Electronic Journal of Probability

Vol. 4 (1999) Paper no. 14, pages 1–25.

Journal URL
http://www.math.washington.edu/~ejpecp
Paper URL
http://www.math.washington.edu/~ejpecp/EjpVol4/paper14.abs.html

RANDOM TIME CHANGES FOR SOCK-SORTING AND OTHER STOCHASTIC PROCESS LIMIT THEOREMS

David Steinsaltz

Department of Statistics, University of California, Berkeley CA 94720
E-mail: dstein@stat.berkeley.edu

Abstract A common technique in the theory of stochastic process is to replace a discrete time coordinate by a continuous randomized time, defined by an independent Poisson or other process. Once the analysis is complete on this Poissonized process, translating the results back to the original setting may be nontrivial. It is shown here that, under fairly general conditions, if the process S_n and the time change ϕ_n both converge, when normalized by the same constant α_n, to limit processes $\widetilde{S}$ and $\widetilde{\Phi}$, then the combined process $S_n \circ \phi_n$ converges to $\widetilde{S} + \widetilde{\Phi} \cdot \frac{d}{dt} \mathrm{E}\, S(t)$ when properly normalized. It is also shown that earlier results on the fine structure of the maxima are preserved by these time changes.

The remainder of the paper then applies these simple results to processes which arise in a natural way from sorting procedures, and from random allocations. The first example is a generalization of "sock-sorting": Given a pile of n mixed-up pairs of socks, we draw out one at a time, laying it on a table if its partner has not yet been drawn, and putting completed pairs away. The question is: What is the distribution of the maximum number of socks ever on the table, for large n? Similarly, when randomly throwing balls into n (a large number) boxes, we examine the distribution of the maximum over all times of the number of boxes that have (for example) exactly one ball.

Keywords maximal inequalities, Poissonization, functional central limit theorem, sorting, random allocations, auxiliary randomization, time change

AMS subject classification Primary: 60F17; Secondary: 60G70,60K30.

Submitted to EJP on March 16, 1998. Final version accepted on May 20, 1999.

1. Introduction

The following home-economics problem has reappeared in various guises over the years ([Ber82], [Lut88], [Ste96], and [LP98]): N pairs of freshly washed socks, each pair having a distinguishing color or pattern, lie thoroughly mixed in a bin. I draw them out one by one, at random, with the object of sorting them into pairs. As each new sock is drawn out in turn, I lay it on my sorting table; if the new sock matches one already on the table, I fold the pair neatly away into my capacious sock drawer. After n draws, what is the expected number of socks on the table?

Since that question was seen off already in the mid-18th century by Daniel Bernoulli [Ber82], a more delicate question suggests itself: How much space do I need on the table? In other words, if s_n is the number of unmatched socks after n draws, what can I say about the distribution of $\max_{0\leq n\leq 2N} s_n$? More generally, I will consider the sorting of N types, each with a random number of elements.

The trick which I will apply here, hardly an uncommon one in probability theory, is a random time-change. (This approach to the sock-sorting problem first appeared in my dissertation [Ste96], but it has since been applied to the same problem by W. Li and G. Pritchard [LP98], apparently independently. For more examples of this sort of embedding, and further references, see [BH91], [Lan96], and [Kin93, pages 14–17].) Concretely, s_n may be represented

$$s_n = \sum_{i=1}^{N} s_n^i \,,$$

where s_n^i is 1 if precisely one sock of type i is included in the first n draws, and 0 otherwise. For convenience, rescale the whole process to occur on the time interval $[0,1]$:

$$\begin{aligned} S_N(t) &= \frac{1}{N} s_{[2Nt]} \,; \\ S_N^i(t) &= S_{[2Nt]}^i \,. \end{aligned}$$

Essentially, what this does is to let the n-th draw occur at time $n/2N$.

This is only mindless rescaling, but it opens the way for an almost magical simplification when this deterministic time-change is replaced by a random one, where each sock is drawn at an independent random time, uniform on $[0,1]$. This i.i.d. randomization of the time automatically generates a uniform permutation of the socks. Let $\phi_N(t)$ be the number of socks drawn up to time t, divided by $2N$; then

$$F_N(t) := \frac{1}{N} s_{2N\phi_N(t)} = \frac{1}{N}\sum_{i=1}^{N} f_i(t) \,,$$

where $f_i(t)$ is defined to be 1 if precisely one of the socks of type i has been drawn by time t and 0 otherwise. The advantage of this representation is that, since each sock pair is concerned only with its own clock, the random functions f_i are independent. This opens up the problem to powerful empirical-process techniques, as described in [Ste96] and [Ste98], to study the limiting behavior of F_N.

It remains, though, to consider how the information that carries over from this earlier analysis is affected by the time change. One result in [Ste98] says that

$$\widetilde{F}_N(t) := N^{1/2}\big(F_N(t) - \mathrm{E}\, F_N(t)\big)$$

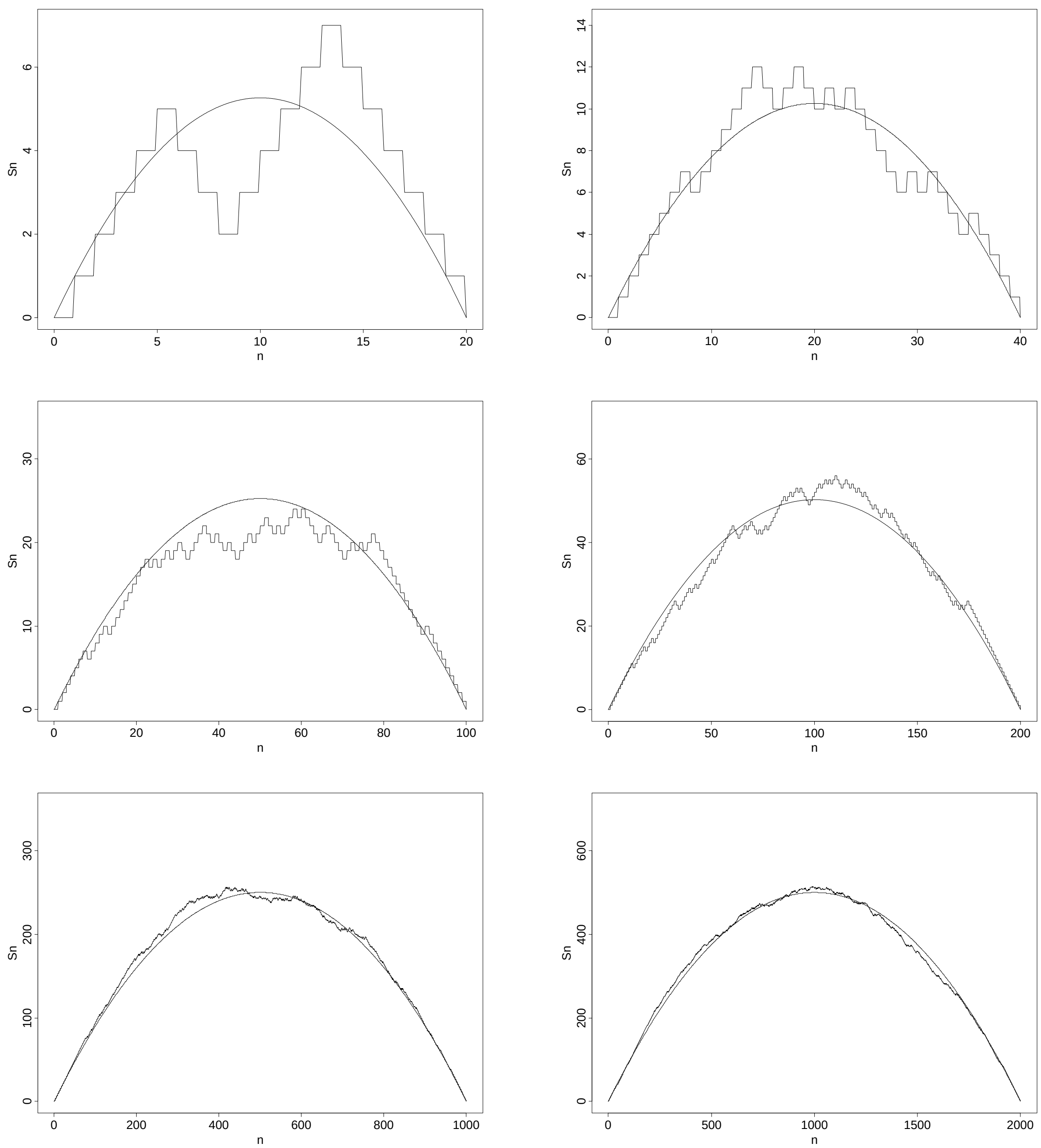

FIGURE 1. Simulations of the socks process $S_N(t)$ with $N = 10, 20, 50, 100, 500, 1000$. The smooth curve in the picture represents the expectation, $\mathrm{E}\big[S_N(t)\big] = 2Nt(1-t)$.

converges weakly to a Gaussian process. Since $\phi_N(t)$ converges uniformly in probability to the identity function, a functional central limit theorem for

$$\widetilde{S}_N(t) = N^{1/2}\big(s_{[2Nt]} - \mathrm{E}\, s_{[2Nt]}\big)$$

would imply a similar theorem for

$$N^{1/2}\Big(F_N(t) - \overline{S}_N\left(\phi_N(t)\right)\Big) = \widetilde{S}_N\left(\phi_N(t)\right).$$

(Here, and occasionally elsewhere, when $X(t)$ is a random function the expectation function $\mathrm{E}\,X(t)$ is represented as $\overline{X}_N(t)$ in order to allow a distinction between $\mathrm{E}\big[X(\phi_N(t))\big]$ and $\overline{X}(\phi_N(t))$; that is, whether or not to integrate over the time change.) The problem here is, first, to move in the opposite direction, from a limit theorem for $\widetilde{F}_N$ to a corresponding one for $\widetilde{S}_N$ (where F_N and ϕ_N are not independent), and second, to replace $\overline{S}_N\big(\phi_N(t)\big)$ by $\overline{S}_N(t)$. Proposition 3.1 states that, asymptotically, $\widetilde{F}_N(t)$ may be neatly decomposed into a sum of two independent Gaussian processes, one which is $\widetilde{S}_N(t)$, and the other corresponding to the fluctuations in the time change itself.

Another result in the earlier paper says that the pair

$$\text{(1)} \qquad \Big(N^{2/3}\big(\max F_N(t) - F_N(\tfrac{1}{2})\big)\,,\, N^{1/3}\big(\arg\max F_N - \tfrac{1}{2}\big)\Big)$$

converges to the pair consisting of the maximum and the location of the maximum of $B_t - 2t^2$, where B_t is a standard Brownian motion.* To put it another way, the maximum may be divided up on three different scales:

$$\text{(2)} \qquad \begin{aligned}\max_t F_N(t) = \frac{1}{2} \;+\; N^{-1/2}\cdot\Big\{N^{1/2}\big(F_N(\tfrac{1}{2}) - \tfrac{1}{2}\big)\Big\} \\ +\; N^{-2/3}\cdot\Big\{N^{2/3}\big(\max F_N(t) - F_N(\tfrac{1}{2})\big)\Big\}\,,\end{aligned}$$

where the terms in braces converge weakly to independent finite-valued random variables.

The distribution of the maximum is unaffected by the time change. Thus, it would certainly be true that

$$\left(N^{-1/3}\big(\max_{0\le n\le 2N} s_n - s_{[2N\phi_N(1/2)]}\big)\,,\, N^{1/3}\Big(\phi_N\left(\frac{\arg\max S}{2N}\right) - \frac{1}{2}\Big)\right).$$

satisfies the same limit law as (1). It would be pleasant to replace this by

$$\Big(N^{-1/3}\big(\max s_n - s_N\big)\,,\, N^{1/3}\big(\frac{\arg\max S}{2N} - \frac{1}{2}\big)\Big).$$

Since the empirical process $\phi_N(t) - t$ is uniformly $O_p(N^{-1/2})$, there is no problem in eliminating ϕ_N from the second term. On the other hand, $s_{[2N\phi_N(1/2)]}$ can differ from s_N by as much as $2N(\phi_N\left(\frac{1}{2}\right) - \frac{1}{2})$, which should be on the order of $N^{1/2}$, swamping the $N^{-1/3}$ normalization. What saves the situation, as I show in Proposition 3.2, is that it occurs at a relative extremum of the expectation. This means that, in these $O(N^{1/2})$ steps between $\frac{1}{2}$ and $\phi_N(\frac{1}{2})$ the process moves like a symmetric random walk, and on average only fluctuates by about $N^{1/4}$, not $N^{1/2}$.

Section 4.1 applies these results to a generalized sorting problem. The procedure is the one described above for socks, only now instead of pairs I admit classes of arbitrary (random) numbers of objects. I show there that both limit theorems hold — the functional central limit theorem as well as the second-order approximation for the maximum (corresponding to (2)) — as long as the class sizes have bounded fourth moments, and their distributions of the class sizes converge in $\mathfrak{L}^4$ when properly normalized.

*For definiteness, for f a cadlag function defined on a real interval $\mathcal{I}$, define $\arg\max f = \inf\big\{t : \sup_{s\le t} f(s) = \sup_{s\in\mathcal{I}} f(s)\big\}$.

In section 4.2 I apply these same methods to a related question about random allocations. Here balls are being thrown successively at random into a large number N of boxes, and I consider the behavior of such functions as the total number of boxes which have exactly one ball, after n have been thrown.

2. Technical lemmas

In what follows, $\mathcal{I}$ will be a subinterval of $\mathbb{R}$ and $\mathcal{D} = \mathcal{D}(\mathcal{I})$ will represent the space of cadlag (right-continuous with limits from the left) functions from $\mathcal{I}$ to $\mathbb{R}$. I will have little to say concretely about the topology, but assume it to be furnished with the Skorokhod topology (cf. section 14 of [Bil68]), and all measurability questions will be referred to the corresponding Borel σ-algebra. The modulus of continuity of a function x is denoted by

$$w_x(\delta) = \sup_{|s-t|\le\delta} \bigl|x(s) - x(t)\bigr|.$$

For nondecreasing functions $\varphi \in \mathcal{D}(\mathcal{I})$, a cadlag pseudoinverse function is defined by

$$\varphi^{-1}(t) = \begin{cases} \inf\{s : \varphi(s) > t\} & \text{if } \inf \operatorname{Range} \varphi \le t < \sup \operatorname{Range} \varphi, \\ \inf \mathcal{I} & \text{if } t < \inf \operatorname{Range} \varphi, \\ \sup \mathcal{I} & \text{if } t \ge \sup \operatorname{Range} \varphi. \end{cases} \tag{3}$$

An elementary fact about pseudoinverses is

Lemma 2.1. *Let φ and ϕ be monotonically nondecreasing functions in $\mathcal{D}[0,1]$, with $\phi(0) = \varphi(0)$. Then*

$$\sup_{t\in\mathbb{R}} \bigl|\varphi^{-1}(t) - \phi^{-1}(t)\bigr| \le \omega_{\phi^{-1}}\Bigl(\sup_{s\in\mathcal{I}} \bigl|\varphi(s) - \phi(s)\bigr| + \sup_{t\in R} \bigl|\varphi\bigl(\varphi^{-1}(t)\bigr) - t\bigr|\Bigr), \tag{4}$$

where $R = \operatorname{Range} \phi \cap \operatorname{Range} \varphi$.

Proof. The difference $\bigl|\varphi^{-1}(t) - \phi^{-1}(t)\bigr|$ attains its maximum on R. Consider then any $t \in R$.

$$\begin{aligned} \bigl|\phi\bigl(\varphi^{-1}(t)\bigr) - t\bigr| &\le \bigl|\phi\bigl(\varphi^{-1}(t)\bigr) - \varphi\bigl(\varphi^{-1}(t)\bigr)\bigr| + \bigl|\varphi\bigl(\varphi^{-1}(t)\bigr) - t\bigr| \\ &\le \sup_{s\in\mathcal{I}} \bigl|\phi(s) - \varphi(s)\bigr| + \sup_{t\in R} \bigl|\varphi\bigl(\varphi^{-1}(t)\bigr) - t\bigr|. \end{aligned}$$

Applying ϕ^{-1} to the two points on the left side gives the result, by the definition of modulus of continuity. □

We will also need some facts about uniform tightness which, while not original, do not seem to be explicitly stated in standard reference works. Let (x_n) be a sequence of random functions in $\mathcal{D}(\mathcal{I})$, for some finite interval $\mathcal{I}$. I will say that (x_n) is *tight in the uniform modulus* if for some $t_0 \in \mathcal{I}$ the random variables $x_n(t_0)$ are uniformly tight, and for every positive ϵ,

$$\lim_{\delta\to 0} \limsup_{n\to\infty} \mathrm{P}\bigl\{w_{x_n}(\delta) > \epsilon\bigr\} = 0.$$

Lemma 2.2. *Let (x_n) and (y_n) be two sequences of random functions (on the same probability space) in $\mathcal{D}(\mathcal{I})$.*

1. *If (x_n) converges weakly to a process that is almost surely continuous, then it is tight in the uniform modulus. Conversely, if (x_n) is tight in the uniform modulus, then it is tight as a sequence of random functions in $\mathcal{D}$, and if it converges weakly the limit law must be almost surely continuous.*

2. *If (x_n) is tight in the uniform modulus then the sequence of random variables $(\sup_{t\in\mathcal{I}}|x_n(t)|)$ is tight.*
3. *If (x_n) and (y_n) both are tight in the uniform modulus, then (x_n+y_n) is as well.*
4. *If $\|x_n-y_n\|_\infty \xrightarrow[P]{n\to\infty} 0$, then if (x_n) is tight in the uniform modulus, (y_n) is as well.*

Proof. The first assertion in 1 is a consequence of the Continuous Mapping Theorem (cf. Theorem IV.12 of [Pol84]), and the fact that the modulus of continuity, seen as a functional of cadlag functions x (with δ fixed), is continuous at continuous x. The second assertion appears as Theorem 15.5 of [Bil68].

For every positive K,

$$\mathrm{P}\big\{\sup_{t\in\mathcal{I}}\big|x_n(t)\big|>K\big\}\le \mathrm{P}\big\{\big|x_n(t_0)\big|>\frac{K}{2}\big\}+\mathrm{P}\big\{w_{x_n}\big(\frac{2|\mathcal{I}|}{K}\big)>1\big\},$$

which, by taking appropriate limits, proves assertion 2.

Assertion 3 follows from the subadditivity of the continuity modulus. Assertion 4 follows from assertion 3, since $y_n=x_n+(y_n-x_n)$, and y_n-x_n is tight in the uniform modulus, since it converges in probability to 0. □

Lemma 2.3. *Let (ϕ_n) be a sequence of random elements of $\mathcal{D}(\mathcal{I})$, which are each almost surely nondecreasing. Let α_n be an increasing sequence of real numbers that go to ∞ with n, and ϕ a continuous real-valued function on $\mathcal{I}$, such that ϕ^{-1} is twice differentiable with bounded second derivative. Suppose that $\widetilde{\phi}_n=\alpha_n(\phi_n-\phi)$ is tight in the unifom modulus. Then $\widetilde{\phi_n^{-1}}=\alpha_n(\phi_n^{-1}-\phi^{-1})$ also is tight in the uniform modulus.*

Proof. Let R_n represent the range of ϕ_n, and let

$$\rho_n=\sup_{t\in R_n}\big|\phi_n\big(\phi_n^{-1}(t)\big)-t\big|.$$

It is immediate from the definition of ϕ_n^{-1} that ρ_n is bounded by the size of the largest jump by ϕ_n. It follows that for every positive δ, $w_{\widetilde{\phi}_n}(\delta)\ge\alpha_n\rho_n$. The tightness condition for $\widetilde{\phi}_n$ implies then that $\alpha_n\rho_n$ converges to 0 in probability, as n goes to ∞. Define also $\sigma_n=\sup_{t\in\mathcal{I}}|\phi_n(t)-\phi(t)|$. Since $\widetilde{\phi}_n$ is tight in the uniform modulus, the sequence of random variables $(\alpha_n\sigma_n)$ is tight. By Lemma 2.1, for any fixed t, $(|\widetilde{\phi_n^{-1}}(t)|)_{n=1}^\infty$ is tight as well.

Let L_1 and L_2 be upper bounds for the absolute values of the first and second derivatives of ϕ^{-1} respectively. Let $\delta>0$ be given, and consider any $s,t\in R_n$ with $|s-t|<\delta$. Let $u=\phi_n^{-1}(t)$, $v=\phi_n^{-1}(s)$. By successive Taylor expansions of ϕ^{-1},

$$\begin{aligned}\big|\widetilde{\phi_n^{-1}}(t)-\widetilde{\phi_n^{-1}}(s)\big| &= \alpha_n\big|\phi^{-1}\big(\phi(u)\big)+\phi^{-1}\big(\phi(v)\big)-\big(\phi^{-1}(t)-\phi^{-1}(s)\big)\big|\\ &\le L_1\alpha_n\big|(t-s)-\big(\phi(u)-\phi(v)\big)\big|\\ &\qquad + L_2\alpha_n\cdot\Big[\big|\phi(u)-t\big|^2+\big|\phi(v)-s\big|^2+\big|s-t\big|\cdot\big|\phi(u)-t\big|\Big]\\ &\le L_1\big|\widetilde{\phi}_n(u)-\widetilde{\phi}_n(v)\big|+2L_1\alpha_n\rho_n\\ &\qquad + L_2\alpha_n\big[2(\sigma_n+\rho_n)^2+\delta(\sigma_n+\rho_n)\big]\,.\end{aligned}$$

At the same time, using Lemma 2.1 and the fact that $w_{\phi^{-1}}(\delta) \le L_1\delta$

$$
\begin{aligned}
|u-v| &\le \bigl|\phi^{-1}(s)-\phi^{-1}(t)\bigr| + \bigl|\phi^{-1}(t)-\phi_n^{-1}(t)\bigr| + \bigl|\phi^{-1}(s)-\phi_n^{-1}(s)\bigr| \\
&\le L_1\bigl(\delta + 2L_1(\sigma_n+\rho_n)\bigr).
\end{aligned}
$$

Thus for every positive ϵ,

$$
\begin{aligned}
&\lim_{\delta\to 0}\limsup_{n\to\infty} \mathrm{P}\bigl\{w_{\widetilde{\phi_n^{-1}}}(2L_1\delta) > \epsilon\bigr\} \\
&\qquad \le \lim_{\delta\to 0}\limsup_{n\to\infty}\Bigl(\mathrm{P}\bigl\{w_{\widetilde{\phi}_n}(2\delta) > \frac{\epsilon}{4L_1}\bigr\} + \mathrm{P}\bigl\{2L_1(\sigma_n+\rho_n) > \delta\bigr\} \\
&\qquad\qquad + \mathrm{P}\bigl\{\alpha_n\rho_n > \frac{\epsilon}{8L_1}\bigr\} \\
&\qquad\qquad\quad + \mathrm{P}\bigl\{\alpha_n^2(\sigma_n+\rho_n)^2 > \frac{\epsilon}{8L_2}\alpha_n\bigr\} \\
&\qquad\qquad\qquad + \mathrm{P}\bigl\{\alpha_n(\sigma_n+\rho_n) > \frac{\epsilon}{4L_2\delta}\bigr\}\Bigr).
\end{aligned}
$$

Remember that $\alpha_n\rho_n \xrightarrow{n\to\infty}_P 0$, while the sequence $(\alpha_n\sigma_n)$ is tight. This implies immediately that the second, third, and fourth terms on the right go to 0 as n goes to ∞, and the fifth term goes to zero when at last δ goes to 0. The first term on the right, meanwhile, is 0, because of the assumption that $\widetilde{\phi}_n$ is tight in the uniform modulus. □

Note: Lemma 2.3 would not be true if the condition on the second derivative of ϕ^{-1} were removed. A particularly simple counterexample, with ϕ^{-1} only once differentiable, is obtained for $\alpha_n = n$ and $\mathcal{I} = [-1,1]$ by letting $\phi(t) = t\sqrt{|t|}$ deterministically, and $\phi_n = \phi + \frac{1}{n}$.
I include here also two elementary lemmas which will be needed for negotiating between the discrete and continuous settings.

Lemma 2.4. *Given positive integers k, and $c \ge b \ge a \ge 1$,*

$$
\frac{\binom{b}{k}-\binom{a}{k}}{\binom{c}{k}} \le \frac{k(b-a)b^{k-1}}{c^k}, \tag{5}
$$

$$
\frac{1}{\binom{a}{k}} - \frac{1}{\binom{b}{k}} \le \frac{k(b-a)}{a\binom{a}{k}} \tag{6}
$$

$$
\left|\frac{\binom{a}{k}}{\binom{c}{k}} - \frac{a^k}{c^k}\right| \le k^2\left(\frac{(a+1)^{k-1}}{2c^k} + \frac{a^k}{(c-k)^{k+1}}\right). \tag{7}
$$

Proof. Assume first $a \ge k$. Clearly,

$$
\begin{aligned}
\frac{\binom{b}{k}-\binom{b-1}{k}}{\binom{c}{k}} &= \frac{\bigl(b-(b-k)\bigr)\bigl((b-1)(b-2)\cdots(b-k+1)\bigr)}{c(c-1)\cdots(c-k+1)} \\
&\le \frac{k}{c}\left(\frac{b}{c}\right)^{k-1}.
\end{aligned}
$$

Iterating this $b-a$ times proves (5).
If $b < k$, the left-hand side of (5) is zero, so the statement is trivial. The last case is $b \ge k > a$. Then

$$
\frac{\binom{b}{k}-\binom{a}{k}}{\binom{c}{k}} = \frac{\binom{b}{k}}{\binom{c}{k}} = \frac{b\cdots(b-k+1)}{c\cdots(c-k+1)} \le \left(\frac{b}{c}\right)^k \le \frac{\bigl(bk-ka\bigr)b^{k-1}}{c^k},
$$

since $k > a$ implies that $b(k-1) \geq (a+1)(k-1) \geq ka$. This completes the proof of (5).

To demonstrate (7), use Taylor's Theorem for the function x^{-k} to see that

$$\begin{aligned}
\left| \frac{\binom{a}{k}}{\binom{c}{k}} - \frac{a^k}{c^k} \right| &\leq \left| \frac{\binom{a}{k}}{\binom{c}{k}} - \frac{a^k}{\binom{c}{k}} \right| + \left| \frac{a^k}{c^k} - \frac{a^k}{\binom{c}{k}} \right| \\
&\leq \frac{\binom{k}{2} a(a-1)\cdots(a-k+2)}{c(c-1)\cdots(c-k+1)} + a^k \left| c^{-k} - (c-k)^{-k} \right| \\
&\leq \binom{k}{2} \frac{(a+1)^{k-1}}{c^k} + k^2 \frac{a^k}{(c-k)^{k+1}}.
\end{aligned}$$

Finally,

$$\frac{1}{\binom{b-1}{k}} - \frac{1}{\binom{b}{k}} = \frac{1}{\binom{b-1}{k}} \left(1 - \frac{b-k}{b} \right) = \frac{k}{b\binom{b-1}{k}},$$

from which iteration yields (6). □

Lemma 2.5. *If a, b, and n are positive real numbers with $b \leq n/2$ and $n > 1$,*

$$\left| (1 - \frac{b}{n})^{an} - e^{-ab} \right| \leq e^{-ab} \frac{ab^2}{2n} \leq \frac{2}{ae^2} \cdot \frac{1}{n}.$$

Proof. In general, if x and y are positive real numbers, by the Theorem of the Mean

$$\begin{aligned}
|x - y| &= |e^{\log x} - e^{\log y}| \\
&\leq |\log x - \log y| \max\{x, y\}.
\end{aligned} \tag{8}$$

Now,

$$\log(1 - \frac{b}{n})^{an} - \log e^{-ab} = an \log(1 - \frac{b}{n}) + ab \leq \frac{-ab^2}{2n},$$

since $\log(1+x) \leq x - x^2/2$ for all $x > -1$. This indicates as well that

$$e^{-ab} > (1 - \frac{b}{n})^{an},$$

so an application of (8) completes the proof. (Note: I have not been able to find a source for this elementary argument, but it was brought to my attention — to replace messier computations — by Brad Mann.) □

3. Theoretical results

The random functions $\phi_n : [0,T] \to [0,T']$ and $S_n : [0,T'] \to \mathbb{R}$ $(1 \leq n < \infty)$ are taken to be in $\mathcal{D}[0,T]$ or $\mathcal{D}[0,T']$ respectively. These define cadlag processes $F_n = S_n \circ \phi_n$. The process ϕ_n is assumed to converge uniformly in probability to a continuous, strictly increasing nonrandom function ϕ, with $\phi_n(0) = \phi(0) = 0$. The expectations of all the random functions exist pointwise, and the expectations $\mathrm{E}\, S_n(t)$ converge uniformly in t to a twice continuously differentiable function $\overline{S}(t)$. Also posit an increasing sequence of positive normalization constants (α_n) which go to infinity with n, and define the normalized processes

$$\begin{aligned}
\widetilde{\phi}_n(t) &= \alpha_n\big(\phi_n(t) - \phi(t)\big)\,,\\
\widetilde{S}_n(t) &= \alpha_n\big(S_n(t) - \mathrm{E}\,S_n(t)\big)\,, \quad \text{and}\\
\widetilde{F}_n(t) &= \alpha_n\big(F_n(t) - \mathrm{E}\,F_n(t)\big)\,.
\end{aligned}$$

Following (fairly) standard nomenclature, I will call the sequences of random variables $\widetilde{S}_n$ and $\widetilde{\phi}_n$ defined on a common probability space *asymptotically independent* if for every bounded uniformly continuous $f : \mathcal{D}[0,T'] \to \mathbb{R}$ and $g : \mathcal{D}[0,T] \to \mathbb{R}$,

$$\lim_{n\to\infty}\Big|\mathrm{E}\big[f(\widetilde{S}_n)g(\widetilde{\phi}_n)\big] \;-\; \mathrm{E}\big[f(\widetilde{S}_n)\big]\,\mathrm{E}\big[g(\widetilde{\phi}_n)\big]\Big| = 0\,. \tag{9}$$

Proposition 3.1. *Suppose that*

$$\lim_{n\to\infty}\alpha_n \sup_{t\in[0,T']}\big|\mathrm{E}\,S_n(t) - \overline{S}(t)\big| = \lim_{n\to\infty}\alpha_n \sup_{t\in[0,T]}\big|\mathrm{E}\,F_n(t) - \overline{S}(\phi(t))\big| = 0. \tag{10}$$

Suppose also that $\widetilde{\phi}_n$ converges uniformly in probability to an almost-surely continuous process $\widetilde{\Phi}$, and that $\widetilde{S}_n$ and $\widetilde{\phi}_n$ are asymptotically independent.

(i) If $\widetilde{S}_n(t)$ converges weakly to an almost-surely continuous process $\widetilde{S}(t)$, then $\widetilde{F}_n(t)$ also converges weakly to an almost-surely continuous process $\widetilde{F}(t)$, and

$$\widetilde{F} =_d \widetilde{S}\circ\phi + (\overline{S}'\circ\phi)\cdot\widetilde{\Phi}\,, \tag{11}$$

where $\overline{S}'(u) = \frac{d\overline{S}}{dt}\big|_{t=u}$.

(ii) Suppose that ϕ_n and S_n satisfy

$$\sup\Big\{\alpha_n\big|S_n(s) - S_n(t)\big| \;:\; [s,t]\cap \mathrm{Range}\,\phi_n = \emptyset\Big\} \xrightarrow{n\to\infty}_P 0 \tag{12}$$

for all n, and that ϕ^{-1} is twice differentiable, with bounded first and second derivatives. If $\widetilde{\phi}_n$ and $\widetilde{F}_n$ converge to continuous Gaussian processes $\widetilde{\Phi}$ and $\widetilde{F}$ respectively, then $\widetilde{S}_n$ converges to a continuous Gaussian process $\widetilde{S}$, and the relation (11) *is satisfied.*

Note: The condition (12) is required to exclude pathological fluctuations which would destroy the tightness of S_n, but which would not be reflected in F_n, because they occur within gaps in the range of ϕ_n. It is trivially satisfied in the examples which I am considering, since S_n is constant between the times that are hit by ϕ_n.

Proof. The function $\widetilde{F}_n$ may be decomposed into

$$\begin{aligned}
\widetilde{F}_n(t) = {} & \alpha_n\Big(S_n\big(\phi(t)\big) - \overline{S}_n\big(\phi(t)\big)\Big) + \alpha_n\overline{S}'\big(\phi(t)\big)\Big(\phi_n(t) - \phi(t)\Big)\\
& + \alpha_n\Big(\overline{S}\big(\phi_n(t)\big) - \overline{S}\big(\phi(t)\big) - \overline{S}'\big(\phi(t)\big)\big(\phi_n(t) - \phi(t)\big)\Big)\\
& + \alpha_n\Big(S_n\big(\phi_n(t)\big) - \overline{S}_n\big(\phi_n(t)\big) - S_n\big(\phi(t)\big) + \overline{S}_n\big(\phi(t)\big)\Big)\\
& + \alpha_n\Big(\overline{S}_n\big(\phi_n(t)\big) - \overline{S}\big(\phi_n(t)\big) + \overline{S}\big(\phi(t)\big) - \overline{S}_n\big(\phi(t)\big)\Big)\\
& + \alpha_n\Big(\overline{S}_n\big(\phi(t)\big) - \overline{F}_n(t)\Big)\\
=: {} & A_n(t) + B_n(t) + \Gamma_n(t) + \Delta_n(t) + E_n(t) + Z_n(t),
\end{aligned}$$

where A_n, B_n, Γ_n, Δ_n, E_n are random functions, and Z_n is a deterministic function. By assumption (10), E_n and Z_n converge uniformly to 0 as $n \to \infty$.

By a Taylor approximation,

$$\left|\Gamma_n(t)\right| \le \frac{1}{2}\|\overline{S}''\|_\infty \alpha_n\big(\phi_n(t) - \phi(t)\big)^2$$

The tightness of $\alpha_n(\phi_n - \phi)$, together with the fact that α_n goes to infinity, shows that the first term converges to 0 in probability.

If $\widetilde{S}_n$ is tight in the uniform modulus (cf. page 5), then the term Δ_n converges to 0 uniformly in probability, since for every positive δ

$$\begin{aligned}\limsup_{n\to\infty} \mathrm{P}\big\{\sup|\Delta_n(t)| > \epsilon\big\} &= \limsup_{n\to\infty} \mathrm{P}\Big\{\sup_{t\in[0,T]}\big|\widetilde{S}_n\big(\phi_n(t)\big) - \widetilde{S}_n\big(\phi(t)\big)\big| > \epsilon\Big\} \\ &\le \limsup_{n\to\infty} \mathrm{P}\big\{\sup_{t\in[0,T]}\big|\phi_n(t) - \phi(t)\big| \ge \delta\big\} \\ &\quad + \limsup_{n\to\infty} \mathrm{P}\big\{\sup_{|s-t|<\delta}\big|\widetilde{S}_n(s) - \widetilde{S}_n(t)\big| > \epsilon\big\}.\end{aligned}$$

The first piece is 0 for every δ by the assumption of uniform convergence in probability of ϕ_n. The second piece converges to 0 as $\delta \to 0$ by the convergence of $\widetilde{S}_n$ to an almost surely continuous process. Thus,

$$\big\|\widetilde{F}_n - \widetilde{S}_n \circ \phi - (\overline{S}' \circ \phi) \cdot \widetilde{\phi}_n\big\|_\infty = \big\|\widetilde{F}_n - A_n - B_n\big\|_\infty \to_P 0. \tag{13}$$

If assumption (i) holds, then $(\widetilde{S}_n)$ is certainly tight in the uniform modulus, since it converges weakly to a continuous process. By Theorem 4.5 of [Bil68] (and the ensuing exercise 7), asymptotic independence and the convergence of $\widetilde{\phi}_n$ in probability suffice to show that the pair $(\widetilde{\phi}_n, \widetilde{S}_n)$ converges weakly. It follows that any linear combination also converges; that is, that

$$\widetilde{S}_n \circ \phi + \overline{S}' \cdot \widetilde{\phi}_n \to_w \widetilde{S} \circ \phi + (\overline{S}' \circ \phi) \cdot \widetilde{\Phi}\,.$$

Now suppose that the assumptions (ii) hold. By Skorokhod's Theorem (e.g., Theorem 3.2.1 of [Sko56]), the sequence of random functions must only be tight and have its finite-dimensional distributions converge. Unfortunately, the proof of the required relation (13) presupposed that $\widetilde{S}_n$ is tight.

Tightness entered only into the proof that $\Delta_n(t)$ converges to 0 uniformly in probability. As a substitute, we may reverse the argument that led to (13). Define

$$\begin{aligned}S_n^*(t) &:= S_n\big(\phi_n(\phi_n^{-1}(t))\big) = F_n\big(\phi_n^{-1}(t)\big), \\ \widetilde{S}_n^*(t) &:= \alpha_n\big(S_n^*(t) - \mathrm{E}\,S_n^*(t)\big).\end{aligned}$$

Carrying out the same sort of decomposition yields terms

$$\begin{aligned}\Gamma_n^*(t) &= \alpha_n\Big(\overline{S}\big(\phi_n^{-1}(t)\big) - \overline{S}\big(\phi^{-1}(t)\big) - \overline{S}'\big(\phi^{-1}(t)\big)\big(\phi_n^{-1}(t) - \phi^{-1}(t)\big)\Big) \\ \Delta_n^*(t) &= \widetilde{F}_n\big(\phi_n^{-1}(t)\big) - \widetilde{F}_n\big(\phi_n^{-1}(t)\big) \\ E_n^*(t) &= \alpha_n\Big(\overline{F}_n\big(\phi_n(t)\big) - \overline{S}\big(\phi_n(t)\big) + \overline{S}\big(\phi(t)\big) - \overline{F}_n\big(\phi(t)\big)\Big) \\ Z_n^*(t) &= \alpha_n\Big(\overline{F}_n\big(\phi^{-1}(t)\big) - \overline{S}_n(t)\Big)\end{aligned}$$

The functions E_n^* and Z_n^* converge to 0 uniformly by assumption (10). The term Δ_n^* converges uniformly to 0 in probability for essentially the same reason that Δ_n does, since $\widetilde{F}_n$ forms a tight sequence, and by Lemma 2.1,

$$\sup_{t\in[0,T']}\left|\phi_n^{-1}(t)-\phi^{-1}(t)\right| \le L_1 \sup_{t\in[0,T]}\left|\phi_n(t)-\phi(t)\right|, \tag{14}$$

where L_1 is the Lipschitz constant for ϕ^{-1}. This also shows that

$$\alpha_n \sup_t\big(\phi_n^{-1}(t)-\phi^{-1}(t)\big)^2 \le L_1^2\alpha_n \sup_t\big(\phi_n(t)-\phi(t)\big)^2 \xrightarrow{n\to\infty}_P 0,$$

which means that Γ_n^* converges uniformly to 0 in probability.

The assumption (12) implies that $\sup_t|\widetilde{S}_n^*(t)-\widetilde{S}_n(t)|\to 0$ as n goes to ∞. It follows that

$$\left\|\widetilde{S}_n - \widetilde{F}_n\circ\phi - \overline{S}'\cdot\widetilde{\phi_n^{-1}}\right\|_\infty \to_P 0. \tag{15}$$

The sequence $(\widetilde{\phi_n^{-1}})$ is tight in the uniform modulus by Lemma 2.3. By Lemma 2.2, this holds for $\widetilde{S}_n$ as well.

The asymptotic characterization (13) implies the convergence of the characteristic functions. Together with the asymptotic independence and the multiplicativity of the characteristic functions of Gaussian variables, this establishes the convergence of the finite-dimensional distributions. □

For the second result, the probability space is assumed to admit a σ-algebra $\mathcal{R}_n$, such that S_n and ϕ_n are independent conditioned on $\mathcal{R}_n$, and have regular conditional probabilities. (For an account of conditional independence, see section 7.3 of [CT78]. In this paper the existence of regular conditional probabilities will be trivial, since $\mathcal{R}_n$ will be generated by a countably-valued random variable.)

The symbols p, q, r, D, c and C will all represent positive real constants, $q\ge\frac{1}{2}$. $\overline{S}(t)$ will generally represent a smooth approximation to the expectation of $S_n(t)$ such that

$$\sup_t\left|\overline{S}(t)-\mathrm{E}\,S_n(t)\right| = o(n^{-p}), \tag{16}$$

and it will be assumed that $\overline{S}$ has a local extremum at t_0, such that

$$\left|\overline{S}(t)-\overline{S}(t_0)\right| \le D|t-t_0|^q. \tag{17}$$

Proposition 3.2. *Suppose that S_n and ϕ_n are independent when conditioned on $\mathcal{R}_n$, have regular conditional distributions, and that the expectation of the fluctuations in S_n satisfies*

$$\mathrm{E}\left|S_n(t)-\mathrm{E}\,S_n(t)-\big(S_n(t_0)-\mathrm{E}\,S_n(t_0)\big)\right| \le cn^{-r}\sqrt{|t-t_0|} \tag{18}$$

for all n positive and $0\le t\le 1$. Suppose, too, that there is a function $\overline{S}$ satisfying conditions (16) and (17), and that

$$\mathrm{E}\Big[\left|\phi_n(t_0')-t_0\right|^q \,\Big|\, \mathcal{R}_n\Big] \le Cn^{-q/2} \text{ almost surely.} \tag{19}$$

Then for $\alpha<\min\{r+\frac{1}{4},\frac{q}{2}\}$ and $\alpha\le p$,

$$n^\alpha\left|F_n(t_0')-S_n(t_0)\right| \to 0 \quad \text{in } \mathfrak{L}^1.$$

Proof. Define $\tau = \left|\phi_n(t_0') - t_0\right|$, and let $\mathcal{R}_n^*$ be the σ-algebra generated by $\mathcal{R}_n$ and the random variable ϕ_n. Define, too, a random variable

$$c(\omega) = \sup_{t \neq t_0} \mathrm{E}\Big[\big|S_n(t) - \mathrm{E}\,S_n(t) - \big(S_n(t_0) - \mathrm{E}\,S_n(t_0)\big)\big| \,\Big|\mathcal{R}_n\Big]\, n^r |t - t_0|^{-1/2}.$$

By (18), $\mathrm{E}\,c(\omega) \leq c$. Now

$$\begin{aligned}\mathrm{E}\big|F_n(t_0') - S_n(t_0)\big| \leq \mathrm{E}\Big[&\big|(S_n - \overline{S}_n)(\phi_n(t_0')) - (S_n - \overline{S}_n)(t_0)\big| \\ &+ \big|\overline{S}(\phi_n(t_0')) - \overline{S}(t_0)\big| + \big|\overline{S}(t_0) - \overline{S}_n(t_0)\big|\Big].\end{aligned}$$

Conditional independence tells us that the regular conditional distribution for (S_n, τ) with respect to $\mathcal{R}_n$ is the product of the regular conditional distributions for S_n and for τ, which means that Fubini's theorem holds for the conditional expectations. That is, if one defines random functionals

$$\begin{aligned}G_n(\phi) &:= \big|(S_n - \overline{S}_n)(\phi(t_0')) - (S_n - \overline{S}_n)(t_0)\big|, \text{ and} \\ g_n(\phi) &:= \mathrm{E}\big[G(S_n, \phi) \,\big|\, \mathcal{R}_n\big],\end{aligned}$$

then

$$\mathrm{E}\big[G_n(\phi_n) \,\big|\mathcal{R}_n^*\big] = g_n(\phi_n) \leq c(\omega) n^{-r} \big|\phi_n(t_0') - t_0\big|^{1/2}.$$

Thus

$$\begin{aligned}\mathrm{E}\big[G_n(\phi_n)\big] &= \mathrm{E}\Big[\mathrm{E}\big[\mathrm{E}\big[G_n(\phi_n) \,\big|\, \mathcal{R}_n^*\big] \,\big|\, \mathcal{R}_n\big]\Big] \\ &\leq \mathrm{E}\Big[\mathrm{E}\big[c(\omega) n^{-r} \big|\phi_n(t_0') - t_0\big|^{1/2} \,\big|\, \mathcal{R}_n\big]\Big] \\ &\leq c n^{-r} \cdot (C n^{-q/2})^{1/2q} \\ &= O\big(n^{-r-1/4}\big).\end{aligned}$$

Similarly,

$$\mathrm{E}\big|\overline{S}(\phi_n(t_0')) - \overline{S}(t_0)\big| \leq D\,\mathrm{E}\big[\big|\phi_n(t_0') - t_0\big|\big]^q \leq C D n^{-q/2}.$$

Finally, the last term is $o(n^{-p})$ by assumption (16). □

4. Applications: socks and boxes

4.1. **Sorting problems.** Returning to the generalized sock-sorting problem described in the introduction, let $\vec{k} = (k_1, \ldots, k_N)$ be an N-tuple of independent random integers, each $k_i \geq 2$, and set $K_i = k_1 + \cdots + k_i$ and $K = K_N$. Let

$$p_k^{(N)} = \frac{1}{N} \sum_{i=1}^{N} \mathrm{P}\big\{k_i = k\big\}.$$

Consider now the set of permutations of K letters, viewed as bijections

$$\mathfrak{S}_{\vec{k}} := \Big\{\pi \colon \big\{(i,j) \;:\; 1 \leq i \leq N,\, 1 \leq j \leq k_i\big\} \to \{1, \ldots, K\}\Big\}.$$

Given a map $\pi \in \mathfrak{S}_{\vec{k}}$, let

$$s_n^i(\pi) = \sum_{j=1}^{k_i} \mathbf{1}_{\{\pi(i,j)\le n\}} - k_i \prod_{j=1}^{k_i} \mathbf{1}_{\{\pi(i,j)\le n\}} ;$$
$$s_n(\pi) = \sum_{i=1}^{N} s_n^i(\pi).$$

When there is no risk of confusion, the argument π will be suppressed. Identifying this with the sequential choice of elements from a set of size K, which is divided up into N classes of the form $\{(i,1),\dots,(i,k_i)\}$, s_n^i represents the number of elements from class i among $\pi^{-1}(1),\dots,\pi^{-1}(n)$, unless the class is complete (that is, the number is k_i), in which case $s_n^i = 0$. Similarly, s_n represents the total number of elements from incomplete classes among $\pi^{-1}(1),\dots,\pi^{-1}(n)$. The stipulation that π be chosen uniformly from $\mathfrak{S}_{\vec{k}}$ defines a stochastic process.

The next step is to embed this in a continuous-time process, by defining

$$S^i(t) = s^i_{[Kt]} \qquad S_N(t) = \frac{1}{N} s_{[Kt]} = \frac{1}{N}\sum_{i=1}^{N} S^i(t) \,.$$

The functions $S^i(t)$ are almost independent, as long as N is much larger than any of the k_i; but not quite, because only one jump can occur at a time. They become independent when the clock which determines the jump times is randomized. For each positive K, let $\big(X(K,\ell) : \ell = 1,\dots,K\big)$ be an array of i.i.d. random variables uniform on $[0,1]$; without further mention, condition on the almost-sure event that the $X(K,\ell)$ are distinct. For $1 \le i \le N$ and $1 \le j \le k_i$ define $x(i,j) = X(K, K_{i-1}+j)$. The empirical process

$$\phi_K(t) = \frac{1}{K}\sum_{\ell=1}^{K} \mathbf{1}_{[X(K,\ell),1]}$$

converges in probability to the identity function as $K \to \infty$, and that

$$\widetilde{\phi}_K(t) = \sqrt{K}\big(\Phi_K(t) - t\big)$$

converges weakly to a Brownian bridge $\widetilde{\Phi}$, with covariance $c_2(s,t) = s(1-t)$ for $s \le t$. The almost-sure representation theorem (Theorem 1.10.3 of [vdVW96]) then allows the convergence to occur uniformly almost surely, on a suitably defined common probability space. This circuitous definition has the advantage of providing an autonomous limit process, independent of $\vec{k}$. As N increases, $\widetilde{\phi}_K$ will simply run through a subsequence of $\big(\widetilde{\phi}_n\big)_{n=1}^{\infty}$, and necessarily will converge to the same limit.

Since the $x(i,j)$ are i.i.d. and distinct, they define a uniform permutation in $\mathfrak{S}_{\vec{k}}$. Define

$$f_i(t) = \sum_{j=1}^{k_i} \mathbf{1}_{[x(i,j),1]}(t) - k_i \prod_{j=1}^{k_i} \mathbf{1}_{[x(i,j),1]}(t) ;$$
$$F_N(t) = \frac{1}{N}\sum_{i=1}^{N} f_i(t) \,.$$

Then $F_N(t) =_d \frac{1}{N} S_N(\phi_K(t))$. Conditioned on the choice of $k_i = k$, the expectation is $\mathrm{E}\, f_i(t) = kt - kt^k$; the covariances $C_k(s,t) := \mathrm{Cov}\big(f_i(s), f_i(t)\big)$ are

$$C_k(s,t) = k(s - st) + k^2(s^k t + st^k - st^{k-1} - s^k t^k) \tag{20}$$

for $s \leq t$.

Assume now that the distribution of $\vec{k}$ has a limit, in the sense that $p_k = \lim_{N\to\infty} p_k^{(N)}$ exists, and impose the following conditions:

$$\kappa := \sup_i \mathrm{E}\big[k_i^4\big] < \infty \tag{21}$$

and

$$\lim_{N\to\infty} N^{2/3} \sum_{k=1}^{\infty} |p_k - p_k^{(N)}| k^4 = 0. \tag{22}$$

We then also have

$$\overline{S}(t) = \lim_{N\to\infty} \mathrm{E}\, F_N(t) = \sum_{k=2}^{\infty} p_k \big(kt - kt^k\big)\,.$$

Since this is a convex combination of strictly concave functions, it has a unique maximum at an internal point t_0.

Define

$$\begin{aligned}
\mu &:= \sum_{k=2}^{\infty} p_k k\,,\\
\widetilde{S}_N(t) &:= N^{-1/2}\big(S_N(t) - \mathrm{E}\, S_N(t)\big)\,,\\
M_N &:= \frac{1}{N} \max_{0\leq n\leq K} S_n\,, \text{ and}\\
\widetilde{M}_N &:= N^{1/2}\big(M_N - \overline{S}(t_0)\big)\,.
\end{aligned}$$

Proposition 4.1. *(i) The random functions $\widetilde{S}_N$ converge weakly to a Gaussian process on $[0,1]$ with continuous sample paths and covariance matrix*

$$c(s,t) = \sum_{k=2}^{\infty} p_k C_k(s,t) \; - \; \frac{s(1-t)}{\mu} \left(\sum_{k=2}^{\infty} p_k k(1 - ks^{k-1})\right) \left(\sum_{k=2}^{\infty} p_k k(1 - kt^{k-1})\right)$$

for $s \leq t$.

(ii) The triple

$$\left(N^{2/3}\big(\max_t S_N(t) - S_N(t_0)\big)\,,\, N^{1/3}\big(\arg\max S_N - t_0\big)\,,\, \widetilde{M}_N\right) \tag{23}$$

converges in distribution to

$$\left(\sup_{u\in\mathbb{R}} G(u)\,,\, \arg\max G\,,\, \widetilde{M}\right)\,,$$

where $\widetilde{M}$ is a normal random variable with variance 1, independent of $G(u)$. $G(u)$ is distributed as $\sigma B_u - Du^2$, where B_u is a standard two-sided Brownian motion,

$$\sigma^2 = \sum_{k=2}^{\infty} p_k (k - k^2(k-2)t_0^{k-1})\,, \tag{24}$$

and

$$D = \sum_{k=2}^{\infty} p_k \frac{k}{2}(1 - k^2(k-1)t_0^{k-2})\,. \tag{25}$$

Proof. By Proposition 7.3 of [Ste98], the assumptions (21) and (22) suffice to prove that the conclusions hold with S_N replaced by F_N, and with the covariance $c(s,t)$ replaced by

$$c_1(s,t) = \sum_{k=2}^{\infty} p_k C_k(s,t)\,.$$

(i) We know that $\widetilde{F}_N(t)$ converges weakly to a Gaussian process with covariance $c_1(s,t)$, and wish to extend this result by an application of Proposition 3.1. As I have already mentioned, the process $\widetilde{\phi}_K(t)$ converges almost surely to the Brownian bridge $\widetilde{\Phi}(t)$. By a slight abuse of indices,

$$\begin{aligned}
\widetilde{\phi}_N(t) &= N^{1/2} \cdot \frac{1}{K_N} \sum_{i=1}^{N} \sum_{j=1}^{k_i} (\mathbf{1}_{[x(i,j),1]}(t) - t) \\
&= \left(\frac{K_N}{N}\right)^{-1/2} \widetilde{\phi}_{K_N}(t)\,.
\end{aligned}$$

By (21) the variables k_i have *a fortiori* bounded second moments, and so satisfy the strong law of large numbers [Fel71, Theorem VII.8.3]. Since every subsequence of a convergent sequence converges to the same limit, $\widetilde{\phi}_N$ converges almost surely, hence also in probability, to $\mu^{-1/2}\widetilde{\Phi}$.

The time change ϕ_N takes on precisely the values m/K_N, $0 \le m \le K_N$, and S_N is constant on $[m/K_N, (m+1)/K_N)$. Thus condition (12) is trivially satisfied.

For any fixed k,

$$\begin{aligned}
\mathrm{E}\left[s^i_{[Kt]} \middle|\, k_i = k\right] &= \sum_{j=1}^{k-1} j \frac{\binom{[Kt]}{j}\binom{K-[Kt]}{k-j}}{\binom{K}{k}} \\
&= \frac{[Kt]}{\binom{K}{k}} \left(\binom{K-1}{k-1} - \binom{[Kt]-1}{k-1}\right) \\
&= \frac{k[Kt]}{K}\left(1 - \frac{\binom{[Kt]-1}{k-1}}{\binom{K-1}{k-1}}\right).
\end{aligned}$$

Since $Kt \ge [Kt] \ge Kt - 1$, an application of Lemma 2.4 gives

$$\begin{aligned}
\left|\mathrm{E}\left[s^i_{[Kt]} \middle|\, k_i = k\right] - k(t - t^k)\right| &\le k\left(\frac{1}{K} + \left|\frac{\binom{[Kt]-1}{k-1}}{\binom{K-1}{k-1}} - t^{k-1}\right|\right) \\
&\le \frac{k}{K} + \frac{k}{K-k}\left(1 + \frac{k}{K-k}\right)^k.
\end{aligned}$$

Integrating over $\vec{k}$, and using the trivial facts $K - k_i \geq 2N - 2$, $s^i_m \leq k_i$ for all m and i, and $(1+\epsilon)^k \leq e^{\epsilon k}$,

$$\begin{aligned}
\left|\mathrm{E}\left[s^i_{[Kt]} - k_i(t - t^{k_i})\right]\right| &= \left|\mathrm{E}\Big[\left|s^i_{[Kt]} - k_i(t - t^{k_i})\right|\left(\mathbf{1}_{\{k_i<\sqrt{N}\}} + \mathbf{1}_{\{k_i\geq\sqrt{N}\}}\right)\Big]\right| \\
&\leq \frac{(1+e)\kappa}{N} + \mathrm{E}\left[k_i \mathbf{1}_{\{k_i\geq\sqrt{N}\}}\right] \\
&\leq \frac{(2+e)\kappa}{N}.
\end{aligned}$$

For the expectation of $S_N(t)$ this yields

$$\mathrm{E}\, S_N(t) = \frac{1}{N}\sum_{i=1}^{N} \mathrm{E}\, S^i_N(t) = \sum_{k=2}^{\infty} p^{(N)}_k\left(t - t^k\right) + O\left(N^{-1}\right).$$

Thus

$$\begin{aligned}
\sup_{t\in[0,1]}\left|\mathrm{E}\, S_N(t) - \overline{S}(t)\right| &= \left|\sum_{k=2}^{\infty}(p_k - p^{(N)}_k)k(t - t^k)\right| \\
&\leq \sum_{k=2}^{\infty} k|p_k - p^{(N)}_k| \\
&= o\left(N^{-2/3}\right)
\end{aligned}$$

by (22). Similarly,

$$\sup_{t\in[0,1]}\left|\mathrm{E}\, F_N(t) - \overline{S}(t)\right| = o\left(N^{-2/3}\right)$$

Since the normalization factor is $\alpha_N = N^{1/2}$, this proves (10).

It follows now by Proposition 3.1 that $\widetilde{S}_N(t)$ converges weakly to a Gaussian process whose covariance is precisely

$$c(s,t) = c_1(s,t) - \frac{1}{\mu}\overline{S}'(s)\overline{S}'(t)c_2(s,t)\,,$$

where $c_2(s,t) = s(1-t)$ is the covariance of the Brownian bridge. In principle it would be possible to compute this covariance directly, using combinatorics of the discrete process and passing to the limit as $N \to \infty$. For the case k_i identically equal to 2 this has been done in [Ste96]; for larger k, though, the calculation becomes complicated and tedious.

(ii) By [Ste98], the triple

$$\left(N^{\frac{2}{3}}\left(\sup_t F_N(t) - F_N(t_0)\right),\, N^{\frac{1}{3}}\left(\arg\max F_N - t_0\right),\, N^{\frac{1}{2}}\left(\sup_t F_N(t) - \overline{F}_n(t_0)\right)\right)$$

converges in distribution to

$$\left(\sup_{u\in\mathbb{R}} G(u)\,,\, \arg\max G\,,\, \widetilde{M}\right).$$

The limit needs to be unchanged when all the F_N are replaced by S_N. This will be achieved if the difference created by this change goes to 0 in probability in the coupled version, where $F_N = S_N \circ \phi_N$, Since $\sup_t F_N(t)$ and $\sup_t S_N(t)$ have the same distribution, we know that the convergence of the third coordinate still holds if F_N is replaced by S_N. In the coupled version,

$$\arg\max S_N = \phi_N\big(\arg\max F_N\big)\,.$$

Since $\sup_{t\in[0,1]} |\phi_N(t) - t|$ is $O_P\big(N^{-1/2}\big)$, it follows that

$$N^{1/3}\big|\arg\max F_N - \arg\max S_N\big| \to_P 0\,,$$

which means that the second coordinate's convergence in distribution is also preserved when F_N is replaced by S_N.

The first coordinate only needs

$$N^{2/3}\big|F_N(t_0) - S_N(t_0)\big| \to_P 0\,,$$

since $\sup_t F_N(t) = \sup_t S_N(t)$. For this apply Proposition 3.2 with $q = 2$, D and $t_0 = t_0'$ as already defined, and $\mathcal{R}_n$ the σ-algebra determined by $\vec{k}$. The condition (17) is obviously satisfied, and (16) has already been established for $p = \frac{2}{3}$. Conditioned on any $\vec{k}$, ϕ_n is just an empirical process with $K \ge 2N$ points; thus

$$\mathrm{E}\Big[\big|\phi_N(t_0) - t_0\big|^2 \,\Big|\, \vec{k}\Big] \le \frac{1}{K} \le \frac{1}{2N},$$

so (19) holds for $q = 2$.

To prove that (18) holds, the individual jumps in the process $S_N(t)$ need to be very nearly independent. Define

$$\xi_n = S_n - S_{n-1},$$

and let $\tilde{\xi}_n$ be normalized to have mean 0:

$$\tilde{\xi}_n = \xi_n - \mathrm{E}\big[\xi_n\big].$$

For $t > t_0$, define $m_1 = [t_0K] + 1$ and $m_2 = [tK]$, and apply Lemma 4.2. Then

$$\begin{aligned}\mathrm{E}\big|S_N(t) - \mathrm{E}\,S_N(t) - \big(S_N(t_0) - \mathrm{E}\,S_N(t_0)\big)\big|& \\ \le N^{-1}\left(\mathrm{E}\left|\sum_{m=m_1}^{m_2}\sum_{n=m_1}^{m_2}\tilde{\xi}_m\tilde{\xi}_n\right|\right)^{1/2}& \\ \le CN^{-1}\sqrt{N|t-t_0|}&\end{aligned} \tag{26}$$

for a constant C (depending only on κ). Exactly the same is true if $t < t_0$. This completes the proof of condition (18) with the bound $\alpha < \frac{3}{4}$. In the present case α is $\frac{2}{3}$. □

It remains only to prove

Lemma 4.2. *If $k_1, k_2, \ldots$ are random variables whose distributions satisfy (21) and (22), and $m_1 = [t_1K] + 1$ and $m_2 = [t_2K]$ for fixed $0 \le t_1 < t_2 \le 1$,*

$$\mathrm{E}\left|\sum_{m=m_1}^{m_2}\sum_{n=m_1}^{m_2}\tilde{\xi}_m\tilde{\xi}_n\right| \le 10\kappa^{3/2}|t_2 - t_1|N \tag{27}$$

for all $N \ge 3$.

Proof. Begin by conditioning on a particular realization of $\vec{k}$. Let $A_{m,i}$ be the event that the draw at time m comes from class number i. By straightforward combinatorics, for $i \neq j$ and $m < n$,

$$\mathrm{E}\left[\xi_m \mathbf{1}_{A_{m,i}} \mid \vec{k}\right] = \frac{k_i}{K}\left(1 - k_i \frac{\binom{m-1}{k_i-1}}{\binom{K-1}{k_i-1}}\right)$$

$$\mathrm{E}\left[\xi_m \xi_n \mathbf{1}_{A_{m,i}\cap A_{n,j}} \mid \vec{k}\right] = \frac{k_i k_j}{K(K-1)} \times \left[1 - k_i \frac{\binom{m-1}{k_i-1}}{\binom{K-2}{k_i-1}} - k_j \frac{\binom{n-1}{k_j-1}}{\binom{K-2}{k_j-1}} + k_i k_j \frac{\binom{m-1}{k_i-1}}{\binom{K-2}{k_i-1}} \frac{\binom{n-k_i-1}{k_j-1}}{\binom{K-k_i-1}{k_j-1}}\right].$$

An application of Lemma 2.4 then shows that for $i \neq j$ and $m \neq n$,

$$\left|\mathrm{E}\left[\xi_m \xi_n \mathbf{1}_{A_{m,i}\cap A_{n,j}} \mid \vec{k}\right] - \mathrm{E}\left[\xi_m \mathbf{1}_{A_{m,i}} \mid \vec{k}\right] \mathrm{E}\left[\xi_n \mathbf{1}_{A_{n,j}} \mid \vec{k}\right]\right| \le \frac{6k_i^3 k_j^3}{K(K-1)N}.$$

This uses the crude approximation $K \ge 2N$ and $K - k_i - 1 \ge 2N - 3 \ge N$ when $N \ge 3$. For $i = j$ and $m \neq n$, it is immediate that

$$\left|\mathrm{E}\left[\xi_m \xi_n \mathbf{1}_{A_{m,i}\cap A_{n,j}} \mid \vec{k}\right] - \mathrm{E}\left[\xi_m \mathbf{1}_{A_{m,i}} \mid \vec{k}\right] \mathrm{E}\left[\xi_n \mathbf{1}_{A_{n,j}} \mid \vec{k}\right]\right| \le \frac{2k_i^2}{K(K-1)}.$$

This yields (for $m \neq n$)

$$\begin{aligned}\left|\mathrm{E}\left[\tilde{\xi}_m \tilde{\xi}_n \mid \vec{k}\right]\right| &= \sum_{i,j=1}^{N} \left|\mathrm{E}\left[\xi_m \xi_n \mathbf{1}_{A_{m,i}\cap A_{n,j}} \mid \vec{k}\right] - \mathrm{E}\left[\xi_m \mathbf{1}_{A_{m,i}} \mid \vec{k}\right] \mathrm{E}\left[\xi_n \mathbf{1}_{A_{n,j}} \mid \vec{k}\right]\right| \\ &\le \sum_{i\neq j=1}^{N} \frac{6k_i^3 k_j^3}{K(K-1)N} + \sum_{i=1}^{N} \frac{2k_i^2}{K(K-1)}\end{aligned}$$

For $m = n$, direct calculation shows that

$$\left|\mathrm{E}\left[\tilde{\xi}_m^2 \mid \vec{k}\right]\right| \le \sum_{i=1}^{N} \frac{k_i^3}{K} + \sum_{i\neq j=1}^{N} \frac{k_i^2 k_j^2}{K^2}.$$

Summing over m and n gives the estimate

$$\begin{aligned}\left|\sum_{m=m_1}^{m_2} \sum_{n=m_1}^{m_2} \tilde{\xi}_m \tilde{\xi}_n\right| &\le (m_2 - m_1)^2 \left(\sum_{i\neq j=1}^{N} \frac{6k_i^3 k_j^3}{K^2 N} + \sum_{i=1}^{N} \frac{2k_i^2}{K(K-1)}\right) \\ &\quad + (m_2 - m_1)\left(\sum_{i\neq j=1}^{N} \frac{k_i^2 k_j^2}{K^2} + \sum_{i=1}^{N} \frac{k_i^3}{K}\right) \\ &\le \frac{6|t_2 - t_1|^2}{N} \sum_{i\neq j=1}^{N} k_i^3 k_j^3 \; + \; 2|t_2 - t_1|^2 \sum_{i=1}^{N} k_i^2 \\ &\quad + \frac{|t_2 - t_1|}{N} \sum_{i\neq j=1}^{N} k_i^2 k_j^2 \; + \; |t_2 - t_1| \sum_{i=1}^{N} k_i^3\end{aligned}$$

Taking the expectation with respect to $\vec{k}$, and using the crude approximations that $\kappa \geq 1$ and $t_2 - t_1 < 1$, yields (27). $\square$

4.2. **Random allocations.** Begin with N empty boxes and an unlimited number of balls. Throw the balls one by one into boxes chosen uniformly at random. For positive integers k, let $c_k(n)$ be the number of boxes which contain exactly k balls, after n balls have been allocated. What is the distribution of $c_k^* = \max_n c_k(n)$, when N is large?

Let

$$\begin{aligned} q(k) &= e^{-k}\frac{k^k}{k!} \ ; \\ S_N(t) &= \frac{1}{N} c_k\big([Nt]\big) \ ; \\ \widetilde{S}_N(t) &= \sqrt{N}\big(S_N(t) - \mathrm{E}\, S_N(t)\big) \end{aligned}$$

Proposition 4.3. *(i) Let k be a fixed positive integer. For $n = [Nt]$,*

$$\mathrm{E}\, S_N(t) = \binom{n}{k}\left(\frac{(N-1)^{n-k}}{N^n}\right) ,$$

The restriction of $\widetilde{S}_N(t)$ to any compact interval $[0,T]$ converges to a Gaussian process with covariance

$$C(s,t) := \mathrm{Cov}\big(\widetilde{S}_k(s), \widetilde{S}_k(t)\big) = \frac{e^{-t}s^k}{k!} - \frac{e^{-(t+s)}(st)^{k-1}}{(k!)^2}\big(st + s(k-s)(k-t)\big)$$

for $s \leq t$.

(ii) For c_k^* defined as above, the variables

$$N^{-1/2}\big(c_k^* - Nq(k)\big)$$

converge weakly to a multivariate normal distribution with mean 0 and covariance

$$c(k_1,k_2) := \mathrm{Cov}\big(c_{k_1}^*, c_{k_2}^*\big) = q(k_1)\big(q(k_2-k_1) - q(k_2)\big)$$

for $k_1 \leq k_2$. The quantities

$$N^{-1/3}\big(c_k^* - c_k(kN)\big)$$

converge jointly to independent random variables distributed as $\sup_u G_k(u)$, where $G_k(u) = \sqrt{2q(k)}B_u - \frac{q(k)}{2k}u^2$.

The following lemma will be helpful, corresponding to Lemma 4.2 in the sock-sorting problem:

Lemma 4.4. *Define $\xi(n) = c_k(n) - c_k(n-1)$, and $\tilde{\xi}(n) = \xi(n) - \mathrm{E}\,\xi(n)$, the variable ξ normalized to have mean 0. Then*

$$\max_{m<n\in\mathbb{Z}^+} \mathrm{E}\big[\tilde{\xi}(m)\tilde{\xi}(n)\big] = O\big(N^{-1}\big).$$

Proof. First,

$$\begin{aligned} \max_{\substack{m<n\in\mathbb{Z}^+ \\ n>N^2/2}} \mathrm{E}\big[\tilde{\xi}(m)\tilde{\xi}(n)\big] &\leq \max_{n>N^2/2} \mathrm{E}\big[\big|\tilde{\xi}(m)\big|\big] \\ &\leq 2 \max_{n>N^2/2} \mathrm{E}\big[\xi(n)\big] \\ &= O\big(N^{-1}\big) , \end{aligned}$$

so we may restrict our attention to m and n smaller than $N^2/2$.

Since ξ takes on only the values $-1, 0, +1$,

$$\begin{aligned}
\mathrm{E}\big[\tilde{\xi}(m)\tilde{\xi}(n)\big] &= \mathrm{E}\big[\xi(m)\tilde{\xi}(n)\big] \\
&= \sum_{\substack{x=\pm 1 \\ y=\pm 1}} \Big(\mathrm{P}\big\{\xi(n) = x \big| \xi(m) = y\big\} - \mathrm{P}\big\{\xi(n) = x\big\}\Big)\, \mathrm{P}\big\{\xi(m) = y\big\} xy
\end{aligned}$$

Thus $\big|\mathrm{E}\big[\tilde{\xi}(m)\tilde{\xi}(n)\big]\big|$ is bounded by twice the maximum absolute value of the expressions in parentheses. Let A be the event that ball number n goes into the same box as ball number m; let $B_m(y)$ be the event that $\xi(m) = y$. Since $n > m$, A is independent of $B_m(y)$. Then

$$\begin{aligned}
&\Big|\mathrm{P}\big(B_n(x) \mid B_m(y)\big) - \mathrm{P}\big(B_n(x)\big)\Big| \\
&\quad \le \Big|\mathrm{P}\big(B_n(x) \cap A^{\complement} \mid B_m(y)\big) - \mathrm{P}\big(B_n(x) \cap A^{\complement}\big)\Big| + \mathrm{P}\big(A \mid B_m(y)\big) + \mathrm{P}\big(A\big) \\
&\quad = \Big|\mathrm{P}\big(B_n(x) \cap A^{\complement} \mid B_m(y)\big) - \mathrm{P}\big(B_n(x) \cap A^{\complement}\big)\Big| + \frac{2}{N},
\end{aligned}$$

so there will be no loss in restricting all events to $A^{\complement}$. This will be done without further comment. Further, by symmetry, nothing will be changed if one conditions on ball number m going into box 1 and ball number n going into box 2, just to facilitate the notation.

The conditional probability $\mathrm{P}\big\{\xi(n) = +1 \,\big|\, \xi(m) = +1\big\}$ is just the chance that there are exactly $k-1$ balls in box 2 at time $n-1$, given that there are $k-1$ balls in box 1 at time $m-1$. Letting r be the number of balls in box 2 at time $m-1$ and summing over all possibilities, it follows that

$$\begin{aligned}
&\mathrm{P}\big\{\xi(n) = 1 \mid \xi(m) = 1\big\} \\
&= \sum_{r=0}^{k-1} \binom{m-k}{r} \binom{n-m-1}{k-r-1} \left(\frac{N-2}{N-1}\right)^{m-k-r} (N-1)^{-r} \\
&\qquad \times \left(\frac{N-1}{N}\right)^{n-m-k+r} N^{-(k-1-r)} \\
&= \big(1 - \frac{2}{N}\big)^{m-k} \big(1 - \frac{1}{N}\big)^{n-2m} \sum_{r=0}^{k-1} N^{-k+1} \binom{m-k}{r} \binom{n-m-1}{k-r-1} (1+\epsilon)^r \\
&= \frac{1}{(k-1)!} e^{-t+\frac{2k}{N}} \big(1 + \frac{a}{N}\big) \sum_{r=0}^{k-1} N^{-k+1} \binom{m-k}{r} \binom{n-m-1}{k-r-1} (1+\epsilon)^r
\end{aligned}$$

where

$$t = \frac{n}{N}, \qquad |a| \le \frac{t^2}{2} + \frac{2k}{N}, \qquad \text{and } \epsilon = \frac{1}{N-2}.$$

Note that the last step used Lemma 2.5, so assumes that $t \le N/2$, which is to say, that $n \le N^2/2$. Meanwhile,

$$\begin{aligned}
\mathrm{P}\big\{\xi(n) = 1\big\} &= \frac{\binom{n-1}{k-1}(N-1)^{n-k}}{N^{n-1}} \\
&= \binom{n-1}{k-1} e^{-t+\frac{k}{N}} \big(1 + \frac{a'}{N}\big) N^{-k+1},
\end{aligned}$$

with

$$|a'| \le \frac{t^2}{2} + \frac{2k}{N}.$$

This yields

$$
\begin{aligned}
&\Big|\mathrm{P}\{\xi(n)=1 \,\big|\, \xi(m)=1\} - \mathrm{P}\{\xi(n)=1\}\Big| \\
&= \frac{e^{-t+\frac{k}{N}}}{N^{k-1}}\left|(1+\frac{a'}{N})\binom{n-1}{k-1} - (1+\frac{a}{N})\sum_{r=0}^{k-1}\binom{m-k}{r}\binom{n-m-1}{k-r-1}(1+\epsilon)^r\right| \\
&\le N^{-1}e^{-t}G(t)
\end{aligned}
$$

where G is a polynomial of degree $k-1$. Since $e^{-t}G(t)$ is bounded in t, this is a bound of order N^{-1} uniform in m,n for the first of the four terms that we want to bound. But the other terms are bounded in exactly the same way: the condition $\xi(n)=-1$ simply means that box 1 has exactly $k+1$ balls rather than k, and similarly for $\xi(m)$. □

I recall here a version of Propositions 7.1 and 6.4 of [Ste98]. The context is random functions $f_i : \mathcal{I} \to \{0,1\}$, where $\mathcal{I}$ is an interval, possibly infinite, and for simplicity the f_i are taken to be i.i.d. The total variation of f_i restricted to the subinterval $\mathcal{I}\cap(-\infty,t)$ (that is, the total number of jumps on this subinterval) is represented by $v_i(t)$, while

$$
c(s,t) = \mathrm{Cov}\big(f_i(s), f_i(t)\big)
$$

for $s \le t$.

I will say that $\mathrm{E}\, f_i$ has a unique quadratic maximum if there is an interior point $t_0 \in \mathcal{I}$ such that for any positive ϵ, $\sup_{|t-t_0|>\epsilon} \bar{f}_i(t) < \bar{f}_i(t_0)$, and for some positive D

$$
\mathrm{E}\, f_i(t) = \mathrm{E}\, f_i(t_0) - D\cdot(t-t_0)^2 + o\big(|t-t_0|^2\big) \quad \text{as } t \to t_0.
$$

I will say that "jumps are not clustered" if for every $t \in [S,T]$,

$$
\lim_{\delta\to 0}\delta^{-1}\sup_i \mathrm{E}\big[\big(v_i(t+\delta)-v_i(t)\big)^2\mathbf{1}_{\{v_i(t+\delta)\ge v_i(t)+2\}}\big] = 0. \tag{28}
$$

Lemma 4.5. *Suppose that the expectation functions $\bar{f}_i(t)$ and $\bar{v}_i(t)$ are twice continuously differentiable, that $\bar{f}_i(t)$ has a unique quadratic maximum, that $v_i(T)$, has a finite fourth moment, and that jumps are not clustered.*

Then $\widetilde{F}_N(t)$ converges weakly on every compact interval to a Gaussian process with covariance $c(s,t)$; the processes

$$
N^{2/3}\big(F_N(t_0+uN^{-1/3}) - F_N(t_0)\big)
$$

converge weakly on every compact interval to the process $G(u)=\sigma B_u - Du^2$, where $\sigma=\sqrt{\bar{v}_i'(t_0)}$ and $D=-\bar{f}_i''(t_0)/2$; and the triple

$$
\Big(N^{2/3}\big(M_N - F_N(t_0)\big)\,,\; N^{1/3}\big(\arg\max F_N - t_0\big)\,,\; \widetilde{M}_N\Big)
$$

converges in distribution to

$$
\big(\sup G(u)\,,\; \arg\max G\,,\; \widetilde{M}\big),
$$

where $\widetilde{M}$ is a normal variable independent of G with mean 0 and the same variance as $f_i(t_0)$.

Furthermore, if $f_i^{(j)}$, $(j=1,\dots,k)$ are random functions satisfying the above conditions separately for each j (with t_j taking the place of t_0), and if, in addition,

$$
\lim_{\gamma\to 0}\gamma^{-1}\,\mathrm{Cov}\big(f_i^{(j)}(t_j+\gamma u) - f_i^{(j)}(t_j)\,,\; f_i^{(j')}(t_{j'}+\gamma u') - f_i^{(j')}(t_{j'})\big) = 0\,,
$$

then the processes $G^{(j)}_{\sigma_j}(u)$ are independent.

Proof of Theorem 4.3. (i) The time change for the allocations process derives from a Poisson process. Define $\left(\mathcal{E}_{i,j}\right)$ for $1 \leq i \leq N$ and $1 \leq j < \infty$ to be i.i.d. exponential variables with expectation 1, and for $t \geq 0$ let

$$f_i(t) = \begin{cases} 1 & \text{if } \mathcal{E}_{i,1} + \cdots + \mathcal{E}_{i,k} \leq t < \mathcal{E}_{i,1} + \cdots + \mathcal{E}_{i,k+1}; \\ 0 & \text{otherwise.} \end{cases}$$

This is the indicator function of the event that a Poisson process with rate 1 has the value k at time t, so

$$\bar{f}_i(t) = e^{-t}\frac{t^k}{k!}\,.$$

Similarly,

$$\bar{v}_i'(t) = e^{-t}\left(\frac{t^{k-1}}{(k-1)!} + \frac{t^k}{k!}\right),$$

so $\bar{v}_i'(t_0) = \bar{v}_i'(k) = 2q(k)$. The maximum of $\mathrm{E}\, f_i$ occurs at $t_0 = k$, and

$$\bar{f}_i''(k) = -\frac{k^{k-1}}{k!}e^{-k} = -\frac{q(k)}{k}\,.$$

By Stirling's formula, $q(k) \sim (2\pi k)^{-1/2}$ for large k. Also, for $s \leq t$, by the "memorylessness" of exponential variables,

$$\begin{aligned} c(s,t) &:= \mathrm{Cov}\big(f_i(s), f_i(t)\big) \\ &= \mathrm{P}\big\{\mathcal{E}_{i,1} + \cdots + \mathcal{E}_{i,k} \leq s \leq t < \mathcal{E}_{i,1} + \cdots + \mathcal{E}_{i,k+1}\big\} - \bar{f}_i(s)\bar{f}_i(t) \\ &= e^{-s}\frac{s^k}{k!}\cdot e^{-(t-s)} - e^{-s}\frac{s^k}{k!}\cdot e^{-t}\frac{t^k}{k!} \\ &= e^{-t}\frac{s^k}{k!}\left(1 - e^{-s}\frac{t^k}{k!}\right). \end{aligned}$$

The time change is $\phi_N(t) = (T+1) \wedge \frac{1}{N}\cdot \#\big\{(i,j) : \mathcal{E}_{i,j} \leq t\big\}$. for $t \in [0,T]$. This is just $\frac{1}{N}$ times a Poisson process with rate N, cut off at $T' = T+1$ in order to match the formal conditions of Proposition 3.1. In fact, we may and shall ignore the truncation, since the probability of reaching it is on the order of e^{-N}.

The first step is to show that $\widetilde{\phi}_N$ converges uniformly in probability to the identity function. Since $\lim_{n\to\infty} \mathrm{E}\big|\widetilde{\phi}_N(t) - \widetilde{\phi}_N(s)\big|^4 = 3|t-s|^2$, Billingsley's moment condition (Theorem 12.3 of [Bil68]) shows that $(\widetilde{\phi}_N)$ is tight in the uniform modulus. In addition, $\mathrm{E}\,\widetilde{\phi}_N(t) = 0$, $\mathrm{E}\,\widetilde{\phi}_N(t)^2 = t$ (for $0 \leq t \leq T$), and $\widetilde{\phi}_N$ has independent increments, so Theorem 19.2 of [Bil68] implies the weak convergence of $\widetilde{\phi}_N$ to standard Brownian motion.

An application of Proposition 3.1 then implies (i), once the condition (10) has been established. (Condition (12) is again trivial, since S_N is constant on the intervals between points of the range of ϕ_N.) These simply say that the expectations of the time-changed process F_N and of the original process are sufficiently close to one another, and to $\overline{S} = \lim_{n\to\infty} \mathrm{E}\, S_N$. Since the f_i are i.i.d., the latter condition is trivial.

For any positive integer n, by Lemma 2.5,

$$\begin{aligned} \binom{n}{k}\frac{(N-1)^{n-k}}{N^n} &= \frac{1}{k!}\cdot\frac{n}{N}\cdot\frac{n-1}{N}\cdot\cdots\cdot\frac{n-k+1}{N}\cdot\left(1-\frac{1}{N}\right)^{N-k} \\ &= \frac{1}{k!}\left(\frac{n}{N}\right)^k e^{-n/N} + O\big(N^{-1}\big)\,, \end{aligned}$$

where the big O terms are uniform in n between 0 and TN for any fixed T. It follows that

$$\lim_{N\to\infty} \sqrt{N} \sup_{t\in[0,T]} \left| e^{-t}\frac{t^k}{k!} - \binom{n}{k}\left(\frac{(N-1)^{n-k}}{N^n}\right)\right| = 0\,,$$

where $n = [Nt]$ completing the proof of (i).

(ii) The asymptotic second-order behavior goes essentially as in section 4.1. The only condition in Lemma 4.5 which is not trivial is (28); but this holds as well, since

$$\begin{aligned}
&\lim_{\delta\to 0} \delta^{-1} \sup_i \mathrm{E}\big[\big(v_i(t+\delta) - v_i(t)\big)^2 \mathbf{1}_{\{v_i(t+\delta)\geq v_i(t)+2\}}\big] \\
&\quad\leq 4 \lim_{\delta\to 0} \delta^{-1}\, \mathrm{P}\big\{2 \text{ jumps in } (t, t+\delta)\big\} \\
&\quad\leq 4 \lim_{\delta\to 0} \delta^{-1}\, \mathrm{P}\big\{\mathcal{E}_{i,1} + \cdots + \mathcal{E}_{i,k-1} \leq t\big\}\, \mathrm{P}\big\{\mathcal{E}_{i,k} \leq 2\delta\big\}\, \mathrm{P}\big\{\mathcal{E}_{i,k+1} < 2\delta\big\} \\
&\quad= 0.
\end{aligned}$$

As before, it remains only to show that $S_N(k)$ may be substituted for $F_N(k)$ in the limit, which will follow if Proposition 3.2 may be applied to show that in the joint probability space $N^{1/3}\big(S_N(k) - F_N(k)\big)$ converges to 0 in probability. In this case, the conditioning σ-algebra $\mathcal{R}_n$ is trivial.

The condition (17) is satisfied for some D with $q = 2$, because $\overline{S}(t) = e^{-t}t^k/k!$ is smooth and has vanishing first derivative 0 and negative second derivative at $t_0 = k$, while being bounded between 0 and $\overline{S}(t_0)$. Condition (19) is simply an elementary statement about the variance of a Poisson distribution.

To establish condition (18), observe that for $t > k$

$$S_N(t) - \mathrm{E}\, S_N(t) - \big(S_N(t_0) - \mathrm{E}\, S_N(t_0)\big) = \frac{1}{N} \sum_{m=kN+1}^{[tN]} \tilde{\xi}(m)\,.$$

The calculation is then the same as (26), with Lemma 4.4 in place of Lemma 4.2; $t < k$ is nearly identical. It follows that $N^{-1/3}\big(c_k(kN) - NF_N(k)\big) \to_P 0$, and so that $N^{-1/3}\big(c_k^* - c_k(kN)\big)$ converges to the desired distribution.

Let us consider now the joint distribution of the c_k^*. Starting from a single sequence of exponential variables $\mathcal{E}_{i,1}, \ldots, \mathcal{E}_{i,K+1}$, we define the random functions $\big(f_i^{(1)}(t), \ldots, f_i^{(K)}(t)\big)$ by the above recipe, with

$$f_i^{(k)}(t) = \begin{cases} 1 & \text{if } \mathcal{E}_{i,1} + \cdots + \mathcal{E}_{i,k} \leq t < \mathcal{E}_{i,1} + \cdots + \mathcal{E}_{i,k+1}\,; \\ 0 & \text{otherwise.} \end{cases}$$

Since the differences

$$N^{-1/2}\big|\sup_t F_N^{(k)}(t) - F_N^{(k)}(k)\big|$$

converge individually in probability to 0 (as $N \to \infty$), any finite collection of them converges jointly to 0. On the scale of $\sqrt{N}$ the c_k^* thus have the same joint distribution as the variables $F_N^{(k)}(k)$, and as $N \to \infty$,

$$\Big(N^{-1/2}\big(c_k^* - q(k)\big)\Big)_{k=1}^{K}$$

converges in distribution to a multivariate normal distribution with mean 0 and covariance matrix

$$c(k_1, k_2) = q(k_1)\big(q(k_2 - k_1) - q(k_2)\big)$$

for $k_1 \le k_2$.

Fix a real γ, and define for real u and positive integers k

$$A(k,u) = \bigl\{ \bigl| f_i^{(k)}(k) - f_i^{(k)}(k+\gamma u) \bigr| = 1 \bigr\}.$$

Then for $k < k'$,

$$\Bigl| \mathrm{Cov}\Bigl(f_i^{(k)}(k+\gamma u) - f_i^{(k)}(k), f_i^{(k')}(k'+\gamma u') - f_i^{(k')}(k') \Bigr) \Bigr| \\ \le \mathrm{P}\bigl(A(k,u) \cap A(k',u')\bigr) + \mathrm{P}\bigl(A(k,u)\bigr)\mathrm{P}\bigl(A(k',u')\bigr).$$

Each of these terms on the right is bounded above by $\mathrm{P}\bigl\{\mathcal{E} \le \gamma(u \vee u')\bigr\}^2$, where $\mathcal{E}$ is an exponential variable with expectation 1. Thus,

$$\lim_{\gamma \to 0} \gamma^{-1} \Bigl| \mathrm{Cov}\Bigl(f_i^{(k)}(k+\gamma u) - f_i^{(k)}(k), f_i^{(k')}(k'+\gamma u') - f_i^{(k')}(k') \Bigr) \Bigr| = 0\,.$$

By Lemma 4.5, this is precisely the condition required for the corrections to c_k^* of order $N^{1/3}$ to be asymptotically independent of one another. □

This example may readily be generalized by changing the parameters of the exponential variables. For instance, if one wishes to have probability p_i of throwing the ball into box i, simply assign the variables $\mathcal{E}_{i,j}$ all to have expectation $1/p_i$, instead of 1. Another possibility is to change the probability of picking a given box according to its contents. For instance, letting $\mathcal{E}_{i,j}$ be exponential with expectation $1/j$ will distribute the balls to the boxes according to Bose-Einstein statistics. In most imaginable variations, the nature of the limiting behavior is unchanged.

Acknowledgment — This work is based on my doctoral dissertation [Ste96]. I would like to express the deep gratitude I owe to my thesis adviser, Persi Diaconis. Part of this research was supported by a postdoctoral fellowship from the *Deutsche Forschungsgemeinschaft.* Thanks are due as well to David Gilat for his assistance in proofreading portions of the final draft.

References

[Ber82] Daniel Bernoulli. De usu algorithmi infinitesimalis in arte coniectandi specimen. In David Speiser, editor, *Die Werke*, volume 2 of *Die Gesammelten Werke der Mathematiker und Physiker der Familie Bernoulli*, pages 276–287. Birkhauser, Basel, Boston, 1982.

[BH91] Gunnar Blom and Lars Holst. Embedding procedures for discrete problems in probability. *Mathematical Scientist*, 16:29–40, 1991.

[Bil68] Patrick Billingsley. *Convergence of Probability Measures.* John Wiley & Sons, New York, 1968.

[CT78] Yuan Shih Chow and Henry Teicher. *Probability Theory: Independence, Interchangeability, Martingales.* Springer-Verlag, New York, Heidelberg, Berlin, 1978.

[Dud89] Richard Dudley. *Real Analysis and Probability.* Wadsorth & Brooks/Cole, 1989.

[Fel71] William Feller. *An Introduction to Probability and its Applications*, volume 2. John Wiley & Sons, New York, 1971.

[Kin93] J F C Kingman. *Poisson Processes.* Oxford University Press, Oxford, 1993.

[Lan96] Kenneth Lange. Illustration of some moment identities for order statistics. *Statistics & Probability Letters*, 27(1):91–98, 1996.

[LP98] W. V. Li and G. Pritchard. A central-limit theorem for the sock-sorting problem. In *Progress in Probability*, volume 43, pages 245–248. Birkhauser, 1998.

[Lut88] Rick Luttman. Matching socks. *American Mathematical Monthly*, 95:357, May 1988.

[Pol84] David Pollard. *Convergence of Stochastic Processes.* Springer-Verlag, New York, Heidelberg, Berlin, 1984.

[Sko56] A. S. Skorokhod. Limit theorems for stochastic processes. *Theory of Probability and its Applications*, 1(2):261–290, 1956.

[Ste96] David Steinsaltz. *Socks & Boxes: Variations on Daniel Bernoulli's Marriage Problem*. PhD thesis, Harvard University, 1996.

[Ste98] David Steinsaltz. Deviation bounds and limit theorems for the maxima of some sorting and storage processes. Unpublished, 1998.

[vdVW96] Aad van der Vaart and Jon Wellner. *Weak convergence and empirical processes: with applications to statistics*. Springer Verlag, New York, Heidelberg, Berlin, 1996.

E-mail address: `dstein@stat.berkeley.edu`

Department of Statistics, University of California, 367 Evans Hall, Berkeley, CA 94720, U.S.A.