

One more approach to the convergence of the empirical process to the Brownian bridge

Jean-François Marckert

CNRS, LaBRI, Université Bordeaux 1
351 cours de la Libération
33405 Talence cedex, France

Abstract: A theorem of Donsker asserts that the empirical process converges in distribution to the Brownian bridge. The aim of this paper is to provide a new and simple proof of this fact.

AMS 2000 subject classifications: Primary 62G30, 60F17.

Keywords and phrases: Empirical process; Donsker Theorem; Brownian bridge.

Let (U_i) be a sequence of i.i.d. random variables uniformly distributed on $[0, 1]$, and let F_n be the so-called cumulative empirical function, associated with the n first U_i 's:

$$F_n(t) := n^{-1} \# \{U_i \leq t, i \in \{1, \dots, n\}\}, \quad t \in [0, 1].$$

The sequence of processes (F_n) converges pointwise a.s. on $[0, 1]$ to F defined by $F(t) = t$; this is a consequence of the strong law of large numbers. The Glivenko-Cantelli theorem asserts that this a.s. convergence stands also for the uniform convergence: a.s. $\sup_{x \in [0, 1]} |F_n(x) - F(x)| \xrightarrow[n]{} 0$. To see this, take $0 = x_1 < \dots < x_k = 1$ and check that by monotonicity of F_n and F , $\sup_{x \in [0, 1]} |F_n(x) - F(x)| \leq \max_j \max(|F_n(x_{j+1}) - F(x_j)|, |F_n(x_j) - F(x_{j+1})|) \xrightarrow[n]{(a.s.)} \max x_{j+1} - x_j$, which may be chosen as small as wanted.

In some sense, Donsker's Theorem [1] provides the second term in this convergence. Consider

$$\mathbf{b}_n(t) := \sqrt{n} (F_n(t) - F(t)), \quad t \in [0, 1]. \quad (1)$$

Theorem 1 (Donsker [1]) *The sequence (\mathbf{b}_n) converges in distribution to the Brownian bridge \mathbf{b} on $D[0, 1]$ the space of càdlàg functions on $[0, 1]$ equipped with the Skorohod topology.*

Note 1 *When the variables U_i 's are not uniform, the study of the empirical process reduces to the uniform case thanks to a classical "time change" involving the inverse of the cumulative function of the U_i 's. Some problems of continuity arise due to the atoms of the U_i 's but roughly speaking one may say that all the difficulties are present in the case of the uniform distribution.*

We recall that the Brownian bridge is the continuous centered Gaussian process such that $\text{cov}(\mathbf{b}(s), \mathbf{b}(t)) = s(1-t)$ when $0 \leq s \leq t \leq 1$. It owns the following trajectorial representation :

$$(\mathbf{b}(t))_{t \in [0,1]} \stackrel{(d)}{=} (\mathbf{B}_t - t \mathbf{B}_1)_{t \in [0,1]}, \quad (2)$$

where \mathbf{B} is the standard Brownian motion. This may immediately be checked using that \mathbf{B} is a centered Gaussian process such that $\text{cov}(\mathbf{B}_s, \mathbf{B}_t) = \min(s, t)$.

In fact, Donsker proves only in details $\max \mathbf{b}_n \xrightarrow[n]{(d)} \max \mathbf{b}$ justifying the Doob's heuristic [3]. One may find in the literature numerous more or less direct proofs of Theorem 1. See e.g. Billingsley [2] (and references therein), Kallenberg [4], and also more advanced proofs and constructions (and stronger results) as that of Komlós, Major and Tusnády [5]. Some books are devoted to the convergence of empirical measures and processes : we send the interested reader to Shorack & Wellner [6], van der Vaart & Wellner [7]. As a matter of fact, usual proofs of Theorem 1 use often some advanced constructions or are treated in probability books when a lot of materials have been introduced, leading to some intricate and complex proofs, quite difficult to be taught entirely to beginners. The aim of this paper is to present a new proof of Theorem 1 using only "simple" arguments: only immediate considerations about the weak convergence in $C[0, 1]$ and $D[0, 1]$, and the other very famous Donsker's Theorem which says that a rescaled random walk converges to the Brownian motion are used. The Appendix recalls this material.

We begin the proof of Theorem 1 following the steps of Donsker [1]. We say that a random vector $(M_i)_{i=1, \dots, n}$ has the multinomial distribution with parameters (k, p_1, \dots, p_n) , we write $(M_i)_{i=1, \dots, n} \sim \text{mult}(k, p_1, \dots, p_n)$, when $\mathbb{P}(M_i = m_i, i = 1, \dots, n) = \frac{k!}{\prod_{i=1}^n m_i!} \prod_{i=1}^k p_i^{m_i}$ for any prescribed non negative integers m_1, \dots, m_k summing to k , and 0 otherwise.

The vector $(N_j)_{j=1, \dots, n}$ defined by

$$N_j := \# \{i \in \{1, \dots, n\}, U_i \in [(j-1)/n, j/n]\},$$

has the $\text{mult}(n, 1/n, \dots, 1/n)$ distribution. The empirical process taken at time k/n for $k \in \{0, \dots, n\}$ is a simple function of this vector :

$$\mathbf{b}_n(k/n) = \sqrt{n} (F_n(k/n) - F(k/n)) = n^{-1/2} \sum_{j=1}^k (N_j - 1). \quad (3)$$

Let $\overline{\mathbf{b}}_n$ be the process obtained by interpolating \mathbf{b}_n between the points $\{k/n, k \in \{0, \dots, n\}\}$.

Let (P_k) be a sequence of i.i.d. Poisson random variables with parameter 1. The distribution of $(P_k)_{k=1, \dots, n}$ under the condition $\sum_{k=1}^n P_k = n$ (or $\sum_{k=1}^n (P_k - 1) = 0$) is also $\text{mult}(n, 1/n, \dots, 1/n)$ as can be straightforwardly

checked. For any $k \in \{0, \dots, n\}$, set

$$\mathbf{S}_k = \sum_{j=1}^k (P_j - 1) \quad (4)$$

and let $\mathbf{S} = (\mathbf{S}_k)_{k=0, \dots, n}$ be the “centered” Poisson random walk, interpolated between integer points. Hence, we have

$$(\overline{\mathbf{b}}_n(t))_{t \in [0,1]} \stackrel{(d)}{=} \left(n^{-1/2} \mathbf{S}_{nt} \right)_{t \in [0,1]} \text{ conditioned by } \mathbf{S}_n = 0, \quad (5)$$

and

$$\sup_{t \in [0,1]} |\overline{\mathbf{b}}_n(t) - \mathbf{b}_n(t)| \leq n^{-1/2} \max_{k=1, \dots, n} N_k. \quad (6)$$

This is controlled as follow: the N_i 's are Binomial($n, 1/n$). By the Markov inequality write $\mathbb{P}(\max_k N_k \geq \varepsilon \sqrt{n}) \leq n \mathbb{P}(N_1 \geq \varepsilon \sqrt{n}) \leq n \mathbb{E}(e^{N_1}) e^{-\varepsilon \sqrt{n}} = n(1 + \frac{e-1}{n})^n e^{-\varepsilon \sqrt{n}} \sim n e^{e-1} e^{-\varepsilon \sqrt{n}} \xrightarrow{n} 0$, and then

$$n^{-1/2} \max_{k=1, \dots, n} N_k \xrightarrow[n]{(proba)} 0. \quad (7)$$

Hence, by (5), (6) and (7), (see also Lemma 7 in Appendix) Theorem 1 stating the convergence of (\mathbf{b}_n) to \mathbf{b} in $D[0, 1]$ is easily implied by the following proposition.

Proposition 2 *The sequence $(n^{-1/2} \mathbf{S}_{nt})_{t \in [0,1]}$ conditioned by $\mathbf{S}_n = 0$ converges in distribution to \mathbf{b} in $C[0, 1]$ equipped with the topology of the uniform convergence.*

The proof we propose for this classical proposition is the real novelty of this paper.

The “correction” of a Poisson random walk

The main line in our approach is the comparison between \mathbf{S} and \mathbf{S} conditioned by $\mathbf{S}_n = 0$. We introduce a correcting process $\mathbf{C} = (\mathbf{C}_k)_{k=0, \dots, n}$, such that the pair (\mathbf{S}, \mathbf{C}) have the following feature :

- \mathbf{S} is the centered Poisson random walk (defined in (4)),
- $\mathbf{S} + \mathbf{C}$ is distributed as \mathbf{S} conditioned by $\mathbf{S}_n = 0$.

Note 2 *Transforming a problem involving n random variables into a problem involving $I_n \sim \text{Poisson}(n)$ random variables is called Poissonization. Taking U_1, \dots, U_{I_n} instead of U_1, \dots, U_n in the construction presented at the beginning of the paper amounts to replacing \mathbf{S} conditioned by $\mathbf{S}_n = 0$ by the centered Poisson random walk \mathbf{S} . This is a Poissonization. The correction of the Poisson random walk we propose, which allows to pass from \mathbf{S} to \mathbf{S} conditioned by $\mathbf{S}_n = 0$ is from our point of view different in nature from the usual depoissonization*

techniques. Here, everything relies on an exact combinatorial correction, when usually, most rely on the convergence in distribution of $(I_n - n)/\sqrt{n}$, ensuring the problem with n variables and with I_n variables being asymptotically equivalent, which is not the case here.

Let us come back to our correction procedure. To fix the details, we will use a classical interpretation of Poisson random walk in term of urns/balls.

Conditionally on $\mathbf{S}_n = s$, the vector $(P_i)_{i=1,\dots,n}$ has the $\text{mult}(s+n, 1/n, \dots, 1/n)$ law. When m balls labeled $1, \dots, m$ are sent independently in n urns according to the uniform distribution, the vector $(N'_i)_{i=1,\dots,n}$ giving the number of balls in the urns follows also the $\text{mult}(m, 1/n, \dots, 1/n)$ distribution.

Let us throw P_i balls in urn i where $(P_i)_{i=1,\dots,n}$ are i.i.d. Poisson random variables with parameter 1. Then three cases arise: $\sum_{i=1}^n P_i = n$ or $\sum_{i=1}^n P_i < n$, or $\sum_{i=1}^n P_i > n$ (recall that $\mathbf{S}_n = \sum_{i=1}^n P_i - n$).

In the first case $\mathbf{S}_n = 0$, no correction are necessary, then set $\mathbf{C}_i = 0$ for any i . The two last cases are treated below. Notice that we focus on the uni-dimensional distributions of the process \mathbf{C} since this will appear to be sufficient.

Case $\mathbf{S}_n < 0$. We work conditionally on $\mathbf{S}_n = s$. Since $-s$ balls are lacking: throw $-s$ new balls and denote by \mathbf{C}_k the number of new balls fallen in the k first urns; for any k ,

$$\mathbf{C}_k \sim \text{Binomial}(-s, k/n). \quad (8)$$

More precisely, $(\Delta \mathbf{C}_k)_{k=1,\dots,n} \sim \text{mult}(-s, 1/n, \dots, 1/n)$ where $\Delta \mathbf{C}_k := \mathbf{C}_k - \mathbf{C}_{k-1}$ is the k th increment of the correcting process \mathbf{C} (with $\mathbf{C}_0 = 0$).

Lemma 3 For any $s < 0$ and any $n \geq 1$, conditionally on $\mathbf{S}_n = s$ the process $\mathbf{S} + \mathbf{C}$ is distributed as \mathbf{S} conditioned by $\mathbf{S}_n = 0$ and $\mathbf{C}_k \sim \text{Binomial}(-s, k/n)$.

Proof. If $X \sim \text{mult}(n+s, 1/n, \dots, 1/n)$, $Y \sim (-s, 1/n, \dots, 1/n)$ and X and Y are independent then $X + Y \sim \text{mult}(n, 1/n, \dots, 1/n)$. \square

Case $\mathbf{S}_n > 0$. We work conditionally on $\mathbf{S}_n = s$. In this case $n+s$ balls have been thrown instead of n and then s balls must be taken out. The vector $(V_k)_{k=1,\dots,n}$ giving the exceeding number of balls in the different urns (those with labels in $n+1, \dots, n+s$) follows the law $\text{mult}(s, 1/n, \dots, 1/n)$. Then given $\mathbf{S}_n = s$, we search a correcting process $(\Delta \mathbf{C}_k)_{k=1,\dots,n} \stackrel{(d)}{=} (-V_k)_{k=1,\dots,n}$. Of course there is a problem to define the correcting process in terms of balls/urns, the balls/urns problem living a priori on a larger probability space than the (P_i) 's. But this gives us the intuition for a right correcting process: we define \mathbf{C} conditionally on the P_i 's as follows. Let $(p_i)_{i=1,\dots,n}$ be non negative integers summing to $n+s$. Set

$$\mathbb{P}(\Delta \mathbf{C}_k = -c_k, k \in \{1, \dots, n\} | P_i = p_i, i \in \{1, \dots, n\}) = \frac{\prod_{i=1}^n \binom{p_i}{c_i} \mathbb{1}_{c_i \leq p_i}}{\binom{\sum_s p_i}{s}} \quad (9)$$

for any given non negative integers c_1, \dots, c_n summing to s , and 0 otherwise.

Lemma 4 For any $s > 0$ and any $n \geq 1$, conditionally on $\mathbf{S}_n = s$ the process $\mathbf{S} + \mathbf{C}$ is distributed as \mathbf{S} conditioned by $\mathbf{S}_n = 0$ and $\mathbf{C}_k \sim -\text{Binomial}(s, k/n)$.

Proof. We have to check that $\mathbf{C} + \mathbf{S}$ is distributed as \mathbf{S} conditioned by $\mathbf{S}_n = 0$: $\mathbb{P}(P_i + \Delta \mathbf{C}_i = j_i, \forall i | \mathbf{S}_n = s)$

$$\begin{aligned} &= \sum_{(p_i) : \sum p_i = n+s, p_i \geq j_i} \frac{\mathbb{P}(\Delta \mathbf{C}_i = -(p_i - j_i), \forall i | P_i = p_i, \forall i) \mathbb{P}(P_i = p_i, \forall i)}{\mathbb{P}(\mathbf{S}_n = s)} \\ &= \sum_{(p_i - j_i) : p_i - j_i \geq 0, \sum p_i - j_i = s} \frac{\prod_{i=1}^n \binom{p_i}{p_i - j_i} e^{-n} \prod_{i=1}^n \frac{1}{p_i!}}{\binom{n+s}{s} \frac{e^{-n} n^{n+s}}{(n+s)!}} \\ &= \frac{e^{-n} \prod_{i=1}^n \frac{1}{j_i!}}{e^{-n} n^n / n!} = \mathbb{P}(P_i = j_i, \forall i | \mathbf{S}_n = 0) \end{aligned}$$

where we have used (9), the fact that $n + \mathbf{S}_n$ is Poisson(n) distributed, and

$$\sum_{(\alpha_i) : \alpha_i \geq 0, \sum \alpha_i = s} \prod_{i=1}^n \frac{s!}{\alpha_i!} = (1 + \dots + 1)^s = n^s. \quad (10)$$

We now show that knowing $\mathbf{S}_n = s$, $(-\Delta \mathbf{C}_k)_{k=1, \dots, n} \sim \text{mult}(s, 1/n, \dots, 1/n)$. This implies the second point. Let c_1, \dots, c_n be non negative integers summing to s . Write $\mathbb{P}(\Delta \mathbf{C}_k = -c_k, \forall k | \mathbf{S}_n = s)$

$$= \sum_{(p_i), p_i \geq c_i, \sum p_i = n+s} \frac{\mathbb{P}(\Delta \mathbf{C}_k = -c_k, \forall k | P_i = p_i, \forall i) \mathbb{P}(P_i = p_i, \forall i)}{\mathbb{P}(\mathbf{S}_n = s)}.$$

By (10), $\mathbf{S}_n + n \sim \text{Poisson}(n)$ and (9), this is easily shown to be equal to $\frac{s!}{\prod_{i=1}^n c_i!} \frac{1}{n^s}$. \square

From now on, consider the process \mathbf{C} as being interpolated between integer points.

Lemma 5 For any $t \in [0, 1]$,

$$n^{-1/2} |\mathbf{C}_{nt} + t\mathbf{S}_n| \xrightarrow[n]{(proba)} 0. \quad (11)$$

Proof. We work with $C_{\lfloor nt \rfloor}$ instead of C_{nt} . For $t = 0$, (11) holds clearly. Let $t \in (0, 1]$, and $\varepsilon > 0$ be fixed. Write

$$\mathbb{P}(|\mathbf{C}_{\lfloor nt \rfloor} + t\mathbf{S}_n| \geq \varepsilon \sqrt{n}) \leq V_n^M + W_n^M$$

where

$$\begin{aligned} V_n^M &= \mathbb{P}(|\mathbf{C}_{\lfloor nt \rfloor} + t\mathbf{S}_n| \geq \varepsilon \sqrt{n}, |\mathbf{S}_n| \in \sqrt{n}[M^{-1}, M]), \\ W_n^M &= \mathbb{P}(|\mathbf{S}_n| \notin \sqrt{n}[M^{-1}, M]). \end{aligned}$$

Let $\alpha > 0$ be a fixed (small) positive real number. The central limit theorem applied to \mathbf{S}_n ensures that there exists M such that $W_n^M \leq \alpha$, for n large enough. Fix this M . To bound V_n^M , use $\mathbb{P}(A | \cup_i B_i) = \sum \mathbb{P}(A|B_i)\mathbb{P}(B_i)/P(\cup B_i) \leq \max_i \mathbb{P}(A|B_i)$ for disjoint sets B_i . Hence

$$V_n^M \leq \max_{k, |k| \in \sqrt{n}[M^{-1}, M]} \mathbb{P}(|\mathbf{C}_{[nt]} + t\mathbf{S}_n| \geq \varepsilon\sqrt{n} \mid |\mathbf{S}_n| = k).$$

By Lemmas 3 and 4, $\mathbb{P}(|\mathbf{C}_{[nt]} + t\mathbf{S}_n| \geq \varepsilon\sqrt{n} \mid |\mathbf{S}_n| = k) = \mathbb{P}(|B(k, [nt]/n) - tk| \geq \varepsilon\sqrt{n})$ where $B(k, [nt]/n)$ is a binomial random variable with parameters k and $[nt]/n$. Further, by the Bienaymé-Tchebichev inequality

$$\max_{k, |k| \in \sqrt{n}[M^{-1}, M]} \mathbb{P}(|B(k, [nt]/n) - tk| \geq \varepsilon\sqrt{n}) \rightarrow 0. \quad \square$$

Proposition 6 (i) *The following convergence holds in $C([0, 1], \mathbb{R}^2)$:*

$$n^{-1/2}(\mathbf{S}_{nt}, \mathbf{C}_{nt})_{t \in [0, 1]} \xrightarrow[n]{(d)} (\mathbf{B}_t, -t\mathbf{B}_1)_{t \in [0, 1]}.$$

(ii) *The following convergence holds in $C[0, 1]$:*

$$n^{-1/2}(\mathbf{S}_{nt} + \mathbf{C}_{nt})_{t \in [0, 1]} \xrightarrow[n]{(d)} (\mathbf{B}_t - t\mathbf{B}_1)_{t \in [0, 1]}.$$

Proposition 2 is a consequence of (ii) thanks to Lemmas 3 and 4.

Proof. Assertion (ii) is a consequence of (i). Proof of (i) : the convergence of $n^{-1/2}\mathbf{S}_n$ to \mathbf{B} in $C[0, 1]$ is given by the other famous Donsker's theorem stating the convergence of rescaled random walks to the Brownian motion (see [2] or [4]). In particular

$$n^{-1/2}\mathbf{S}_n \rightarrow \mathbf{B}_1. \quad (12)$$

The finite dimensional distribution of $n^{-1/2}\mathbf{C}_n$ converges to those of the process $(t\mathbf{B}_1)_{t \in [0, 1]}$. Indeed by Lemma 5 and (12),

$$n^{-1/2}(\mathbf{S}_n, \mathbf{C}_{nt_1}, \dots, \mathbf{C}_{nt_k}) \xrightarrow[n]{(d)} (\mathbf{B}_1, -t_1\mathbf{B}_1, \dots, -t_k\mathbf{B}_1)$$

for any $0 \leq t_1 \leq \dots \leq t_k \leq 1$. Then the family $(n^{-1/2}\mathbf{C}_n)$ is tight since it is a sequence of monotone processes whose finite dimensional distribution converge to those of the a.s. continuous process $(t\mathbf{B}_1)_{t \in [0, 1]}$ (this is Lemma 8(ii)). Hence the family $n^{-1/2}(\mathbf{S}_n, \mathbf{C}_n)$ is tight. The limit is identified again thanks to Lemma 5. \square

Conclusion

The idea of this proof appeared after a discussion with Philippe Duchon, where he explained me his algorithm to generate uniformly a Bernoulli bridge with $2n$ steps, that is a random walk $\mathbf{S} = (\mathbf{S}_k)_{k=0, \dots, 2n}$ with increments ± 1 , conditioned by $\mathbf{S}_{2n} = 0$: build first a simple random walk with $2n$ steps, choosing i.i.d.

increments $+1$, or -1 with probability $1/2$. If $\mathbf{S}_{2n} = 0$ then it's done. If not, assume that $\mathbf{S}_{2n} = 2k > 0$. Then pick up at random indices I_1, I_2, \dots in $\llbracket 1, 2n \rrbracket$. If I_i is the index of a positive increment, change it into a negative one; if it is negative then do nothing. Stop when you have changed k increments. By a simple symmetry argument the path obtained is uniform in the set of Bernoulli bridges of size $2n$. I found that this was a nice way to prove that rescaled Bernoulli bridge converges to the Brownian bridge; this can be proved using the same argument than the ones exposed above: the correction procedure will asymptotically and "eventually removes a straight line of the Brownian motion". Therefore, I tried to find other increment distributions for which a similar correction procedure would have been possible. It appears to be not so general, or at least, not so agreeable. The problem is the following one: in general there does not exist any simple correction procedure that conserves at each step of the correction the property of the trajectory to have conditionally on its terminal position k , the law of a simple random walk conditioned by $\mathbf{S}_n = k$.

Appendix

A simple link between the convergence in $C[0, 1]$ and in $D[0, 1]$

Lemma 7 *Let (X_n) be a sequence of processes taking their values in $D[0, 1]$. Assume that for any n , $X_n = Y_n + Z_n$ where Y_n is a continuous process, and Z_n is a càdlag process. If (Y_n) converges in distribution to Y in $C[0, 1]$, and if $\sup |Z_n| \xrightarrow[n]{(d)} 0$ then (X_n) converges in distribution to Y in $D[0, 1]$.*

Proof. As a matter of fact, this statement is easy if one knows:

(a) $(Y_n) \xrightarrow[n]{(d)} Y$ in $C[0, 1]$ implies $(Y_n) \xrightarrow[n]{(d)} Y$ in $D[0, 1]$,

(b) (Z_n) càdlag, $\sup |Z_n| \xrightarrow[n]{(d)} 0$ implies $Z_n \xrightarrow[n]{(d)} 0$ (the null process) on $D[0, 1]$.

Indeed, knowing this, Lemma 7 is a consequence of the following classical result: let A_n, B_n, C_n be random variables in a metric space (S, ρ) . If $A_n = B_n + C_n$, $A_n \xrightarrow[n]{(d)} A$, $\rho(B_n) \xrightarrow[n]{(d)} 0$ then $A_n + B_n \xrightarrow[n]{(d)} A$ (see e.g. Billingsley [2, Theorem 3.1]).

In order to understand why (a) and (b) hold true, recall that the Skorohod topology on $D[0, 1]$ is defined by a metric: let Λ be the set of strictly increasing continuous functions λ , satisfying $\lambda(0) = 0, \lambda(1) = 1$. The metric is

$$d_s(f, g) = \inf_{\lambda \in \Lambda} \left\{ \sup\{|\lambda(t) - t|, t \in [0, 1]\} \vee \sup\{|f(\lambda(t)) - g(t)|, t \in [0, 1]\} \right\}.$$

Hence $d_s(f, g) \leq \sup\{|f(t) - g(t)|, t \in [0, 1]\}$; this yields immediately to (a) and (b). \square

Convergence in $C[0, 1]$ and in $C[0, 1]^2$

We recall some classical facts concerning the weak convergence in $C[0, 1]$ and $(C[0, 1])^2$. First tightness and relative compactness are equivalent in these sets by Prohorov's theorem, since they are both Polish spaces.

Lemma 8 (i) Let (X_n, Y_n) be a sequence of pairs of processes in $(C[0, 1])^2$. The tightnesses of both families (X_n) and (Y_n) imply that of (X_n, Y_n) .

(ii) Let (X_n) be a sequence of monotone processes in $C[0, 1]$. If the finite dimensional distributions of (X_n) converge to those of an a.s. continuous process X , then (X_n) is tight and then $X_n \xrightarrow[n]{(d)} X$ in $C[0, 1]$.

Proof. (i) Take two compacts K_1 and K_2 of $C[0, 1]$ such that $\mathbb{P}(X_n \in K_1) \geq 1 - \varepsilon$ and $\mathbb{P}(Y_n \in K_2) \geq 1 - \varepsilon$, then $\mathbb{P}((X_n, Y_n) \in K_1 \times K_2) \geq 1 - 2\varepsilon$ and $K_1 \times K_2$ is compact in $(C[0, 1])^2$.

(ii) Only the tightness of (X_n) in $C[0, 1]$ has to be checked. For any function $f : [0, 1] \rightarrow \mathbb{R}$, and $\delta > 0$, the global modulus of continuity of f is

$$\omega_\delta(f) = \sup\{|f(x) - f(y)|, x, y \in [0, 1], |x - y| \leq \delta\}.$$

Since X_n is increasing, for any positive integer m ,

$$\omega_{1/m}(X_n) \leq A_{m,n} := 2 \max \left\{ \left| X_n\left(\frac{k}{m}\right) - X_n\left(\frac{k-1}{m}\right) \right|, k = 1, \dots, m \right\}.$$

Since the finite dimensional distributions of (X_n) converge to those of X ,

$$A_{m,n} \xrightarrow[n]{(d)} A_m := 2 \max \left\{ \left| X\left(\frac{k}{m}\right) - X\left(\frac{k-1}{m}\right) \right|, k = 1, \dots, m \right\}$$

and by the uniform continuity of X , $A_m \xrightarrow[m]{proba.} 0$. Hence $\lim_m \limsup_n \mathbb{P}(\omega_{1/m}(X_n) \geq \varepsilon) = 0$ for any $\varepsilon > 0$. \square

Acknowledgments

I thanks Philippe Duchon who gives me the inspiration of this approach. I thank also Bernard Bercu for his comments on a preliminary version of this text.

References

- [1] M.D. Donsker, (1952) *Justification and extension of Doob's heuristic approach to the Kolmogorov-Smirnov theorems*, Annals of Mathematical Statistics., 23:277–281, 1952
- [2] P. Billingsley, (1968) *Convergence of Probability measure*, John Wiley and Sons, New York, third edition.

- [3] J. L. Doob, (1949) *Heuristic approach to the Kolmogorov-Smirnov theorems*, Ann. Math. Statistics 20, 393–403.
- [4] O. Kallenberg (1997), *Foundations of Modern Probability*. Probability and Its Applications. Springer, New York, NY.
- [5] J. Komlós, P. Major, G. Tusnády, (1975) *An approximation of partial sums of independent RV's and the sample DF*, I. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 32, 111–131.
- [6] G.R. Shorack, J.A. Wellner, (1986) *Empirical processes with applications to statistics*, Wiley Series in Probab. and Math. Stat.: Probab. and Math. Stat. John Wiley & Sons, Inc., New York.
- [7] A.W. van der Vaart, J.A. Wellner, Jon A. (1996) *Weak convergence and empirical processes. With applications to statistics*. Springer Series in Stat. Springer-Verlag, New York.