1)\big) \text{ modulo } 1,$$

that is, $A_1 = [0, \log 2)$, $A_2 = [\log 2, 1) \cup [0, (\log 3) - 1)$ and so on. Define also $X_0 = 0$ and $X_n = I_{A_n}$ for $n \geq 1$. Since $P(A_n) = \log((n+1)/n)$, then $X_n \xrightarrow{P} 0$, or equivalently X_n converges $\mathcal{A}$-stably to the point mass at 0 (see Section 2). Thus, condition (b*) holds with $\mathcal{G} = \mathcal{A}$ and K the point mass at 0. Given $\epsilon > 0$,

$$P\big(M_n(\delta) > \epsilon,\, X_n = 0\big) \leq P\Big(\bigcup_{j:|n-j|\leq n\,\delta} A_j\Big) \leq \sum_{j:|n-j|\leq n\,\delta} P(A_j) \leq \log \frac{[n\,(1+\delta)] + 1}{[n\,(1-\delta)]}.$$

Since $P(X_n = 0) \to 1$, it follows that

$$\limsup_n P\big(M_n(\delta) > \epsilon\big) = \limsup_n P\big(M_n(\delta) > \epsilon,\, X_n = 0\big) \leq \log \frac{1+\delta}{1-\delta},$$

that is, condition (c) holds. Finally, define $U(\omega) = \exp(\omega)$ for all $\omega \in [0,1)$ and

$$N_n = [U\, \exp(r_n)],$$

where the r_n are non-negative integers such that $r_n \to \infty$. Condition (a*) is trivially true. Further, for each n, one obtains $\{N_n = k\} \subset A_k$ for all k, so that $X_{N_n} = 1$. Thus, X_{N_n} fails to converge $\mathcal{A}$-stably to the point mass at 0.

We next prove that condition (e) does not imply condition (c*). We give two examples. The first is just a modification of Example 4, while the second (which requires some more calculations) concerns the exchangeable CLT. Recall that (d) and (e) are equivalent under (a*).

Example 5. (Example 4 revisited). Conditions (b*)-(c)-(c*) depend on (X_n) and $\mathcal{G}$ only. In view of Theorem 1, condition (c*) fails in Example 4. Hence, to build an example where (c*) fails but (a*)-(b*)-(c)-(d) hold, it suffices to suitably modify the random times N_n of Example 4. Precisely, suppose $(\Omega, \mathcal{A}, P)$, U, (X_n) and $\mathcal{G}$ are as in Example 4, but the random times are now

$$N_n = \Big[\frac{T_{n-1} + T_n}{2}\Big] \quad \text{where } T_n = \inf\{j : j > T_{n-1} \text{ and } X_j = 1\} \text{ and } N_0 = T_0 = 0.$$

Then, (c*) fails while (b*)-(c) hold. It is not hard to see that $T_n = [\exp(n-1)\, U]$ for $n \geq 1$. Thus, conditions (a*) and (d) are both trivially true. (As to (d), just note that $T_{n-1} < N_n\,(1-\delta) < N_n\,(1+\delta) < T_n$ for large n and small δ).

Example 6. (Exchangeable CLT). Let $(Z_n : n \geq 1)$ be an exchangeable sequence of real random variables with tail σ-field $\mathcal{T}$. By de Finetti's theorem, (Z_n) is i.i.d. conditionally on $\mathcal{T}$. Basing on this fact, if $E(Z_1^2) < \infty$, it is not hard to see that

$$\frac{\sum_{i=1}^n \{Z_i - E(Z_1 \mid \mathcal{T})\}}{\sqrt{n}} \xrightarrow{\mathcal{A}-stably} N(0, L)$$

where $L = E(Z_1^2 \mid \mathcal{T}) - E(Z_1 \mid \mathcal{T})^2$ and $N(0, \sigma^2)$ denotes the Gaussian law with mean 0 and variance σ^2 (with $N(0,0)$ the point mass at 0); see e.g. Theorem 3.1 of [1] and the subsequent remark. Fix a $\mathcal{T}$-measurable random variable $U > 0$ and define

$$N_n = [n\, U], \quad X_0 = 0, \quad X_n = \frac{\sum_{i=1}^n \{Z_i - E(Z_1 \mid \mathcal{T})\}}{\sqrt{n}}.$$

Then, conditions (a*)-(b*)-(c)-(d) are satisfied (with $\mathcal{G} = \mathcal{A}$ and $K = N(0, L)$) so that

$$\frac{\sum_{i=1}^{N_n} \{Z_i - E(Z_1 \mid \mathcal{T})\}}{\sqrt{N_n}} \xrightarrow{\mathcal{A}-stably} N(0, L)$$

because of Theorem 2. Indeed, (a*)-(b*) are obvious and (c) can be checked precisely as (d). As to (d), given $\epsilon > 0$, just note that

$$\limsup_n P\big(M_{N_n}(\delta) > \epsilon \mid \mathcal{T}\big) \leq \limsup_n P\big(M_n(\delta) > \epsilon \mid \mathcal{T}\big) \quad \text{a.s.}$$

for N_n is $\mathcal{T}$-measurable, and

$$\limsup_n P\big(M_n(\delta) > \epsilon \mid \mathcal{T}\big) \xrightarrow{a.s.} 0 \quad \text{as } \delta \to 0$$

for (Z_n) is i.i.d. conditionally on $\mathcal{T}$. Thus,

$$\begin{aligned}\limsup_n P\big(M_{N_n}(\delta) > \epsilon\big) &\leq \int \limsup_n P\big(M_{N_n}(\delta) > \epsilon \mid \mathcal{T}\big)\, dP \\ &\leq \int \limsup_n P\big(M_n(\delta) > \epsilon \mid \mathcal{T}\big)\, dP \longrightarrow 0 \quad \text{as } \delta \to 0.\end{aligned}$$

It remains to see that condition (c*) may fail. We verify this fact for

$$\mathcal{G} = \sigma(U) \quad \text{and} \quad Z_n = U\, V_n$$

where

- U is any random variable such that $U > 0$, $E(U^2) < \infty$ and $P(U > u) > 0$ for all $u > 0$;
- (V_n) is i.i.d., $V_1 \sim N(0, 1)$, and (V_n) is independent of U.

Such a sequence (Z_n) is exchangeable and $E(Z_1^2) = E(U^2) < \infty$. Furthermore, $E(Z_1 \mid \mathcal{T}) = 0$ a.s. and U is $\mathcal{T}$-measurable (up to modifications on P-null sets) for

$$\frac{\sum_{i=1}^n Z_i}{n} = U\, \frac{\sum_{i=1}^n V_i}{n} \xrightarrow{a.s.} 0 \quad \text{and} \quad \frac{\sum_{i=1}^n Z_i^2}{n} = U^2\, \frac{\sum_{i=1}^n V_i^2}{n} \xrightarrow{a.s.} U^2.$$

Next, a direct calculation shows that

$$\frac{\sum_{i=1}^n V_i}{\sqrt{n}} - \frac{\sum_{i=1}^m V_i}{\sqrt{m}} \sim N\big(0,\, 2 - 2\sqrt{n/m}\big) \quad \text{for } 1 \leq n \leq m.$$

Thus, conditionally on U,

$$X_n - X_{[n\,(1-\delta)]} = U \left\{ \frac{\sum_{i=1}^n V_i}{\sqrt{n}} - \frac{\sum_{i=1}^{[n\,(1-\delta)]} V_i}{\sqrt{[n\,(1-\delta)]}} \right\} \sim N\big(0,\, U^2\, \sigma_n^2(\delta)\big)$$

where $\delta \in (0, 1)$ and

$$\sigma_n^2(\delta) = 2 - 2\sqrt{\frac{[n\,(1-\delta)]}{n}} \geq 2 - 2\sqrt{1-\delta}.$$

Define $H = \{U > u\}$ and $f(\delta) = 2\sqrt{2 - 2\sqrt{1-\delta}}$ for some $u > 0$ and $\delta \in (0, 1/2)$. Letting Φ denote the standard normal distribution function, for each n such that

$n - [n\,(1-\delta)] \leq n\,2\,\delta$, one obtains

$$P\big(M_n(2\,\delta) > 1/2 \mid H\big) \geq P\Big(|X_n - X_{[n\,(1-\delta)]}| > 1/2 \mid H\Big)$$
$$= P(H)^{-1} \int_H P\Big(|X_n - X_{[n\,(1-\delta)]}| > 1/2 \mid U\Big)\,dP$$
$$= P(H)^{-1} \int_H 2\,\Phi\big(-\frac{1}{2\,U\,\sigma_n(\delta)}\big)\,dP$$
$$\geq 2\,P(H)^{-1} \int_H \Phi\big(-\frac{1}{U\,f(\delta)}\big)\,dP \geq 2\,\Phi\big(-\frac{1}{u\,f(\delta)}\big).$$

Since $P(U > u) > 0$ for all $u > 0$, condition (c*) (applied with $\epsilon = 1/2$) would imply $\Phi\big(-\frac{1}{u\,f(\delta)}\big) < 1/4$ for some fixed δ and all $u > 0$. But this is absurd for $\lim_{u\to\infty} \Phi\big(-\frac{1}{u\,f(\delta)}\big) = \Phi(0) = 1/2$. Therefore, (c*) fails in this example.

Our last example deals with empirical processes for non independent data. Let $l^\infty(\mathbb{R})$ denote the space of real bounded functions on $\mathbb{R}$ equipped with uniform distance.

Example 7. (Exchangeable empirical processes). Again, let $(Z_n : n \geq 1)$ be an exchangeable sequence of real random variables with tail σ-field $\mathcal{T}$. Let F be a random distribution function satisfying

$$F(t) = P(Z_1 \leq t \mid \mathcal{T}) \quad \text{a.s. for all } t \in \mathbb{R}.$$

The n-th empirical process can be defined as

$$X_n(t) = \sqrt{n}\,\Big\{(1/n) \sum_{i=1}^n I_{\{Z_i \leq t\}} - F(t)\Big\} \quad \text{for } t \in \mathbb{R}.$$

Define also the process $X(t) = \mathbb{B}\big(F(t)\big)$, $t \in \mathbb{R}$, where $\mathbb{B}$ is a Brownian-bridge process independent of F. (Such a $\mathbb{B}$ is available up to enlarging the basic probability space $(\Omega, \mathcal{A}, P)$). If $P(Z_1 = Z_2) = 0$ or if Z_1 is discrete, then $X_n \xrightarrow{d} X$ in the metric space $l^\infty(\mathbb{R})$; see [1]-[2] for details. But $l^\infty(\mathbb{R})$ is not separable and working with it yields various measurability issues. So, to avoid technicalities, we assume $0 \leq Z_1 \leq 1$ and we take S to be the space of real cadlag functions on $[0,1]$ equipped with Skorohod distance. Then, $X_n \xrightarrow{d} X$ in the separable metric space S; see e.g. Theorem 3 of [2]. Actually, basing on de Finetti's theorem, it can be shown that X_n converges $\mathcal{A}$-stably to a certain kernel K on S. Precisely, for each distribution function H, let Q_H denote the probability distribution (on the Borel sets of S) of the process $X_H(t) = \mathbb{B}\big(H(t)\big)$, $t \in [0,1]$. Then, K can be written as

$$K(A) = Q_F(A) \quad \text{for all Borel sets } A \subset S.$$

Finally, let $N_n = [n\,U]$ where $U > 0$ is any $\mathcal{T}$-measurable random variable. Then, condition (a*) is trivially true, (b*) holds with $\mathcal{G} = \mathcal{A}$, and (d) can be checked as in Example 6. Thus, Theorem 2 implies $X_{N_n} \xrightarrow{\mathcal{A}-stably} K$. This fact can not be deduced by Theorem 1, however, for condition (c*) may fail.

References

[1] Berti P., Pratelli L., Rigo P. (2004) Limit theorems for a class of identically distributed random variables, *Ann. Probab.*, 32, 2029-2052.

[2] Berti P., Pratelli L., Rigo P. (2006) Asymptotic behaviour of the empirical process for exchangeable data, *Stoch. Proc. Appl.*, 116, 337-344.
[3] Crimaldi I., Letta G., Pratelli L. (2007) A strong form of stable convergence, *Seminaire de Probabilites* XL, Lect. Notes in Math., 1899, 203-225.
[4] Gnedenko B.V., Korolev V.Y. (1996) *Random summation: limit theorems and applications*, CRC Press.
[5] Gut A. (2011) Anscombe's theorem 60 years later, *U.U.D.M. Report* 2011:14.
[6] Nzi M., Theodorescu R. (1990) Anscombe's condition revisited, *Bull. Polish Acad. Sci. Math.*, 38, 55-60.
[7] Silvestrov D.S. (2004) *Limit theorems for randomly stopped stochastic processes*, Springer.
[8] Zhang Bo (2000) Stable convergence of random sequences with random indices, *J. Math. Sciences*, 99, 1515-1526.

Patrizia Berti, Dipartimento di Matematica Pura ed Applicata "G. Vitali", Universita' di Modena e Reggio-Emilia, via Campi 213/B, 41100 Modena, Italy
E-mail address: patrizia.berti@unimore.it

Irene Crimaldi, IMT Institute for Advanced Studies, Piazza San Ponziano 6, 55100 Lucca, Italy
E-mail address: irene.crimaldi@imtlucca.it

Luca Pratelli, Accademia Navale, viale Italia 72, 57100 Livorno, Italy
E-mail address: pratel@mail.dm.unipi.it

Pietro Rigo (corresponding author), Dipartimento di Matematica "F. Casorati", Universita' di Pavia, via Ferrata 1, 27100 Pavia, Italy
E-mail address: pietro.rigo@unipv.it