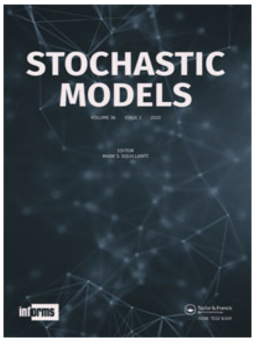

Stochastic Models

ISSN: 1532-6349 (Print) 1532-4214 (Online) Journal homepage: https://www.tandfonline.com/loi/lstm20

Asymptotics on the number of walks until no shoes when the number of doors is large

May-Ru Chen, Shoou-Ren Hsiau, Jia-Ching Tsai & Yi-Ching Yao

To cite this article: May-Ru Chen, Shoou-Ren Hsiau, Jia-Ching Tsai & Yi-Ching Yao (2020) Asymptotics on the number of walks until no shoes when the number of doors is large, Stochastic Models, 36:3, 428-451, DOI: 10.1080/15326349.2020.1745081

To link to this article: https://doi.org/10.1080/15326349.2020.1745081

Published online: 16 Apr 2020.

STOCHASTIC MODELS
2020, VOL. 36, NO. 3, 428–451
https://doi.org/10.1080/15326349.2020.1745081

Asymptotics on the number of walks until no shoes when the number of doors is large

May-Ru Chen[a], Shoou-Ren Hsiau[b], Jia-Ching Tsai[b], and Yi-Ching Yao[c]

[a]Department of Applied Mathematics, National Sun Yat-sen University, Kaohsiung, Taiwan, ROC; [b]Department of Mathematics, National Changhua University of Education, Changhua City, Taiwan, ROC; [c]Institute of Statistical Science, Academia Sinica, Taipei, Taiwan, ROC

ABSTRACT

A man has a house with n doors. Initially he places k pairs of walking shoes at each door. For each walk, he chooses one door at random, and puts on a pair of shoes, returns after the walk to a randomly chosen door and takes off the shoes at the door. Let T_n be the first time a door is chosen to walk out but with no shoes available. We show that as $n \to \infty$, T_n has the same asymptotic distribution and moments as the number of choices required to choose among n equally likely alternatives repeatedly until any one of the alternatives has appeared $k+1$ times. To derive these results, we need to consider a more general setting where the numbers of pairs of shoes initially placed at the doors (initial configuration) are not necessarily equal. We show that T_n increases in the sense of stochastic ordering if the initial configuration is more evenly distributed in the sense of majorization.

ARTICLE HISTORY
Received 26 August 2018
Accepted 17 March 2020

KEYWORDS
Coupling; majorization; number of walks until no shoes; stochastic ordering

AMS MSC 2010
60C05; 60F99

1. Introduction

A man has a house with $n \geq 2$ doors. Initially he places $k \geq 1$ pairs of walking shoes at each door. For each walk, he chooses one door at random, and puts on a pair of shoes, returns after the walk to a randomly chosen door and takes off the shoes at the door. Sooner or later he discovers that no shoes are available at the door he has chosen for a further walk (and has to walk barefoot). We are interested in the asymptotic behavior as $n \to \infty$ of the distribution and moments of the number of finished walks (before walking barefoot). When $n = 2$ (a house with 2 doors), this problem is referred to as "Number of walks until no shoes" in the well-known book "Problems and Snapshots from the World of Probability" by Blom et al.[2]. The problem first appeared in *The American Mathematical Monthly*[1].

Our study of this problem was originally motivated by modeling, design and analysis of bicycle-sharing systems[3,8], which have become popular in

CONTACT May-Ru Chen mayru@faculty.nsysu.edu.tw Department of Applied Mathematics, National Sun Yat-sen University, 70 Lien-hai Rd., Kaohsiung 804, Taiwan, ROC.

major cities worldwide. In a bicycle-sharing system, for example the YouBike system[7], a user can rent a bike at one station and return it at another station. For such a system to be user-friendly, the number of stations where a user can rent and return a bike needs to be large and the number of bikes available at a station should depend on the demand at the station. In particular, it is costly if a user discovers that no bikes are available at the chosen station. A bicycle-sharing system may be formulated mathematically as follows. Suppose that there are n stations labeled $1, 2, \ldots, n$. For $i = 1, 2, \ldots,$ let a_i be the arrival time of the i-th customer at station $s_i \in \{1, \ldots, n\}$ with $0 < a_1 < a_2 < \cdots$. The i-th customer rents a bike (if available) at station s_i and returns it at station $s'_i \in \{1, \ldots, n\}$ at the departure time b_i with $b_i > a_i$. The i-th customer is lost if no bike is available at station s_i. To reduce the rate of customers lost, bikes need to be shipped between stations periodically according to a policy that takes various costs into account. While stochastic modeling of a bicycle-sharing system and finding a cost-effective policy is a challenging issue in operations management, our doors-shoes problem corresponds to a much simplified special case with $b_i < a_{i+1}$ (i.e., the next customer arrives after the current customer returns the bike).

More specifically, with the n doors labeled $1, \ldots, n$, let $D_t, D'_t, t = 1, 2, \ldots$ be i.i.d. uniformly distributed on $\{1, \ldots,, n\}$, where D_t denotes the labeling of the door which the man chooses to go out for the t-th walk and D'_t denotes the labeling of the door which he chooses to return upon completing the t-th walk. Let

$$L_n = \inf\left\{ t \geq k+1 : \sum_{r=1}^{t} \mathbf{1}_{\{D_r = D_t\}} = k+1 \right\}, \tag{1}$$

the first time a door has been chosen $k+1$ times to walk out. Note that L_n can be viewed as the number of choices required to choose among n equally likely alternatives repeatedly until any one of the alternatives has appeared $k+1$ times (cf. Ref.[4]). Let T_n be the first time a door is chosen to walk out but with no shoes available (i.e., the first time the man walks barefoot). Thus $T_n - 1$ is the number of finished walks (before walking barefoot).

It is instructive to consider a more general setting where the initial numbers of pairs of shoes placed at the n doors are not necessarily equal. Let $V_n = (v_1, \ldots, v_n)$ where v_i denotes the initial number of pairs of shoes placed at door i. We will refer to V_n as the *initial configuration* (of the numbers of pairs of shoes placed at the n doors). Let $\mathcal{L}(T_n|V_n)$ and $E(T_n^r|V_n)$ denote, respectively, the distribution and r-th moment of T_n when the initial configuration is V_n. Note that $\mathcal{L}(T_n|V_n)$ and $E(T_n^r|V_n)$ are invariant with respect to permutations of V_n.

In Section 2, we first describe a useful relationship between T_n and L_n (Proposition 1) and then give the main results (Theorems 2–5) which show under suitable conditions on V_n that T_n has the same asymptotic distribution and moments as L_n. Moreover, we also consider the case $k=1$ and derive (with $V_n = (1, ..., 1)$) the asymptotic (conditional) distribution and moments of $T_n - L_n$ given $T_n > L_n$, a consequence of which is

$$\lim_{n\to\infty} E(T_n - L_n) = \frac{2}{3}.$$

The proofs of Proposition 1 and Theorems 3–5 are contained in Section 3. To prove Theorem 3, we need to show that $\mathcal{L}(T_n|V_n)$ is stochastically smaller than $\mathcal{L}(T_n|V'_n)$ if V_n majorizes V'_n (see Theorem 6), which is of independent interest. The proof of Theorem 6 is given in Appendix.

2. Main results

In this section, we first give a useful relationship between T_n and L_n. Then we will show under suitable conditions on V_n that T_n has the same asymptotic distribution and moments as L_n. The proofs of Proposition 1 and Theorems 3–5 are postponed to Section 3.

Proposition 1. *Suppose* $V_n = (k, ..., k)$.

(i) *Then* $P(T_n > L_n \geq \ell) = E\Big[\big\{1 - (1 - 1/n)^{L_n - 1}\big\}\mathbf{1}_{\{L_n \geq \ell\}}\Big]$. *Consequently,* $P(T_n = L_n) = E\big[(1 - 1/n)^{L_n - 1}\big]$.

(ii) *As* $n \to \infty$, *we have* $P(T_n = L_n) \to 1$ *as* $n \to \infty$.

Recall that Dwass[4] has shown that

$$\lim_{n\to\infty} P\left(\frac{L_n}{n^{k/(k+1)}} \leq x\right) = 1 - \exp\left\{-\frac{x^{k+1}}{(k+1)!}\right\}, \quad \text{for } x > 0. \tag{2}$$

That is, $L_n/n^{k/(k+1)}$ has asymptotically a Weibull distribution with shape parameter $k+1$ and scale parameter $[(k+1)!]^{1/(k+1)}$, which we denote by $\text{Weibull}(k+1, [(k+1)!]^{1/(k+1)})$. Thus, an immediate consequence of Proposition 1(ii) and Equation (2) is as follows.

Theorem 2. *Suppose* $V_n = (k, ..., k)$. *As* $n \to \infty$, *we have*

$$\frac{T_n}{n^{k/(k+1)}} \xrightarrow{d} \text{Weibull}(k+1, [(k+1)!]^{1/(k+1)}).$$

The following theorem considers a more general configuration of the numbers of pairs of shoes initially placed at the n doors.

Theorem 3. *Suppose* $V_n = (v_1^{(n)}, \dots, v_n^{(n)})$ *has the following properties:*

(a) $\sum_{i=1}^{n} v_i^{(n)} = kn.$

(b) $\left|\left\{i : v_i^{(n)} < k\right\}\right| = o(n^{1/(k+1)}),$

where $|A|$ *denote the cardinality number of a set* A*. Then*

$$\frac{T_n}{n^{k/(k+1)}} \xrightarrow{d} \text{Weibull}(k+1, [(k+1)!]^{1/(k+1)}).$$

Notice that Dwass[4] has also shown that for $r = 1, 2, \dots,$

$$\lim_{n\to\infty} E\left[\left(\frac{L_n}{n^{k/(k+1)}}\right)^r\right] = \mu_{k,r} := \int_0^\infty \frac{x^{k+r}}{k!} \exp\left\{-\frac{x^{k+1}}{(k+1)!}\right\} dx, \tag{3}$$

the r-th moment of Weibull $(k+1, [(k+1)!]^{1/(k+1)})$. Theorem 3 suggests that for $r = 1, 2, \dots,$ the limit of $E([T_n/n^{k/(k+1)}]^r | V_n)$ exists and is equal to $\mu_{k,r}$, which is formally stated in the following theorem.

Theorem 4. *Suppose* $V_n = (v_1^{(n)}, \dots, v_n^{(n)})$ *satisfies conditions* (a) *and* (b) *as given in Theorem 3. Then we have*

$$\lim_{n\to\infty} E\left(\left[\frac{T_n}{n^{k/(k+1)}}\right]^r \;\middle|\; V_n\right) = \mu_{k,r}, r = 1, 2, \dots.$$

While Theorems 3 and 4 show under conditions (a) and (b) on V_n that L_n and T_n have the same asymptotic distribution and moments, we have $P(T_n \geq L_n) = 1$ for $V_n = (k, \dots, k)$ so that it is of interest to study the asymptotic behavior of $T_n - L_n$. We next consider the case $k=1$ and derive (with $V_n = (1, \dots, 1)$) the asymptotic (conditional) distribution and moments of $T_n - L_n$ as well as the asymptotic (unconditional) moments of $T_n - L_n$.

Theorem 5. *Let* U *denote a random variable with density*

$$f(x) = \sqrt{\frac{2}{\pi}} \int_0^\infty y^2 (x+y) e^{-\frac{1}{2}(x+y)^2} dy, \quad x > 0.$$

Let $V_n = (1, \dots, 1)$*. Then*

(i) $\mathcal{L}(n^{-1/2}(T_n - L_n) | T_n > L_n, V_n) \xrightarrow{d} U$. *Equivalently, for* $x > 0$

$$\lim_{n\to\infty} P\left(n^{-1/2}(T_n - L_n) > x \mid T_n > L_n, V_n\right) = P(U > x)$$

$$= \sqrt{\frac{2}{\pi}} \int_0^\infty y^2 e^{-\frac{1}{2}(x+y)^2} dy. \tag{4}$$

(ii) $$\lim_{n\to\infty} E\left(\left[\frac{T_n - L_n}{n^{1/2}}\right]^r \middle| T_n > L_n, V_n\right) = E(U^r), \quad r = 1, 2, \ldots. \tag{5}$$

(iii) $$\lim_{n\to\infty} n^{1/2} E\left(\left[\frac{T_n - L_n}{n^{1/2}}\right]^r \middle| V_n\right) = \sqrt{\frac{\pi}{2}} E(U^r), \quad r = 1, 2, \ldots. \tag{6}$$

In particular, for $r = 1$,

$$\lim_{n\to\infty} E(T_n - L_n | V_n) = \sqrt{\frac{\pi}{2}} E(U) = \frac{2}{3}. \tag{7}$$

3. Proofs of Proposition 1 and Theorems 3–5

In the following proofs, it is convenient to adopt the useful notations O_p and o_p, which are defined below for completeness (cf. the item of "Big O in probability notation" in Wikipedia).

Definition 1. For a sequence of random variables $\{X_n\}$ and a sequence of non-zero constants $\{c_n\}$, we write $X_n = O_p(c_n)$ if for any $\epsilon > 0$, there exist finite $M > 0$ and $n_0 \in \mathbb{N}$ such that $P(|X_n/c_n| > M) < \epsilon$ for all $n \geq n_0$; and we write $X_n = o_p(c_n)$ if for any $\epsilon > 0$, $\lim_{n\to\infty} P(|X_n/c_n| > \epsilon) = 0$.

Proof of Proposition 1. (i) Note that $P(T_n \geq L_n) = 1$ and that D_t and D'_t, $t = 1, 2, \ldots$, are all independent and L_n is independent of $D'_t, t = 1, 2, \ldots$. Given $L_n = t$ and $D_t = i$, $T_n = L_n$ if and only if $D'_r \neq i$ for $r = 1, \ldots, t-1$. We have

$$\begin{aligned} P(T_n > L_n | L_n = t, D_t = i) &= 1 - P(D'_r \neq i, r = 1, \ldots, t-1 | L_n = t, D_t = i) \\ &= 1 - P(D'_r \neq i, r = 1, \ldots, t-1) \\ &= 1 - \left(1 - \frac{1}{n}\right)^{t-1}, \end{aligned}$$

so that

$$\begin{aligned} P(T_n > L_n = t) &= \sum_i P(T_n > L_n | L_n = t, D_t = i) P(L_n = t, D_t = i) \\ &= \sum_i \left[1 - \left(1 - \frac{1}{n}\right)^{t-1}\right] P(L_n = t, D_t = i) \\ &= \left[1 - \left(1 - \frac{1}{n}\right)^{t-1}\right] P(L_n = t) \\ &= E\left[\left\{1 - \left(1 - \frac{1}{n}\right)^{L_n - 1}\right\} 1_{\{L_n = t\}}\right]. \end{aligned}$$

Hence

$$P(T_n > L_n \geq \ell) = \sum_{t \geq \ell} P(T_n > L_n = t) = E\left[\left\{1 - \left(1 - \frac{1}{n}\right)^{L_n - 1}\right\} 1_{\{L_n \geq \ell\}}\right].$$

Consequently,

$$P(T_n > L_n) = E\left[1 - \left(1 - \frac{1}{n}\right)^{L_n - 1}\right]$$

and

$$P(T_n = L_n) = 1 - P(T_n > L_n) = E\left[\left(1 - \frac{1}{n}\right)^{L_n - 1}\right].$$

For Part (ii), for any $\epsilon > 0$, we can choose x_0 such that $\exp\left\{-x_0^{k+1}/(k+1)!\right\} < \epsilon/2$ which combining with (2) implies that

$$\lim_{n \to \infty} P\left(\frac{L_n}{n^{k/(k+1)}} > x_0\right) = \exp\left\{-\frac{x_0^{k+1}}{(k+1)!}\right\} < \frac{\epsilon}{2}.$$

So there exists $n_0 \in \mathbb{N}$ such that $P\left(\frac{L_n}{n^{k/(k+1)}} > x_0\right) < \epsilon$ for all $n \geq n_0$. That is, for any $\epsilon > 0$, there exist $x_0 > 0$ and $n_0 \in \mathbb{N}$ such that $P\left(\frac{L_n}{n^{k/(k+1)}} > x_0\right) < \epsilon$ for all $n \geq n_0$. Thus, $L_n = O_p(n^{k/(k+1)})$, implying that

$$\begin{aligned}(L_n - 1) \ln\left(1 - \frac{1}{n}\right) &= \frac{L_n - 1}{n} \ln\left(1 - \frac{1}{n}\right)^n \\ &= O_p(n^{-1/(k+1)}) = o_p(1) \ \ (\text{as } n \to \infty).\end{aligned} \tag{8}$$

Moreover, since $P((1 - 1/n)^{L_n - 1} \leq 1) = 1$ for all n, it follows from part (i), (8) and bounded convergence theorem that

$$P(T_n = L_n) = E\left[\left(1 - \frac{1}{n}\right)^{L_n - 1}\right] = E\left[\exp\left\{(L_n - 1) \ln\left(1 - \frac{1}{n}\right)\right\}\right] \to 1,$$

proving part (ii). □

To prove Theorem 3, we need Lemma 1 and Theorem 6 below, whose proofs are relegated to Appendix.

Lemma 1. *For $1 \leq n' < n$, we have*

$$P(D_i \leq n' \text{ for all } i \leq L_n) = E\left[\left(\frac{n'}{n}\right)^{L_{n'}}\right],$$

where $L_{n'} = \inf\left\{t \geq k+1 : \sum_{r=1}^{t} \mathbf{1}_{\{D_r^* = D_t^*\}} = k+1\right\}$ and where $D_1^*, D_2^*, \ldots$ are i.i.d. uniformly distributed on $\{1, \ldots, n'\}$.

To state Theorem 6, we first define majorization and a stochastic ordering for completeness.

Definition 2. A vector $V_n = (v_1, \ldots, v_n)$ is said to majorize another vector $V'_n = (v'_1, \ldots, v'_n)$ if

$$\sum_{i=1}^{n} v_i = \sum_{i=1}^{n} v_i{}', \text{ and } \sum_{i=1}^{j} v_{(i)} \geq \sum_{i=1}^{j} v_{(i)}{}', j = 1, \ldots, n-1,$$

where $v_{(i)}$ and $v'_{(i)}$ denote the i-th largest components of V_n and V'_n, respectively.

Definition 3. A random variable X is said to be stochastically smaller than another random variable X' if $P(X \leq x) \geq P(X' \leq x)$ for all x.

Notice that majorization provides a partial order and $V = (k, k, \ldots, k)$ is majorized by any other configuration with a total of kn pairs of shoes. Notice also that X is stochastically smaller than X' if and only if for every non-decreasing function f one has $E[f(X)] \leq E[f(X')]$. See e.g. Ref.[5] for the theory of majorization and Ref.[6] for various stochastic orders and their applications.

Theorem 6. *Suppose $V_n = (v_1, \ldots, v_n)$ majorizes $V'_n = (v'_1, \ldots, v'_n)$ (denoted $V_n \succ V'_n$). Then $\mathcal{L}(T_n | V_n)$ is stochastically smaller than $\mathcal{L}(T_n | V'_n)$.*

Proof of Theorem 3. Let $K_n = \left| \left\{ i : v_i^{(n)} < k \right\} \right|$, which is the number of doors with initial numbers of pairs of shoes less than k. Since $\mathcal{L}(T_n | V_n)$ is invariant with respect to permutations of V_n, we assume without loss of generality that $v_i^{(n)} \geq k$ for $i \leq n - K_n$ and $v_i^{(n)} < k$ for $i > n - K_n$, which implies that $\{D_i \leq n - K_n \text{ for all } i \leq L_n\} \subset \{T_n \geq L_n\}$. By Lemma 1 (with $n' = n - K_n$),

$$\begin{aligned} P(T_n \geq L_n) &\geq P(D_i \leq n - K_n \text{ for all } i \leq L_n) \\ &= E\left[\exp\left\{(L_{n-K_n}) \ln\left(1 - \frac{K_n}{n}\right)\right\}\right]. \end{aligned} \tag{9}$$

For any $\epsilon > 0$, let x_0 be such that $\exp\{-x_0^{k+1}/(k+1)!\} < \epsilon/2$. Since assumption (b) implies that $n - K_n \to \infty$, we see that by (2)

$$\lim_{n\to\infty} P\left(\frac{L_{n-K_n}}{(n-K_n)^{k/(k+1)}} > x_0\right) = \exp\left\{-\frac{x_0^{k+1}}{(k+1)!}\right\} < \frac{\epsilon}{2}.$$

It follows that there exists $n_0 \in \mathbb{N}$ such that

$$P\left(\frac{L_{n-K_n}}{(n-K_n)^{k/(k+1)}} > x_0\right) < \epsilon$$

for all $n \geq n_0$. Thus for $n \geq n_0$,

$$P\left(\frac{L_{n-K_n}}{n^{k/(k+1)}} > x_0\right) = P\left(\frac{L_{n-K_n}}{(n-K_n)^{k/(k+1)}} > x_0\left(\frac{n}{n-K_n}\right)^{k/(k+1)}\right)$$
$$\leq P\left(\frac{L_{n-K_n}}{(n-K_n)^{k/(k+1)}} > x_0\right) < \epsilon.$$

This shows that $L_{n-K_n} = O_p(n^{k/(k+1)})$. Moreover,

$$\frac{\ln(1-K_n/n)}{n^{-k/(k+1)}} = \left[\frac{n}{K_n}\ln\left(1-\frac{K_n}{n}\right)\right]\frac{K_n}{n^{1/(k+1)}} \to 0,$$

since $K_n = o(n^{1/(k+1)})$. Thus, we see that

$$L_{n-K_n} = O_p(n^{k/(k+1)}) \quad \text{and} \quad \ln\left(1-\frac{K_n}{n}\right) = o(n^{-k/(k+1)}),$$

which together with (9) implies that

$$P(T_n \geq L_n) \geq E\left[\exp\left\{(L_{n-K_n})\ln\left(1-\frac{K_n}{n}\right)\right\}\right] \to 1,$$
$$\text{i.e.} \quad \lim_{n\to\infty} P(T_n \geq L_n) = 1. \tag{10}$$

Since $V_n \succ V'_n := (k, ..., k)$, we have by Theorem 6 that $\mathcal{L}(T_n|V_n)$ is stochastically smaller than $\mathcal{L}(T_n|V'_n)$, i.e.

$$P\left(\frac{T_n}{n^{k/(k+1)}} \leq x \;\middle|\; V_n\right) \geq P\left(\frac{T_n}{n^{k/(k+1)}} \leq x \;\middle|\; V'_n\right), \quad \text{for } x > 0. \tag{11}$$

By Theorem 2,

$$\lim_{n\to\infty} P\left(\frac{T_n}{n^{k/(k+1)}} \leq x \;\middle|\; V'_n\right) = 1 - \exp\left\{-\frac{x^{k+1}}{(k+1)!}\right\}, \quad \text{for } x > 0. \tag{12}$$

We have for $x > 0$,

$$\limsup_{n\to\infty} P\left(\frac{T_n}{n^{k/(k+1)}} \leq x \;\middle|\; V_n\right) \leq \lim_{n\to\infty} P\left(\frac{L_n}{n^{k/(k+1)}} \leq x\right) \text{ (by (10))}$$
$$= 1 - \exp\left\{-\frac{x^{k+1}}{(k+1)!}\right\} \text{ (by (2))}, \tag{13}$$

and

$$\liminf_{n\to\infty} P\left(\frac{T_n}{n^{k/(k+1)}} \leq x \;\middle|\; V_n\right) \geq \lim_{n\to\infty} P\left(\frac{T_n}{n^{k/(k+1)}} \leq x \;\middle|\; V'_n\right) \text{ (by (11))}$$
$$= 1 - \exp\left\{-\frac{x^{k+1}}{(k+1)!}\right\} \text{ (by (12)).} \tag{14}$$

By (13) and (14), $T_n/n^{k/(k+1)} \xrightarrow{d} \text{Weibull}(k+1, [(k+1)!]^{1/(k+1)})$, completing the proof. □

Proof of Theorem 4. Since by Theorem 3, $\mathcal{L}(T_n/n^{k/(k+1)}|V_n) \xrightarrow{d} \text{Weibull}(k+1, [(k+1)!]^{1/(k+1)})$, an application of Fatou's lemma yields

$$\liminf_{n\to\infty} E\left(\left[\frac{T_n}{n^{k/(k+1)}}\right]^r \middle| V_n\right) \geq \mu_{k,r}. \tag{15}$$

If

$$\limsup_{n\to\infty} E\left(\left[\frac{T_n}{n^{k/(k+1)}}\right]^r \middle| V'_n\right) \leq \mu_{k,r}, \quad r = 1, 2, \ldots, \tag{16}$$

holds where $V'_n = (k, \ldots, k)$, then by Theorem 6,

$$\limsup_{n\to\infty} E\left(\left[\frac{T_n}{n^{k/(k+1)}}\right]^r \middle| V_n\right) \leq \limsup_{n\to\infty} E\left(\left[\frac{T_n}{n^{k/(k+1)}}\right]^r \middle| V'_n\right) \leq \mu_{k,r},$$

which together with (15) implies that $\lim_{n\to\infty} E([T_n/n^{k/(k+1)}]^r|V_n) = \mu_{k,r}$.

It remains to establish the claim (16). In the proof below, note that the initial configuration is $V'_n = (k, \ldots, k)$, so that Theorem 2 can apply. Notice that

$$E\left(\left[\frac{T_n}{n^{k/(k+1)}}\right]^r \middle| V'_n\right) = E\left(\left[\frac{T_n}{n^{k/(k+1)}}\right]^r \mathbf{1}_{\{T_n > L_n\}} \middle| V'_n\right) + E\left(\left[\frac{L_n}{n^{k/(k+1)}}\right]^r \mathbf{1}_{\{T_n = L_n\}} \middle| V'_n\right). \tag{17}$$

Since $T_n = (T_n - L_n) + L_n \leq 2\max\{L_n, T_n - L_n\}$ on $\{T_n > L_n\}$, we have

$$\left[\frac{T_n}{n^{k/(k+1)}}\right]^r \leq 2^r \max\left\{\left[\frac{T_n - L_n}{n^{k/(k+1)}}\right]^r, \left[\frac{L_n}{n^{k/(k+1)}}\right]^r\right\}$$
$$\leq 2^r\left\{\left[\frac{T_n - L_n}{n^{k/(k+1)}}\right]^r + \left[\frac{L_n}{n^{k/(k+1)}}\right]^r\right\} \qquad \text{on } \{T_n > L_n\},$$

which together with (17) implies that

$$E\left(\left[\frac{T_n}{n^{k/(k+1)}}\right]^r \middle| V'_n\right) \le 2^r \left\{E\left(\left[\frac{T_n - L_n}{n^{k/(k+1)}}\right]^r \mathbf{1}_{\{T_n > L_n\}} \middle| V'_n\right) + E\left(\left[\frac{L_n}{n^{k/(k+1)}}\right]^r \mathbf{1}_{\{T_n > L_n\}} \middle| V'_n\right)\right\} + E\left(\left[\frac{L_n}{n^{k/(k+1)}}\right]^r \mathbf{1}_{\{T_n = L_n\}} \middle| V'_n\right). \tag{18}$$

For $t = 1, 2, \ldots,$ let $\mathcal{V}_t$ denote the (random) configuration after the tth walk (and before the $(t+1)$-th walk). Note that $\mathcal{V}_0 = V'_n = (k, \ldots, k)$ is the initial configuration. Let $V^* = \left\{(v_1, \ldots, v_n) : v_i \ge 0 \text{ and } \sum_{i=1}^n v_i = kn\right\}$ be the set of all configurations whose components sum up to kn. For fixed t and fixed $(v_1, \ldots, v_n) \in V^*$, given $T_n > L_n = t$ and $\mathcal{V}_t = (v_1, \ldots, v_n)$, the conditional distribution of $T_n - L_n = T_n - t$ (the number of additional walks up until the first barefoot walk) depends only on the configuration $(v_1, \ldots, v_n)$. This conditional distribution is the same as the distribution of T_n with $(v_1, \ldots, v_n)$ as the initial configuration. (To make it clearer, we may think of the clock being reset after the t-th walk. Then, the configuration after the t-th walk $(v_1, \ldots, v_n)$ becomes the "new" initial configuration, and $T_n - L_n = T_n - t$ is the "new" T_n with $(v_1, \ldots, v_n)$ as the initial configuration.) That is,

$$\mathcal{L}(T_n - L_n | T_n > L_n = t, \mathcal{V}_t = (v_1, \ldots, v_n), V'_n) = \mathcal{L}(T_n | (v_1, \ldots, v_n)),$$

implying that

$$\begin{aligned} E\left(\left[\frac{T_n - L_n}{n^{k/(k+1)}}\right]^r \middle| \, T_n > L_n = t, \mathcal{V}_t = (v_1, \ldots, v_n), V'_n\right) \\ = E\left(\left[\frac{T_n}{n^{k/(k+1)}}\right]^r \middle| \, (v_1, \ldots, v_n)\right) \\ \le E\left(\left[\frac{T_n}{n^{k/(k+1)}}\right]^r \middle| \, V'_n\right), \end{aligned} \tag{19}$$

where the inequality is by Theorem 6 since $V'_n = (k, \ldots, k)$ is majorized by $(v_1, \ldots, v_n)$. Multiplying both sides of (19) by $P(L_n = t, \mathcal{V}_t = (v_1, \ldots, v_n) | T_n > L_n, V'_n)$ and summing over $t \in \{k+1, k+2, \ldots\}$ and $(v_1, \ldots, v_n) \in V^*$ yields

$$E\left(\left[\frac{T_n - L_n}{n^{k/(k+1)}}\right]^r \middle| T_n > L_n, V'_n\right) \le E\left(\left[\frac{T_n}{n^{k/(k+1)}}\right]^r \middle| V'_n\right),$$

which implies that

$$E\left(\left[\frac{T_n - L_n}{n^{k/(k+1)}}\right]^r \mathbf{1}_{\{T_n>L_n\}}\middle|V'_n\right) = P(T_n > L_n|V'_n)E\left(\left[\frac{T_n - L_n}{n^{k/(k+1)}}\right]^r\middle|T_n > L_n, V'_n\right)$$
$$\leq P(T_n > L_n|V'_n)E\left(\left[\frac{T_n}{n^{k/(k+1)}}\right]^r\middle|V'_n\right). \tag{20}$$

Then from (18) to (20), it follows that

$$\begin{aligned}
&\left[1 - 2^r\ P(T_n > L_n|V'_n)\right]\ E\left(\left[\frac{T_n}{n^{k/(k+1)}}\right]^r\middle|V'_n\right)\\
&\leq 2^r\ E\left(\left[\frac{L_n}{n^{k/(k+1)}}\right]^r \mathbf{1}_{\{T_n>L_n\}}\middle|V'_n\right) + E\left(\left[\frac{L_n}{n^{k/(k+1)}}\right]^r \mathbf{1}_{\{T_n=L_n\}}\middle|V'_n\right)\\
&= E\left(\left[\frac{L_n}{n^{k/(k+1)}}\right]^r\middle|V'_n\right) + (2^r - 1)\ E\left(\left[\frac{L_n}{n^{k/(k+1)}}\right]^r \mathbf{1}_{\{T_n>L_n\}}\middle|V'_n\right).
\end{aligned} \tag{21}$$

By Proposition 1(ii), $1 - 2^r P(T_n > L_n) \to 1$, and by Cauchy-Schwarz's inequality,

$$\left\{E\left(\left[\frac{L_n}{n^{k/(k+1)}}\right]^r \mathbf{1}_{\{T_n>L_n\}}\right)\right\}^2 \leq E\left[\frac{L_n}{n^{k/(k+1)}}\right]^{2r} P(T_n > L_n) \to 0.$$

Thus, by (3) and (21), we see that

$$\limsup_{n\to\infty} E\left(\left[\frac{T_n}{n^{k/(k+1)}}\right]^r\middle|V'_n\right) \leq \lim_{n\to\infty} E\left(\left[\frac{L_n}{n^{k/(k+1)}}\right]^r\middle|V'_n\right) = \mu_{k,r},$$

which establishes (16) and the proof is complete. □

Proof of Theorem 5. Assume that part (i) holds. Noting that $V_n = (1, \ldots, 1)$ is majorized by all configurations, we have by Theorem 6 that $\mathcal{L}(T_n - L_n|T_n > L_n, V_n)$ is stochastically smaller than $\mathcal{L}(T_n|V_n)$, implying that for $x > 0$

$$P\left(\left[\frac{T_n - L_n}{n^{1/2}}\right]^r > x\middle|T_n > L_n, V_n\right) \leq P\left(\left[\frac{T_n}{n^{1/2}}\right]^r > x\middle|V_n\right). \tag{22}$$

By (4) and Theorem 3, the left and right sides of (22) converge, respectively, to $P(U^r > x)$ and $P(W^r > x)$, where W has the Weibull $(2, 2^{1/2})$ distribution.

By Theorem 4,

$$\lim_{n\to\infty}\int_0^\infty P\left(\left[\frac{T_n}{n^{1/2}}\right]^r > x\middle| V_n\right)dx = \lim_{n\to\infty} E\left(\left[\frac{T_n}{n^{1/2}}\right]^r\middle| V_n\right)$$
$$= \mu_{1,r} \tag{23}$$
$$= E(W^r)$$
$$= \int_0^\infty P(W^r > x)dx.$$

We have

$$\lim_{n\to\infty} E\left(\left[\frac{T_n - L_n}{n^{1/2}}\right]^r\middle| T_n > L_n, V_n\right)$$
$$= \lim_{n\to\infty}\int_0^\infty P\left(\left[\frac{T_n - L_n}{n^{1/2}}\right]^r > x\middle| T_n > L_n, V_n\right)dx$$
$$= \int_0^\infty P(U^r > x)\ dx$$
$$= E(U^r),$$

where the second equality is due to dominated convergence theorem together with (22) and (23). (More precisely, the left side of (22) is dominated by the right side of (22). Moreover, by (23), the integral of the right side of (22) converges as $n\to\infty$ to the integral of the limit of the right side of (22). Then the second equality follows from dominated convergence theorem. This proves (5).

To show (6), we have

$$\lim_{n\to\infty} n^{1/2}E\left(\left[\frac{T_n - L_n}{n^{1/2}}\right]^r\middle| V_n\right)$$
$$= \lim_{n\to\infty} n^{1/2}P(T_n > L_n|V_n)E\left(\left[\frac{T_n - L_n}{n^{1/2}}\right]^r\middle| T_n > L_n, V_n\right)$$
$$= \lim_{n\to\infty} n^{1/2}P(T_n > L_n|V_n)\ E(U^r)$$
$$= \sqrt{\frac{\pi}{2}}E(U^r),$$

where we have used the fact (cf. (33) below) that $\lim_{n\to\infty} n^{1/2}P(T_n > L_n|V_n) = \sqrt{\pi/2}$. This proves (6). For $r=1$ in (6), we have

$$\begin{aligned}\lim_{n\to\infty} E(T_n - L_n|V_n) &= \sqrt{\frac{\pi}{2}}E(U)\\ &= \sqrt{\frac{\pi}{2}}\int_0^\infty P(U>x)dx\\ &= \int_0^\infty\int_0^\infty y^2 e^{-\frac{1}{2}(x+y)^2}dxdy\\ &= \frac{1}{2}\int_0^\infty\int_0^\infty (x^2+y^2)e^{-\frac{1}{2}(x+y)^2}dxdy\\ &= \frac{1}{2}\int_0^{\frac{\pi}{2}}\int_0^\infty r^2 e^{-\frac{1}{2}r^2(\cos\theta+\sin\theta)^2} rdrd\theta\\ &= \int_0^{\frac{\pi}{2}}\frac{1}{(\cos\theta+\sin\theta)^4}d\theta\\ &= \int_0^{\frac{\pi}{2}}\frac{\sec^4\theta}{(1+\tan\theta)^4}d\theta\\ &= \int_0^\infty \frac{1+u^2}{(1+u)^4}du = \frac{2}{3},\end{aligned}$$

proving (7).

It remains to establish (4). Below for notational simplicity, we will suppress V_n in $P\big(n^{-1/2}(T_n - L_n) > x \mid T_n > L_n, V_n\big)$. Let $S_\ell = \{D_1, ..., D_\ell\}$, the set of the labelings of the doors which the man chooses to go out for the first ℓ walks. On the event $\{T_n > L_n = \ell\}$ (with $\ell \geq 2$), we have that

$$|S_\ell| = \ell - 1; \quad D_t = D_\ell = D'_\tau \in S_\ell \text{ for some } 1 \leq t, \tau < \ell \tag{24}$$

and that before the $(\ell+1)$-th walk, there is at least a pair of shoes available at door i for all $i \notin S_\ell$. Let

$$M_{n,\ell} := \inf\{t \geq 1 : D_{t+\ell} \in S_\ell \text{ or } D_{t+\ell} = D_{t'+\ell} \notin S_\ell \text{ for some } 1 \leq t' < t\}.$$

Plainly, on the event $\{T_n > L_n = \ell\}$, we have $T_n - L_n \geq M_{n,\ell}$. Thus

$$P(T_n - L_n \geq M_{n,L_n} \mid T_n > L_n) = 1. \tag{25}$$

It is not difficult to show that for $t \geq 1$,

$$\begin{aligned}&P(M_{n,\ell} > t \mid |S_\ell| = \ell - 1)\\ &\quad= \left(1 - \frac{\ell-1}{n}\right)\left(1 - \frac{\ell}{n}\right)\cdots\left(1 - \frac{\ell+t-2}{n}\right)\\ &\quad= (1 + O(n^{-1/2}))\exp\left\{-\frac{t(\ell+t/2)}{n}\right\}\\ &\qquad(\text{for } \ell = O(n^{1/2}) \text{ and } t = O(n^{1/2})),\end{aligned} \tag{26}$$

where the $O(n^{-1/2})$ term is uniform in $2 \le \ell \le Cn^{1/2}$ and $1 \le t \le Cn^{1/2}$ for any constant $C > 0$. More precisely (26) is equivalent to

$$\sup_{2\le\ell\le Cn^{1/2},\,1\le t\le Cn^{1/2}} \left|P(M_{n,\ell} > t \mid |S_\ell| = \ell-1)\exp\left\{t(\ell+t/2)/n\right\} - 1\right| = O(n^{-1/2}). \tag{27}$$

For the reader's convenience, we give some details on the derivation of (27). Let $C > 0$ be fixed. Since $\ln(1-x) = -\sum_{n=1}^{\infty}\frac{x^n}{n}$ as $|x| < 1$, it follows that

$$|\ln(1-x)+x| \le \sum_{n=2}^{\infty}\frac{|x|^n}{n} \le x^2\sum_{n=2}^{\infty}\frac{|x|^{n-2}}{2} \le x^2\sum_{n=2}^{\infty}\left(\frac{1}{2}\right)^{n-1} = x^2 \text{ for all } |x| < \frac{1}{2}.$$

Thus, for $n > 16C^2$, $2 \le \ell \le Cn^{1/2}$ and $1 \le t \le Cn^{1/2}$, we have

$$0 < \frac{\ell-1}{n} < \frac{\ell}{n} < \cdots < \frac{\ell+t-2}{n} < \frac{2C}{n^{1/2}} < \frac{1}{2}.$$

Letting $\alpha_{n,\ell,t} = \sum_{i=0}^{t-1}[\ln\left(1-\frac{\ell-1+i}{n}\right)+\frac{\ell-1+i}{n}]$, we have

$$|\alpha_{n,\ell,t}| \le \sum_{i=0}^{t-1}\left(\frac{\ell-1+i}{n}\right)^2 \le \frac{1}{n^2}\int_{\ell-1}^{\ell+t-1}x^2dx \le \frac{(\ell+t-1)^3}{3n^2} \le \frac{8C^3}{3n^{1/2}}.$$

Since $\sum_{i=0}^{t-1}\frac{\ell-1+i}{n} = \frac{t(\ell+t/2)}{n} - \frac{3t}{2n}$, it follows that

$$\left|\left(1-\frac{\ell-1}{n}\right)\left(1-\frac{\ell}{n}\right)\cdots\left(1-\frac{\ell+t-2}{n}\right)\exp\left\{\frac{t(\ell+t/2)}{n}\right\}-1\right| = \left|\exp\left\{\alpha_{n,\ell,t}+\frac{3t}{2n}\right\}-1\right| \le \exp\left\{\frac{8C^3}{3n^{1/2}}+\frac{3C}{2n^{1/2}}\right\}-1 = O(n^{-1/2}),$$

establishing (27).

Note that (26) also holds for $t = 0$. While the left side of (26) is undefined for $\ell = 1$ due to $P(|S_1| = 0) = 0$, it is convenient to let $P(M_{n,1} > t \mid |S_1| = 0) := e^{-t(1+t/2)/n}$, so that (27) remains to hold when the supremum is taken over $1 \le \ell \le Cn^{1/2}$ and $0 \le t \le Cn^{1/2}$. Let $\lceil c \rceil$ denote the smallest integer not less than c. For $x, y > 0$, we have by (26)

$$P\left(n^{-1/2}M_{n,\lceil n^{1/2}y\rceil} > x \;\middle|\; |S_{\lceil n^{1/2}y\rceil}| = \lceil n^{1/2}y\rceil - 1\right) = (1 + O(n^{-1/2}))e^{-x(x/2+y)}, \tag{28}$$

where the $O(n^{-1/2})$ term is uniform in $0 < x, y \le C$ for any constant $C > 0$. Furthermore, we have by Proposition 1(i)

$$\begin{aligned} P(L_n \ge \lceil n^{1/2}y\rceil | T_n > L_n) &= \frac{P(T_n > L_n \ge \lceil n^{1/2}y\rceil)}{P(T_n > L_n)} \\ &= \frac{E\left[\left\{1 - (1-1/n)^{L_n-1}\right\}\mathbf{1}_{\{L_n \ge \lceil n^{1/2}y\rceil\}}\right]}{E\left[1 - (1-1/n)^{L_n-1}\right]}. \end{aligned} \tag{29}$$

Since $\ln(1 - 1/n) > -1/n - 1/n^2$ for all $n \ge 2$ and since $e^x \ge 1 + x$ for all x, we have

$$\begin{aligned} \left(1 - \frac{1}{n}\right)^{L_n-1} &= \exp\left\{(L_n - 1)\ln\left((1 - \frac{1}{n}\right)\right\} \\ &\ge \exp\left\{(L_n - 1)\left(-\frac{1}{n} - \frac{1}{n^2}\right)\right\} \\ &\ge 1 - (L_n - 1)\left(\frac{1}{n} + \frac{1}{n^2}\right), \end{aligned}$$

so that

$$n^{1/2}\left[1 - \left(1 - \frac{1}{n}\right)^{L_n-1}\right] \le \left(\frac{L_n}{n^{1/2}}\right)\left(1 + \frac{1}{n}\right). \tag{30}$$

On the other hand, since $\ln(1 - 1/n) < -1/n$ for all $n \ge 2$ and $e^x \le 1 + x + x^2/2$ for all $x \le 0$, we have

$$\begin{aligned} & n^{1/2}\left[1 - \left(1 - \frac{1}{n}\right)^{L_n-1}\right] \\ &= n^{1/2}\left[1 - \exp\left\{(L_n - 1)\ln\left(1 - \frac{1}{n}\right)\right\}\right] \\ &\ge n^{1/2}\left[1 - \exp\left\{(L_n - 1)\left(-\frac{1}{n}\right)\right\}\right] \\ &\ge n^{1/2}\left[\frac{L_n - 1}{n} - \frac{(L_n - 1)^2}{2n^2}\right] \\ &= \frac{L_n}{n^{1/2}} - \frac{1}{n^{1/2}} - \frac{(L_n/n^{1/2} - 1/n^{1/2})^2}{2n^{1/2}}. \end{aligned} \tag{31}$$

By (2), $L_n/n^{1/2} \xrightarrow{d} W$ where W has the Weibull $(2, 2^{1/2})$ distribution, implying that

$$n^{1/2}\left[1-\left(1-\frac{1}{n}\right)^{L_n-1}\right]=\frac{L_n}{n^{1/2}}+O_p(n^{-1/2})\xrightarrow{d}W. \tag{32}$$

It follows from (29) to (32) and dominated convergence theorem that

$$\begin{aligned} &n^{1/2}E\left[1-\left(1-\frac{1}{n}\right)^{L_n-1}\right]\xrightarrow{n\to\infty}E(W)=\sqrt{\frac{\pi}{2}}\,, \\ &n^{1/2}E\left(\left[1-\left(1-\frac{1}{n}\right)^{L_n-1}\right]\mathbf{1}_{\{L_n\geq\lceil n^{1/2}y\rceil\}}\right)\xrightarrow{n\to\infty}E\left[W\mathbf{1}_{\{W\geq y\}}\right], \\ &P(L_n\geq\lceil n^{1/2}y\rceil|T_n>L_n)\xrightarrow{n\to\infty}\frac{E\left[W\mathbf{1}_{\{W>y\}}\right]}{E(W)}=\sqrt{\frac{2}{\pi}}\int_y^\infty v^2e^{-v^2/2}dv. \end{aligned} \tag{33}$$

So,

$$\mathcal{L}\left(\frac{L_n}{n^{1/2}}\;\middle|\;T_n>L_n\right)\xrightarrow{d}W', \tag{34}$$

where W' has density $\sqrt{\frac{2}{\pi}}y^2e^{-y^2/2}$ for $y>0$. By (28) and (34), for $0<C<\infty$,

$$\begin{aligned} &\lim_{n\to\infty}P(n^{-1/2}M_{n,L_n}>x,L_n\leq Cn^{1/2}|T_n>L_n) \\ &\quad=\lim_{n\to\infty}\sum_{\ell\leq Cn^{1/2}}P(n^{-1/2}M_{n,\ell}>x|T_n>L_n=\ell)P(L_n=\ell|T_n>L_n) \\ &\quad=\lim_{n\to\infty}\sum_{\ell\leq Cn^{1/2}}P(n^{-1/2}M_{n,\ell}>x||S_\ell|=\ell-1)P(L_n=\ell|T_n>L_n) \\ &\quad=\int_0^C e^{-x(x/2+y)}\sqrt{\frac{2}{\pi}}y^2e^{-y^2/2}dy \\ &\quad=\sqrt{\frac{2}{\pi}}\int_0^C y^2e^{-(x+y)^2/2}dy, \end{aligned} \tag{35}$$

where the second equality is a consequence of the result that

$$P(M_{n,\ell}>x|T_n>L_n=\ell)=P(M_{n,\ell}>x||S_\ell|=\ell-1). \tag{36}$$

To show (36), let

$$\begin{aligned} \Gamma_\ell=&\{(I_1,\ldots,I_\ell):1\leq I_1,\ldots,I_{\ell-1}\leq n\text{ are distinct integers} \\ &\text{and }I_\ell=I_j\text{ for some }1\leq j\leq\ell-1\}. \end{aligned} \tag{37}$$

For $\gamma=(I_1,\ldots,I_\ell)\in\Gamma_\ell$, let

$$B_\gamma=\{D_t=I_t,t=1,\ldots,\ell,D'_t=I_\ell\text{ for some }1\leq t\leq\ell-1\}. \tag{38}$$

It is readily seen that $B_\gamma\cap B_{\gamma'}=\emptyset$ for $\gamma\neq\gamma'$ and $\{T_n>L_n=\ell\}=\cup_{\gamma\in\Gamma_\ell}B_\gamma$. Since $M_{n,\ell}$ depends only on the D_t's but not on the D'_t's, we have for

$\gamma = (I_1, \ldots, I_\ell) \in \Gamma_\ell,$

$$P(M_{n,\ell} > x \mid B_\gamma) = P(M_{n,\ell} > x \mid D_t = I_t, t = 1, \ldots, \ell).$$

Moreover, by symmetry, $P(M_{n,\ell} > x \mid D_t = I_t, t = 1, \ldots, \ell)$ is constant for all $\gamma = (I_1, \ldots, I_\ell) \in \Gamma_\ell,$ so that

$$\begin{aligned} P(M_{n,\ell} > x \mid B_\gamma) &= P(M_{n,\ell} > x \mid D_t = I_t, t = 1, \ldots, \ell) \\ &= P(M_{n,\ell} > x \mid D_t = t, t = 1, \ldots, \ell - 1, D_\ell = 1). \end{aligned}$$

It follows that

$$\begin{aligned} P(M_{n,\ell} > x \mid T_n > L_n = \ell) &= \sum_{\gamma \in \Gamma_\ell} P(M_{n,\ell} > x \mid B_\gamma) P(B_\gamma \mid T_n > L_n = \ell) \\ &= P(M_{n,\ell} > x \mid D_t = t, t = 1, \ldots, \ell - 1, D_\ell = 1). \end{aligned}$$

It is also readily seen that

$$P(M_{n,\ell} > x \mid |S_\ell| = \ell - 1) = P(M_{n,\ell} > x \mid D_t = t, t = 1, \ldots, \ell - 1, D_\ell = 1).$$

This proves (36).

While the equation (35) has been shown to hold for all $0 < C < \infty,$ it in fact holds for $C = \infty$ as well. To see this, we have

$$\begin{aligned} &\liminf_{n\to\infty} P(n^{-1/2} M_{n,L_n} > x \mid T_n > L_n) \\ &\qquad \geq \liminf_{n\to\infty} P(n^{-1/2} M_{n,L_n} > x, L_n \leq C n^{1/2} \mid T_n > L_n) \\ &\qquad = \sqrt{\frac{2}{\pi}} \int_0^C y^2 e^{-(x+y)^2/2} dy, \end{aligned}$$

for all $0 < C < \infty,$ implying that

$$\liminf_{n\to\infty} P(n^{-1/2} M_{n,L_n} > x \mid T_n > L_n) \geq \sqrt{\frac{2}{\pi}} \int_0^\infty y^2 e^{-(x+y)^2/2} dy. \tag{39}$$

On the other hand, for all $0 < C < \infty,$

$$\begin{aligned} &\limsup_{n\to\infty} P(n^{-1/2} M_{n,L_n} > x \mid T_n > L_n) \\ &= \limsup_{n\to\infty} \Big\{ P(n^{-1/2} M_{n,L_n} > x, L_n \leq C n^{1/2} \mid T_n > L_n) \\ &\qquad + P(n^{-1/2} M_{n,L_n} > x, L_n > C n^{1/2} \mid T_n > L_n) \Big\} \\ &\leq \limsup_{n\to\infty} \Big\{ P(n^{-1/2} M_{n,L_n} > x, L_n \leq C n^{1/2} \mid T_n > L_n) \\ &\qquad + P(L_n > C n^{1/2} \mid T_n > L_n) \Big\} \\ &= \sqrt{\frac{2}{\pi}} \int_0^C y^2 e^{-(x+y)^2/2} dy + P(W' > C), \end{aligned} \tag{40}$$

where the last equality is due to (34) and (35). Letting $C \to \infty$ in (40) yields

$$\limsup_{n\to\infty} P(n^{-1/2}M_{n,L_n} > x \mid T_n > L_n) \le \sqrt{\frac{2}{\pi}}\int_0^\infty y^2 e^{-(x+y)^2/2}dy,$$

which together with (39) implies that

$$\lim_{n\to\infty} P(n^{-1/2}M_{n,L_n} > x \mid T_n > L_n) = \sqrt{\frac{2}{\pi}}\int_0^\infty y^2 e^{-(x+y)^2/2}dy = P(U > x). \tag{41}$$

We claim that

$$\lim_{n\to\infty} P(T_n - L_n = M_{n,L_n} \mid T_n > L_n) = 1. \tag{42}$$

(Note by (25) that (42) is equivalent to $\lim_{n\to\infty} P(T_n - L_n > M_{n,L_n} \mid T_n > L_n) = 0$.) Then (4) follows from (41) to (42).

It remains to establish the claim (42). We first show that

$$P(D_{L_n} = D_{L_n+M_{n,L_n}} \mid T_n > L_n) \le E\left(\frac{1}{L_n - 1} \;\middle|\; T_n > L_n\right) \xrightarrow{n\to\infty} 0. \tag{43}$$

Let $M'_{n,\ell} := \inf\{t \ge 1 : D_{t+\ell} \in S_\ell\} \ge M_{n,\ell}$. If $M_{n,\ell} < M'_{n,\ell}$, then $D_{\ell+M_{n,\ell}} \notin S_\ell$, implying that $D_{\ell+M_{n,\ell}} \ne D_\ell$ $(\in S_\ell)$. Thus we have

$$\{D_{L_n} = D_{L_n+M_{n,L_n}}\} \subset \{D_{L_n} = D_{L_n+M'_{n,L_n}}\}.$$

It is readily seen that

$$\begin{aligned}
&P(D_{L_n} = D_{L_n+M_{n,L_n}} \mid T_n > L_n)\\
&\quad\le \sum_{\ell=2}^{\infty} P(D_{L_n} = D_{L_n+M'_{n,L_n}}, L_n = \ell \mid T_n > L_n)\\
&\quad= \sum_{\ell=2}^{\infty} P(D_{L_n} = D_{L_n+M'_{n,L_n}} \mid T_n > L_n = \ell)P(L_n = \ell \mid T_n > L_n)\\
&\quad= \sum_{\ell=2}^{\infty} \frac{1}{\ell - 1} P(L_n = \ell \mid T_n > L_n)\\
&\quad= E\left(\frac{1}{L_n - 1} \;\middle|\; T_n > L_n\right),
\end{aligned} \tag{44}$$

where the second equality is a consequence of

$$P(D_{L_n} = D_{L_n+M'_{n,L_n}} \mid T_n > L_n = \ell) = \frac{1}{\ell - 1} \tag{45}$$

To show (45), note by (24) that $|S_{L_n}| = \ell - 1$ if $L_n = \ell$. Recall the definitions of Γ_ℓ and B_γ in (37) and (38), respectively. We have $\{T_n > L_n = \ell\} =$

$\cup_{\gamma\in\Gamma_\ell} B_\gamma$. Since M'_{n,L_n} depends only on the D_t's but not on the D'_t's, we have for $\gamma = (I_1, \ldots, I_\ell) \in \Gamma_\ell$,

$$\begin{aligned}
P(D_{L_n} = D_{L_n+M'_{n,L_n}} \mid B_\gamma) \\
&= P(D_{L_n} = D_{L_n+M'_{n,L_n}} \mid D_t = I_t, t = 1, \ldots, \ell, \\
&\quad \text{and } D'_t = I_\ell \text{ for some } 1 \le t < \ell) \\
&= P(D_{L_n} = D_{L_n+M'_{n,L_n}} \mid D_t = I_t, t = 1, \ldots, \ell) \\
&= \frac{1}{\ell - 1},
\end{aligned}$$

implying that

$$\begin{aligned}
P(D_{L_n} = D_{L_n+M'_{n,L_n}} \mid T_n > L_n = \ell) \\
&= \sum_{\gamma\in\Gamma_\ell} P(D_{L_n} = D_{L_n+M'_{n,L_n}} \mid B_\gamma) P(B_\gamma \mid T_n > L_n = \ell) \\
&= \frac{1}{\ell - 1},
\end{aligned}$$

establishing (45). Noting that $L_n \ge 2$ a.s., we have for any (large) constant $C > 0$,

$$\begin{aligned}
\limsup_{n\to\infty} E\left(\frac{1}{L_n - 1} \;\middle|\; T_n > L_n\right) \\
&\le \limsup_{n\to\infty} \left\{ P(L_n \le C \mid T_n > L_n) + \frac{1}{C-1} P(L_n > C \mid T_n > L_n) \right\} \\
&\le \limsup_{n\to\infty} P(L_n \le C \mid T_n > L_n) + \frac{1}{C-1} \\
&= P(W' = 0) + \frac{1}{C-1} \quad (\text{by } (34)) \\
&= \frac{1}{C-1}.
\end{aligned}$$

Since the upper bound $1/(C-1)$ can be made arbitrarily small, we have

$$\lim_{n\to\infty} E\left(\frac{1}{L_n - 1} \;\middle|\; T_n > L_n\right) = 0,$$

which together with (44) proves (43). By (34, 41) and (43), we have for a sufficiently small (fixed) $\delta > 0$

$$\begin{aligned}
&\lim_{n\to\infty} P(\max\{L_n, M_{n,L_n}\} < n^{1/2+\delta}, D_{L_n} \ne D_{L_n+M_{n,L_n}} \mid T_n > L_n) \\
&\quad = \lim_{n\to\infty} P(\max\{L_n, M_{n,L_n}\} < n^{1/2+\delta} \mid T_n > L_n, D_{L_n} \ne D_{L_n+M_{n,L_n}}) = 1.
\end{aligned} \tag{46}$$

We next show that

$$\lim_{n\to\infty} P(T_n - L_n = M_{n,L_n}, \max\{L_n, M_{n,L_n}\} < n^{1/2+\delta} \mid T_n > L_n, D_{L_n} \neq D_{L_n+M_{n,L_n}}) = 1, \tag{47}$$

which together with (46) implies (42). For $1 \le i \neq j \le n$, denote by $\alpha_n(i,j)$ the probability

$$P(T_n - L_n = M_{n,L_n}, \max\{L_n, M_{n,L_n}\} < n^{1/2+\delta} \mid T_n > L_n, D_{L_n} = i, D_{L_n+M_{n,L_n}} = j).$$

It is easily seen that the value of $\alpha_n(i,j)$ is the same for all pairs of (i, j) with $i \neq j$. It follows that

$$\alpha_n(1,2) = P(T_n - L_n = M_{n,L_n}, \max\{L_n, M_{n,L_n}\} < n^{1/2+\delta} \mid T_n > L_n, D_{L_n} \neq D_{L_n+M_{n,L_n}}).$$

For the same reason, we have by (46)

$$\begin{aligned} &\lim_{n\to\infty} P(\max\{L_n, M_{n,L_n}\} < n^{1/2+\delta} \mid T_n > L_n, D_{L_n} = 1, D_{L_n+M_{n,L_n}} = 2) \\ &= \lim_{n\to\infty} P(\max\{L_n, M_{n,L_n}\} < n^{1/2+\delta} \mid T_n > L_n, D_{L_n} \neq D_{L_n+M_{n,L_n}}) = 1. \end{aligned} \tag{48}$$

It now suffices to show

$$\lim_{n\to\infty} \alpha_n(1,2) = 1, \tag{49}$$

(which implies (47) which in turn implies (42)).

We have

$$\begin{aligned} &\alpha_n(1,2) \\ &= P\Big(T_n - L_n = M_{n,L_n}, \max\{L_n, M_{n,L_n}\} < n^{1/2+\delta} \mid T_n > L_n, D_{L_n} = 1, D_{L_n+M_{n,L_n}} = 2\Big) \\ &= \sum_{\ell, m < n^{1/2+\delta}} P\big(T_n - L_n = M_{n,L_n} = m, L_n = \ell \mid T_n > L_n, D_{L_n} = 1, D_{L_n+M_{n,L_n}} = 2\big) \\ &= \sum_{\ell, m < n^{1/2+\delta}} \beta_n(\ell, m) P\big(M_{n,L_n} = m, L_n = \ell \mid T_n > L_n, D_{L_n} = 1, D_{L_n+M_{n,L_n}} = 2\big), \end{aligned} \tag{50}$$

where

$$\beta_n(\ell, m) := P\big(T_n - L_n = m \mid T_n > L_n, D_{L_n} = 1, D_{L_n+M_{n,L_n}} = 2, L_n = \ell, M_{n,L_n} = m\big).$$

Since the event $\big\{T_n > L_n, D_{L_n} = 1, D_{L_n+M_{n,L_n}} = 2, L_n = \ell, M_{n,L_n} = m\big\}$ is the same as

$$A_n(\ell, m) := \big\{L_n = \ell, M_{n,\ell} = m, D_\ell = 1 = D'_t \text{ for some } t < \ell, D_{\ell+m} = 2\big\},$$

we have

$$\begin{aligned} \beta_n(\ell, m) &= P(T_n - L_n = m \mid A_n(\ell, m)) \\ &= P\big(D'_t \neq 2 \text{ for all } t < \ell + m \mid A_n(\ell, m)\big). \end{aligned}$$

Note that L_n and M_{n,L_n} depend solely on the process $\{D_t\}$, so that

$$\begin{aligned}\beta_n(\ell,m) &= P\left(D'_t \neq 2 \text{ for all } t < \ell + m \mid D'_t = 1 \text{ for some } t < \ell\right)\\ &= \frac{P\left(D'_s \neq 2 \text{ for all } s < \ell + m, D'_t = 1 \text{ for some } t < \ell\right)}{P(D'_t = 1 \text{ for some } t < \ell)}\\ &= \frac{(1-1/n)^m\left[(1-1/n)^{\ell-1} - (1-2/n)^{\ell-1}\right]}{1-(1-1/n)^{\ell-1}}\\ &= \frac{(1-1/n)^m(1-1/n)^{\ell-1}\left[1-(1-1/(n-1))^{\ell-1}\right]}{1-(1-1/n)^{\ell-1}}\\ &= \frac{\left(1+O(n^{-1/2+\delta})\right)\left(1+O(n^{-1/2+\delta})\right)\left(\frac{\ell-1}{n}\right)\left(1+O(n^{-1/2+\delta})\right)}{\left(\frac{\ell-1}{n}\right)\left(1+O(n^{-1/2+\delta})\right)}\\ &= 1+O(n^{-1/2+\delta}),\end{aligned}$$

where the big O terms are all uniform in $\ell, m < n^{1/2+\delta}$. In particular,

$$\lim_{n\to\infty} \min\left\{\beta_n(\ell,m) : \ell, m < n^{1/2+\delta}\right\} = 1. \tag{51}$$

By (50),

$$\begin{aligned}&\alpha_n(1,2)\\ &= \sum_{\ell,m<n^{1/2+\delta}} \beta_n(\ell,m) P\left(M_{n,L_n} = m, L_n = \ell \mid T_n > L_n, D_{L_n} = 1, D_{L_n+M_{n,L_n}} = 2\right)\\ &\geq \min\left\{\beta_n(\ell,m) : \ell, m < n^{1/2+\delta}\right\}\\ &\qquad \times \sum_{\ell,m<n^{1/2+\delta}} P\left(M_{n,L_n} = m, L_n = \ell \mid T_n > L_n, D_{L_n} = 1, D_{L_n+M_{n,L_n}} = 2\right)\\ &= \min\left\{\beta_n(\ell,m) : \ell, m < n^{1/2+\delta}\right\}\\ &\qquad \times P\left(\max\{L_n, M_{n,L_n}\} < n^{1/2+\delta} \mid T_n > L_n, D_{L_n} = 1, D_{L_n+M_{n,L_n}} = 2\right)\\ &\xrightarrow{n\to\infty} 1 \quad (\text{by } (48) \text{ and } (51)),\end{aligned}$$

establishing (49). The proof is complete. □

Appendix

Proof of Lemma 1. For $1 \leq n' < n$, define

$$S_{t,n'} := \left\{(\ell_1,\ldots,\ell_t) : \ell_r \leq n' \text{ for } r = 1,\ldots,t,\ \sum_{r=1}^{s} \mathbf{1}_{\{\ell_r=\ell_s\}} < k+1 \text{ for } s = 1,\ldots,t-1,\right.$$
$$\left.\text{and } \sum_{r=1}^{t} \mathbf{1}_{\{\ell_r=\ell_t\}} = k+1\right\}.$$

Note that

$$\{D_r \le n' \text{ for all } r \le L_n\} = \bigcup_t \{L_n = t, D_r \le n' \text{ for all } r \le t\}$$
$$= \bigcup_t \bigcup_{(\ell_1,\ldots,\ell_t)\in S_{t,n'}} \{D_r = \ell_r, r = 1,\ldots,t\}, \tag{52}$$

and that for $(\ell_1,\ldots,\ell_t) \in S_{t,n'}$,

$$P(D_r = \ell_r, r = 1,\ldots,t) = P(D_r^* = \ell_r, r = 1,\ldots,t)\left(\frac{n'}{n}\right)^t, \tag{53}$$

since $D_1^*, D_2^*, \ldots$ are i.i.d. uniformly distributed on $\{1,\ldots,n'\}$. Also,

$$\{L_{n'} = t\} = \{L_{n'} = t, D_r^* \le n' \text{ for all } r \le t\}$$
$$= \bigcup_{(\ell_1,\ldots,\ell_t)\in S_{t,n'}} \{D_r^* = \ell_r, r = 1,\ldots,t\}. \tag{54}$$

We have

$$\begin{aligned} P(D_r \le n' \text{ for all } r \le L_n) &= \sum_t \sum_{(\ell_1,\ldots,\ell_t)\in S_{t,n'}} P(D_r = \ell_r, r = 1,\ldots,t) \ (\text{by } (52)) \\ &= \sum_t \sum_{(\ell_1,\ldots,\ell_t)\in S_{t,n'}} P(D_r^* = \ell_r, r = 1,\ldots,t)\left(\frac{n'}{n}\right)^t \ (\text{by } (53)) \\ &= \sum_t P(L_{n'} = t)\left(\frac{n'}{n}\right)^t \ (\text{by } (54)) \\ &= E\left[\left(\frac{n'}{n}\right)^{L_{n'}}\right]. \end{aligned}$$

The proof is complete. □

Proof of Theorem 6. By [5, Lemma 2.B.1], it suffices to prove the theorem for the case that V_n and V'_n differ only in 2 components. Recall the assumption that V_n majorizes V'_n. Without loss of generality, assume $v_1 > v'_1 \ge v'_2 > v_2$, and $v_i = v'_i$ for $i = 3,\ldots,n$. In particular, $v_1 \ge 2$ and $v'_1 \ge v'_2 \ge 1$. As a consequence of these assumptions, $v_i > 0$ implies $v'_i > 0$. We will use a coupling device to construct two random variables T and T' on the same probability space in such a way that $T \le T'$ a.s., and $\mathcal{L}(T) = \mathcal{L}(T_n|V_n)$ and $\mathcal{L}(T') = \mathcal{L}(T_n|V'_n)$. Consider two houses (called the V-house and V'-house) each with a different owner, where there are v_i (v'_i, resp.) pairs of shoes available initially at door i of the V-house (V'-house, resp.). Let $K = \sum_i v_i = \sum_i v'_i$, the total number of pairs of shoes available for each owner. Let $D_t, D'_t, t = 1,2,\ldots$ be i.i.d. uniformly distributed on $\{1,\ldots,n\}$ where for *both* houses, D_t denotes the *current* labeling of the door chosen for the t-th walk and D'_t the *current* labeling of the door chosen to leave the shoes after the t-th walk.

In the construction of T and T' below, the labelings of the n doors of the V-house and the labelings of doors i ($i = 3,\ldots,n$) of the V'-house remain the same throughout, while the labelings of doors 1 and 2 of the V'-house may be exchanged at some t *only* when $D_t = 1$ or $D'_t = 2$. For the V'-house, an exchange of the labelings of doors 1 and 2 may be made during the t-th walk if $D_t = 1$ and before the $(t+1)$-th walk if $D'_t = 2$. The total number of exchanges of the labelings of doors 1 and 2 of the V'-house depends on the observed values of $D_t, D'_t, t = 1,2,\ldots$. More details are described below. Then T (T', resp.) denotes the first time the V-house owner (V'-house owner, resp.) discovers that no shoes are available at the door (currently) labeled D_T ($D_{T'}$, resp.) for the T-th (T'-th, resp.) walk.

For the V-house, let $v_i(t)$ ($v_i(t+)$, resp.) be the number of pairs of shoes available at the door initially (and always) labeled i before the t-th walk (during the t-th walk, resp.). The notation $v_i'(t)$ and $v_i'(t+)$, $i = 3, \ldots, n$, is defined similarly for the V'-house. Let $v_i'(t), i = 1, 2$ ($v_i'(t+), i = 1, 2,$ resp.) be the number of pairs of shoes available at the door *currently* labeled i for the V'-house before the t-th walk (during the t-th walk, resp.). Note that $v_i(1) = v_i$ and $v_i'(1) = v_i'$, $i = 1, \ldots, n$ and that

$$\sum_{i=1}^{n} v_i(t) = K \text{ for } t \leq T, \ \sum_{i=1}^{n} v_i(t+) = K - 1 \text{ for } t < T,$$
$$\sum_{i=1}^{n} v_i'(t) = K \text{ for } t \leq T', \ \sum_{i=1}^{n} v_i'(t+) = K - 1 \text{ for } t < T'.$$

We now describe exactly when an exchange of the labelings of doors 1 and 2 of the V'-house is made. Essentially, the labeling exchanges ensure that the majorization relation assumed at the outset continues to hold (until the stopping time $\min\{T, T'\}$). Specifically, for $t = 1$, if $v_{D_1} = 0$ (i.e., no shoes are available at the door labeled D_1 of the V-house), then $T = 1$. In this case, no exchanges of door labelings are needed for the V'-house. Since necessarily $T' \geq 1$, we have $T = 1 \leq T'$ as required. Suppose $v_{D_1} > 0$, implying $T > 1$. Since $v_{D_1} > 0$ implies $v_{D_1}' > 0$, we also have $T' > 1$. During the first walk (or more precisely, before both owners return from the first walk), by exchanging the labelings of doors 1 and 2 of the V'-house if (and only if) $D_1 = 1$ and $v_1' = v_2'$, we have $v_1(1+) \geq v_1'(1+) \geq v_2'(1+) \geq v_2(1+)$ and $v_i(1+) = v_i'(1+), i = 3, \ldots, n$. As a consequence, $V_n(1+) = (v_1(1+), \ldots, v_n(1+)) \succ V_n'(1+) = (v_1'(1+), \ldots, v_n'(1+))$. If $v_1(1+) = v_1'(1+)$ (hence $v_2(1+) = v_2'(1+)$ and $V_n(1+) = V_n'(1+)$), the two configurations are identical and no further labeling exchanges will be made for the V'-house. As the same sequence $D_1', D_2, D_2', \ldots$ applies to both houses, we have $T = T'$. Suppose $v_1(1+) > v_1'(1+) \geq v_2'(1+) > v_2(1+)$. After both owners return from the first walk, by exchanging the labelings of doors 1 and 2 of the V'-house if (and only if) $D_1' = 2$ and $v_1'(1+) = v_2'(1+)$, the numbers of pairs of shoes available at the n doors for both houses (before the second walk) satisfy that $v_1(2) \geq v_1'(2) \geq v_2'(2) \geq v_2(2)$ and $v_i(2) = v_i'(2)$ for $i = 3, \ldots, n$.

More generally, suppose that T and T' are both greater than $t - 1$ and that before the t-th walk, $V_n(t) = (v_1(t), \ldots, v_n(t))$ and $V_n'(t) = (v_1'(t), \ldots, v_n'(t))$ satisfy $v_1(t) \geq v_1'(t) \geq v_2'(t) \geq v_2(t)$ and $v_i(t) = v_i'(t), i = 3, \ldots, n$ (i.e., $V_n(t) \succ V_n'(t)$). If $V_n(t) = V_n'(t)$, then the two configurations are identical and no further labeling exchanges will be made for the V'-house. As the sequence $D_t, D_t', D_{t+1}, \ldots$ applies to both houses, we have $T = T'$. Suppose $v_1(t) > v_1'(t) \geq v_2'(t) > v_2(t)$. If $D_t (\neq 1)$ is such that $v_{D_t}(t) = 0$, then $T = t \leq T'$. (In this case, no further exchanges of door labelings are needed for the V'-house.) Suppose $v_{D_t}(t) > 0$, implying $T > t$. Since $v_{D_t}(t) > 0$ implies $v_{D_t}'(t) > 0$, we also have $T' > t$. Before both owners return from the t-th walk, by exchanging the labelings of doors 1 and 2 of the V'-house if (and only if) $D_t = 1$ and $v_1'(t) = v_2'(t)$, we have $v_1(t+) \geq v_1'(t+) \geq v_2'(t+) \geq v_2(t+)$ and $v_i(t+) = v_i'(t+)$, $i = 3, \ldots, n$. [*Note* 1: Exchanging the labelings (when $D_t = 1$ and $v_1'(t) = v_2'(t)$) does not depend on $D_t', D_{t+1}, D_{t+1}', \ldots$. In particular, each of the n doors of the V'-house is equally likely to be the door currently labeled D_t' where the V'-house owner chooses to leave the shoes upon returning from the t-th walk.] If $v_1(t+) = v_1'(t+)$, then $V_n(t+) = V_n'(t+)$ and the two configurations are identical. No further labeling exchanges will be made for the V'-house. As the sequence $D_t', D_{t+1}, D_{t+1}', \ldots$ applies to both houses, we have $T = T'$. Suppose $v_1(t+) > v_1'(t+) \geq v_2'(t+) > v_2(t+)$. After each owner returns from the t-th walk and leaves shoes at the door currently labeled D_t', by exchanging the labelings of doors 1 and 2 of the V'-house if (and only if) $D_t' = 2$ and $v_1'(t+) = v_2'(t+)$, the numbers of pairs of shoes at the n doors for both houses (before the $(t+1)$-th walk) satisfy that $v_1(t+1) \geq v_1'(t+1) \geq v_2'(t+1) \geq v_2(t+1)$ and $v_i(t+1) = v_i'(t+1)$ for $i = 3, \ldots, n$. [*Note* 2: Exchanging the labelings (when $D_t' = 2$ and $v_1'(t+) = v_2'(t+)$) does not depend on $D_{t+1}, D_{t+1}', D_{t+2}, \ldots$. In particular, each of the n doors of the V'-house is equally likely to be the door currently labeled D_{t+1} which the V'-house owner chooses to go out for the $(t+1)$-th walk.]

The above construction yields that $T \le T'$ and $\mathcal{L}(T) = \mathcal{L}(T_n|V_n)$ and $\mathcal{L}(T') = \mathcal{L}(T_n|V'_n)$. While $\mathcal{L}(T) = \mathcal{L}(T_n|V_n)$ is obvious, $\mathcal{L}(T') = \mathcal{L}(T_n|V'_n)$ follows from Notes 1 and 2 in the preceding paragraph. The proof is complete. □

Remark 1. In the proof of Theorem 6, for the V-house, the labelings of the n doors are fixed. For the V'-house, the labelings of doors $3, \ldots, n$ are also fixed, while the labelings of doors 1 and 2 may be exchanged a number of times in order to preserve the majorization relations $V_n(t) \succ V'_n(t)$ and $V_n(t+) \succ V'_n(t+)$ for all t. Right before the t-th walk, we generate D_t which is uniform on $\{1, \ldots, n\}$ and independent of $D_1, \ldots, D_{t-1}, D'_1, \ldots, D'_{t-1}$. The owner of the V'-house chooses the door currently labeled D_t to walk out. Note that the current labelings of the doors of the V'-house before the t-th walk depend only on $D_1, \ldots, D_{t-1}, D'_1, \ldots, D'_{t-1}$. Since D_t is uniform and independent of $D_1, \ldots, D_{t-1}, D'_1, \ldots, D'_{t-1}$, given the history of the doors chosen for the i-th walk, $i = 1, \ldots, t-1$ and the doors chosen to leave shoes after the i-th walk, $i = 1, \ldots, t-1$, the conditional probability that any door is chosen for the t-th walk equals $1/n$. Next, during the t-th walk, we generate D'_t which is uniform on $\{1, \ldots, n\}$ and independent of $D_1, \ldots, D_t, D'_1, \ldots, D'_{t-1}$. The owner of the V'-house chooses the door currently labeled D'_t to leave shoes upon completing the t-th walk. Note that the the current labelings of the doors of the V'-house during the t-th walk depend only on $D_1, \ldots, D_t, D'_1, \ldots, D'_{t-1}$. Consequently, given the history of the doors chosen for the i-th walk, $i = 1, \ldots, t$ and the doors chosen to leave shoes after the i-th walk, $i = 1, \ldots, t-1$, the conditional probability that any door is chosen to leave shoes after the t-th walk equals $1/n$. This shows that for the V'-house, the doors chosen for the t-th walk, $t = 1, 2, \ldots$ and the doors chosen to leave shoes after the t-th walk, $t = 1, 2, \ldots,$ are i.i.d. with the uniform distribution.

Acknowledgements

We are grateful to the referee for a careful reading of the paper and a number of constructive comments. We gratefully acknowledge support by the Ministry of Science and Technology of Taiwan, R.O.C.

References

[1] Blom, G.; Dou, J.; Popescu, C. P. Elementary problems: E3043. *Am. Math. Monthly* **1984**, *91*, 310. DOI: 10.2307/2322680.

[2] Blom, G.; Holst, L.; Sandell, D. *Problems and Snapshots from the World of Probability*; Springer: New York, **1994**.

[3] DeMaio, P. Bike-sharing: history, impacts, models of provision, and future. *J. Public Transp.* **2009**, *12*, 41–56. DOI: 10.5038/2375-0901.12.4.3.

[4] Dwass, M. More birthday surprises. *J. Comb. Theory* **1969**, *7*, 258–261. DOI: 10.1016/S0021-9800(69)80019-0.

[5] Marshall, A. W.; Olkin, I.; Arnold, B. C. *Inequalities: Theory of Majorization and Its Applications*, 2nd ed.; Springer: New York, **2011**.

[6] Shaked, M.; Shanthikumar, J. G. *Stochastic Orders*; Springer: New York, **2007**.

[7] What is YouBike, https://taipei.youbike.com.tw/about/youbike (accessed Mar 12, 2020).

[8] Zhang, J.; Pan, X.; Li, M.; Anad Yu, P. Bicycle-sharing system analysis and trip prediction. *In* 17th IEEE International Conference on Mobile Data Management, Porto, Portugal, **2016**; 174–179.