ON KNUTH'S GENERALIZATION OF BANACH'S MATCHBOX PROBLEM

By

KEN DUFFY
Communications Network Research Institute, Dublin Institute of Technology

and

W.M.B. DUKES*
School of Theoretical Physics, Dublin Institute for Advanced Studies

[Received 31^{st} October 2002.]

ABSTRACT

We revisit a simply stated problem of Knuth. Previous approaches rely on the Bernoulli nature of the underlying stochastic process to recover the systems mean behavior. We show that limiting results hold for a wide range of stochastic processes. A Large Deviation Principle (LDP) is proved, allowing estimates to be made for the probability of rare–events. From the LDP, a weak law of large numbers is deduced.

1. Introduction

In D. Knuth's 1984 paper [3], he describes a stochastic system which he calls the *'The Toilet Paper Problem'*. He considers the toilet paper dispensers in Stanford University's computer science building, in which each cubicle has two distinct rolls of toilet–paper from which a "user" can choose. Initially there are two toilet–rolls, each with n pieces. He considers the stochastic system where users have two distinct behaviors: a user is a *little–chooser* if they take from the roll with the least number of sheets; a user is a *big–chooser* if they take from the roll with the most number of sheets.

Assuming users arrive to the cubicle independently, use the same unit amount of toilet–roll and are little–choosers with probability p, big–choosers with probability $q := 1 - p$, we wish to know the expected amount of toilet–roll remaining just after one of the two rolls has emptied, called the *residue* R_n. In the case $p = 1/2$, the system reduces to the last problem in the Scottish Book [6], known as Banach's matchbox problem. Judicious use of combinatorial techniques lead Knuth to prove:

$$\mathbb{E}[R_n] = \begin{cases} q/(q-p) + O(r^n) & \text{if } p < q, \\ n(p-q)/p + q/(p-q) + O(r^n) & \text{if } p > q, \end{cases}$$

where r is any value greater than $4pq$. If big–choosers predominate, then on average the system keeps self–leveling and there is very little residue. If little–choosers predominate, then on average one roll drains quicker than the other, leaving a more substantial residue.

*Corresponding author, e-mail: dukes@stp.dias.ie

In D. Stirzaker's 1988 paper, [7], the problem is reconsidered using martingale techniques. This approach enables a solution to the slightly more general problem where the two rolls may be initially unequal in size. He comments:

> These methods provide alternative derivations of Knuth's results. Whether they provide the simple and easy proof he requests is perhaps a matter of taste. This approach does allow quite a few more results to obtained at little extra cost...

With this comment in mind, we reconsider Knuth's problem using large deviations. Let $\{X_i : i \in \mathbb{N}\}$ denote the chooser process: $X_i = +1$ if the i^{th} user is a little–chooser; $X_i = -1$ if the i^{th} user is a big–chooser. Define the partial sum $S_n := \sum_{i=1}^n X_i$. We identify general conditions on $\{X_i\}$ under which the Large Deviation Principle (LDP) and a Weak Law of Large Numbers (WLLN) can be deduced for $\{R_n/n\}$.

In particular, we make two assumptions. The first ensures that a sample–path LDP holds in the topology of uniform convergence. The second is a stability condition.

Assumption 1. *$\{S_n/n\}$ satisfies the LDP with good convex rate–function I. For each fixed $m \in \mathbb{N}$ and $0 = t_0 < t_1 < \cdots < t_m \leq 1$, define $Y_n := (S_n(t_1), S_n(t_2) - S_n(t_1), \ldots, S_n(t_m) - S_n(t_{m-1}))$. $\{Y_n\}$ satisfies the LDP with good rate–function:*

$$I_m(y) := \sum_{i=1}^{m} (t_i - t_{i-1}) I\left(\frac{y_i}{t_i - t_{i-1}}\right),$$

where $y = (y_1, \ldots, y_m)$.

Assumption 2. *The rate–function I is finite on $[0,1]$ and there exists unique $m \in [-1,1]$ such that $I(m) = 0$.*

Using these assumptions, we prove the following LDP:

Proposition 1. *Under assumptions 1 and 2, the residue process $\{R_n/n\}$ satisfies the LDP in $[0,1]$ with good convex rate–function L where, if $m \geq 0$,*

$$L(x) = \begin{cases} (2-x) I\left(\dfrac{x}{2-x}\right) & \text{if } x \in [0,1], \\ +\infty & \text{otherwise,} \end{cases}$$

and if $m < 0$,

$$L(x) = \begin{cases} (2-x) \displaystyle\inf_{a \in [x/(2-x), 1]} a I\left(\dfrac{x}{(2-x)a}\right) & \text{if } x \in [0,1], \\ +\infty & \text{otherwise.} \end{cases}$$

As a corollary to this theorem, the following WLLN is deduced:

Corollary 2. *Under assumptions 1 and 2, $\{R_n\}$ satisfies a weak law of large num-*

bers with

$$\lim_{n\to\infty} \mathbb{E}[R_n/n] = \begin{cases} 2m/(1+m) & \textit{if } m > 0, \\ 0 & \textit{if } m \leq 0. \end{cases}$$

As Knuth's approach involved moment generating function techniques, it is surprising that the Gärtner–Ellis theorem does not appear to provide the best means of tackling this problem. The approach we take uses two distinct methodologies from large deviations: the first involves a relatively routine application of sample–path results; the second method is that of Ruelle and Lanford (a clear exposition of which appears in Lewis and Pfister [4]) developed by Ruelle to give a rigorous treatment of statistical thermodynamics.

The most significant advance this approach admits is that it allows us to extend the little–chooser/big–chooser process from Bernoulli to a range of processes including finite–state irreducible Markov chains. This paper is organized as follows: in Section 2 we give a definition of the LDP; the proof of the main results appears in Section 3; Section 4 contains examples.

2. Large Deviations

Let $\mathcal{X}$ be a Hausdorff topological space with Borel σ–algebra $\mathcal{B}$ and let $\{\mu_n : n \in \mathbb{N}\}$ be a sequence of probability measures on $(\mathcal{X}, \mathcal{B})$. We say that μ_n satisfies the *Large Deviation Principle* (LDP) with rate–function $I : \mathcal{X} \to \mathbb{R}^+ \cup \{+\infty\}$ if I is lower semi–continuous and, for all $B \in \mathcal{B}$,

$$-\inf_{x\in B^\circ} I(x) \leq \liminf_{n\to\infty} \frac{1}{n}\log \mu_n(B) \leq \limsup_{n\to\infty} \frac{1}{n}\log \mu_n(B) \leq -\inf_{x\in \bar{B}} I(x),$$

where B° denotes the interior of B and $\bar{B}$ denotes the closure of B. We say that $\{Z_n\}$ satisfies the LDP if for each n, Z_n is a realization of μ_n. A rate–function is *good* if its level sets $\{x : I(x) \leq \alpha\}$ are compact for all α.

For a general process, $\{Z_n : n \in \mathbb{N}\}$, we define its *sample–path process* $\{\tilde{Z}_n(\cdot)\}$ as follows: for $t \in [0,2]$, let $Z_n(t) := Z_{[nt]}/n$, then

$$\tilde{Z}_n(t) := Z_n(t) + \left(t - \frac{[nt]}{n}\right)\left(Z_n\left(\frac{[nt]+1}{n}\right) - Z_n\left(\frac{[nt]}{n}\right)\right). \tag{2.1}$$

Note that $\tilde{Z}_n(\cdot)$ is a polygonal approximation to $Z_n(\cdot)$. We consider the LDP for sample–paths in $\mathcal{C}[0,2]$, the space of continuous functions on $[0,2]$ equipped with the topology of uniform convergence (whose norm is: $\|\phi\| = \sup_{t\in[0,2]} |\phi(t)|$).

The LDP is a "covariant" principle, in the sence that if μ_n satisfies the LDP in $\mathcal{X}$ with good rate–function I and $f : \mathcal{X} \to \mathcal{Y}$ is continuous, where $\mathcal{X}$ and $\mathcal{Y}$ are Hausdorff, then $\mu_n \circ f^{-1}$ satisfies the LDP in $\mathcal{Y}$ with good rate–function given by $J(y) := \inf\{I(x) : f(x) = y\}$. This fact is called the *Contraction Principle* and a proof can be found in Dembo and Zeitouni, Theorem 4.2.1 [2].

3. Large Deviations for the Residue

Consider the following interpretation of the problem in terms of a random walk with a reflecting barrier. The position of the walk at time $k \geq 0$ is denoted by S_k^* and represents the difference in roll lengths after k users have taken a piece each. If we denote by $\{X_k\}$ the underlying chooser process, taking values in $\{-1, +1\}$, then S_k^* evolves according to the rules:

$$\begin{array}{lcll} S_0^* & := & 0; & \\ S_{k+1}^* & := & S_k^* + X_k, & \text{if } S_k^* > 0; \\ S_{k+1}^* & := & 1, & \text{if } S_k^* = 0. \end{array} \tag{3.1}$$

If $S_k^* = 0$, then the rolls are of equal length and S_{k+1}^* must be 1. Define

$$T_n := \min\{k \in [n, \ldots, 2n-1] : S_k^* = 2n - k\} \tag{3.2}$$

as the number of users that arrive until one of the rolls empties. As each roll initially had n pieces, the residue is given by:

$$R_n = 2n - T_n. \tag{3.3}$$

In order to prove the LDP for $\{R_n/n\}$, we first prove that the sample–path process of the reflected random walk $\{S_n^*/n\}$ satisfies the LDP. We then prove that $\{T_n/n\}$ satisfies the LDP using the Ruelle–Lanford approach. As subtraction is continuous, the LDP for $\{R_n/n\}$ follows by applying the contraction principle.

For each $n \in \mathbb{N}$, define the partial sum of the chooser process, $\{X_k\}$, by

$$S_n := \sum_{k=1}^{n} X_k,$$

and let $\tilde{S}_n(\cdot)$ denote its sample–path.

Theorem 3 (Dembo and Zajic, [1]). *Under Assumption 1, the partial sums sample–path process $\{\tilde{S}_n(\cdot)\}$ satisfies the LDP in $\mathcal{C}[0,2]$ with the good convex rate–function*

$$I_\infty(\phi) = \begin{cases} \int_0^2 I(\dot{\phi}(t))dt & \text{if } \phi \in \mathcal{A}[0,2], \\ +\infty & \text{otherwise}, \end{cases} \tag{3.4}$$

where $\mathcal{A}[0,2]$ is the space of absolutely continuous functions, ϕ, on $[0,2]$ with $\phi(0) = 0$.

We bound the process $\{S_k^*\}$ by two processes which are equivalent on the scale of large deviations; for $i \in \{0,1\}$, define the process $\{S_k^i\}$ by the evolution:

$$\begin{array}{lcll} S_0^i & := & i; & \\ S_{k+1}^i & := & S_k^i + X_k, & \text{if } S_k^i > i; \\ S_{k+1}^i & := & i, & \text{if } S_k^i + X_k \leq i. \end{array}$$

Note that $S^0_k \leq S^*_k \leq S^1_k$ for all $k \in \mathbb{Z}^+$. With $a \vee b := \max\{a, b\}$, by induction,

$$S^i_n = \max_{0 \leq k \leq n-1} \left\{ \sum_{j=n-k}^{n} (X_j) + i \right\} \vee i. \tag{3.5}$$

Theorem 4. *Under Assumption 1, $\{\tilde{S}^0_n(\cdot)\}$ satisfies the LDP in $\mathcal{C}[0,2]$ with good rate–function*

$$J_\infty(\psi) := \inf\{I_\infty(\phi) : f(\phi) = \psi\},$$

where $f : \mathcal{C}[0,2] \to \mathcal{C}[0,2]$ is defined by:

$$(f(\phi))(t) := \phi(t) - \inf_{t^* \in [0,t]} \phi(t^*), \qquad \textit{for } t \in [0,2].$$

PROOF. Note that $f(S_k) = S^0_k$ for all $k \in \mathbb{N}$. That is, at the sample–path level, f represents Equation (3.5) for $i = 0$. As f is continuous, the result follows by invoking the contraction principle. ∎

Two processes $\{X_n\}$ and $\{Y_n\}$ are *exponentially equivalent* (definition 4.2.10 of [2]) if for each $\delta > 0$,

$$\limsup_{n \to \infty} \frac{1}{n} \log \mathbb{P}[|X_n - Y_n| > \delta] = -\infty.$$

Theorem 4.2.13 of [2] proves that if two processes are exponentially equivalent then, if one satisfies the LDP with rate–function I, the other does.

Lemma 5. *The processes $\{\tilde{S}^0_n(\cdot)\}$, $\{\tilde{S}^*_n(\cdot)\}$ and $\{\tilde{S}^1_n(\cdot)\}$ are exponentially equivalent.*

PROOF. $\tilde{S}^0_k \leq \tilde{S}^*_k \leq \tilde{S}^1_k$ and, from (3.5), it is clear that $S^1_k - S^0_k \leq 1$, for all $k \in \mathbb{Z}^+$. Hence

$$\limsup_{n \to \infty} \frac{1}{n} \log \mathbb{P}[|\tilde{S}^0_n(\cdot) - \tilde{S}^1_n(\cdot)| > \delta] = -\infty,$$

for all $\delta > 0$. ∎

From Theorem 4, Lemma 5 and Theorem 4.2.13 of [2] we deduce the following:

Theorem 6. *Under Assumption 1, $\{\tilde{S}^*_n(\cdot)\}$ satisfies the LDP in $\mathcal{C}[0,2]$ with good rate–function J_∞ defined in Theorem 4.*

Corollary 7. *Under Assumptions 1 and 2, $\{S^*_n/n\}$ satisfies the LDP in $[0,1]$ with good convex rate–function J where, if $m \geq 0$,*

$$J(x) = \begin{cases} I(x) & \textit{if } x \in [0,1], \\ +\infty & \textit{otherwise}, \end{cases}$$

and if $m < 0$,

$$J(x) = \begin{cases} \inf_{a \in [x,1]} aI\left(\frac{x}{a}\right) & \textit{if } x \in [0,1], \\ +\infty & \textit{otherwise.} \end{cases}$$

PROOF. As $\psi \to \psi(1)$ is continuous, J is a good rate–function defined by:

$$\begin{aligned} J(x) &:= \inf\{J_\infty(\psi) : \psi(1) = x\} \\ &= \inf\{I_\infty(\phi) : f(\phi)(1) = x\} \\ &= \inf\{I_\infty(\phi) : \phi(1) - \inf_{0 \le t^* \le 1} \phi(t^*) = x\} \\ &\geq \inf_{0 \le t^* \le 1} \inf\{I_\infty(\phi) : \phi(t^*) \le 0, \phi(1) - \phi(t^*) = x\} \\ &\geq \inf_{0 \le t^* \le 1}\{\inf_{h \le 0} t^* I(h/t^*) + (1-t^*)I(x/(1-t^*))\}, \end{aligned}$$

where the last inequality follows by Jensen's inequality. If $m \le 0$, then $\inf_{h \le 0} t^* I(h/t^*) = 0$ and $J(x) \geq \inf_{0 \le t^* \le 1}(1-t^*)I(x/(1-t^*))$. Defining the function

$$\hat{\psi}(t) := mt + \left(\frac{x}{1-t^*} - m\right) \int_0^t 1_{[t^*,1]}(t)dt,$$

gives $f(\hat{\psi})(1) = x$ and $I_\infty(\hat{\psi}) = (1-t^*)I(x/(1-t^*))$. Thus

$$J(x) = \inf_{0 \le t^* \le 1} (1-t^*) I\left(\frac{x}{1-t^*}\right) = \inf_{a \in [x,1]} aI\left(\frac{x}{a}\right),$$

since $I(x) = \infty$ for $x > 1$. If $m = 0$, this infimum occurs at $a = 1$ since I is convex.

If $m > 0$, then $\inf_{h \le 0} t^* I(h/t^*) = t^* I(0) > 0$. Hence $J(x) \geq I(x)$ and equality is obtained by the function $\hat{\psi}(t) := xt$. In both cases, the convexity of J is readily checked.

■

Define the function $\tau : \mathcal{C}[0,2] \to [1,2]$ by $\tau(\phi) := \inf\{t \in [1,2] : \phi(t) = 2 - t\}$. Note that $\tau(\tilde{S}_n^*) = T_n$, where T_n is defined in equation (3.2). Thus τ is the sample–path equivalent of T_n, but is not continuous. For example, consider the sequence of $\mathcal{C}[0,2]$ functions

$$\eta_n(t) := \begin{cases} 0 & \text{if } t \in [0,1/n] \cup [2-1/n, 2], \\ t - 1/n & \text{if } t \in (1/n, 1], \\ 2 - t - 1/n & \text{if } t \in (1, 2-1/n), \end{cases}$$

for $n \geq 1$. The discontinuity arises since $\lim_{n\to\infty} \tau(\eta_n) = 2$, but $\tau(\lim_{n\to\infty} \eta_n) = 1$ and hence the contraction principle cannot be invoked. In order to prove the LDP for $\{T_n/n\}$, we need extra information from the process that led to the LDP for $\{\tilde{S}_n^*(\cdot)\}$. We use the Ruelle–Lanford approach.

Theorem 8. *Under Assumptions 1 and 2,* $\{T_n/n\}$ *satisfies the LDP with good*

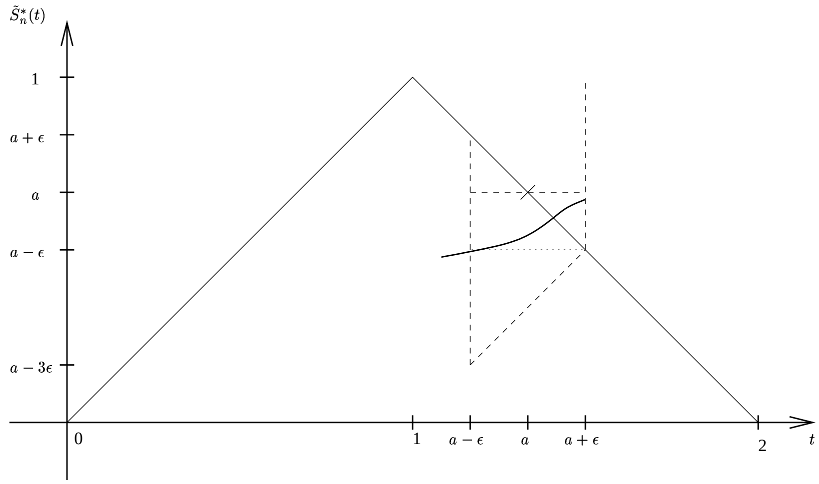

FIG. 1—Equivalence of events.

convex rate-function

$$K(x) := \begin{cases} xJ\left(\dfrac{2-x}{x}\right) & \text{if } x \in [1,2], \\ +\infty & \text{otherwise,} \end{cases}$$

where J is defined in Corollary 7.

PROOF. As T_n/n takes values within the compact interval $[1,2]$, if one can prove the large deviations upper bound for compact sets, then it holds for all closed sets. Thus, in order to prove the full LDP, by Theorem 3.1 of [4] (or Theorem 4.1.11 of [2]), it suffices to show that for all $a \in [1,2]$

$$\lim_{\epsilon\to 0}\liminf_{n\to\infty}\frac{1}{n}\log\mathbb{P}\left[\frac{T_n}{n}\in B_\epsilon(a)\right] = \lim_{\epsilon\to 0}\limsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}\left[\frac{T_n}{n}\in B_\epsilon(a)\right],$$

where $B_\epsilon(a) := (a-\epsilon, a+\epsilon)$.

For $a > \epsilon > 0$, consider the following characterization of the event $\{T_n/n \in B_\epsilon(a)\}$, bearing Figure 1 in mind,

- At $a-\epsilon$, $\tilde{S}^*_n(\cdot)$ must not have crossed, or touched, the line $2-x$. If it did cross or touch the line, it could not return below it and the first crossing would have occurred before $a-\epsilon$. However, at $a-\epsilon$, $\tilde{S}^*_n(\cdot)$ must be no further than 4ϵ below the line or else it could not cross until after $a+\epsilon$.
- At $a+\epsilon$, the process $\tilde{S}^*_n(\cdot)$ must be on or above the line $2-x$, in order that the first crossing have occurred before $a+\epsilon$.

Formally, for $0 < \epsilon < a$, the event $\{T_n/n \in B_\epsilon(a)\}$ is equivalent to

$$\{2-[n(a+\epsilon)]/n < \tilde{S}^*_n(a-\epsilon) < 2-[n(a-3\epsilon)]/n\}\cap\{\tilde{S}^*_n(a+\epsilon) \geq 2-[n(a+\epsilon)]/n\}. \quad (3.6)$$

For the upper bound, drop the second term in (3.6),

$$\mathbb{P}\left[\frac{T_n}{n} \in B_\epsilon(a)\right] \leq \mathbb{P}\left[\tilde{S}_n^*(a-\epsilon) \in \left(2 - \frac{[n(a-3\epsilon)]}{n}, 2 - \frac{[n(a+\epsilon)]}{n}\right)\right].$$

Define the sets

$$\begin{aligned} A'(a,\epsilon) &:= ((2-a-\epsilon)/(a-\epsilon), (2-a+3\epsilon)/(a-\epsilon)), \\ A(a,\epsilon) &:= \{\phi : \phi(a-\epsilon) \in A'(a,\epsilon)\}. \end{aligned}$$

Taking limits gives,

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left[\frac{T_n}{n} \in B_\epsilon(a)\right] \leq -\inf_{\phi \in A(a,\epsilon)} J_\infty(\phi) = -(a-\epsilon) \inf_{x \in A'(a,\epsilon)} J(x),$$

where J is defined in the statement of Corollary 7. Thus, as J is continuous on $[0,1]$,

$$\lim_{\epsilon\to 0} \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left[\frac{T_n}{n} \in B_\epsilon(a)\right] \leq -aJ\left(\frac{2-a}{a}\right) \qquad \text{for all } a \in [1,2].$$

If $a = 0$, the lower bound is trivially equivalent to the upper bound. For $a > \epsilon > 0$, the set defined in (3.6) contains the set

$$\{2 - [n(a+\epsilon)]/n < \tilde{S}_n^*(a-\epsilon) < 2 - [n(a-3\epsilon)]/n\} \cap \{\tilde{S}_n^*(a+\epsilon) - \tilde{S}_n^*(a-\epsilon) \geq 2\epsilon\}.$$

Define the set

$$B := \{\phi : \phi(a+\epsilon) - \phi(a-\epsilon) \geq 2\epsilon\}.$$

Again, taking limits gives,

$$\begin{aligned} \liminf_{n\to\infty} n^{-1} \log \mathbb{P}[T_n/n \in B_\epsilon(a)] &\geq -\inf_{\phi \in A \cap B} J_\infty(\phi) \\ &= -\inf\{I_\infty(\psi) : f(\psi)(a-\epsilon) \in A, \\ &\qquad \psi(a+\epsilon) - \psi(a-\epsilon) \geq 2\epsilon\} \\ &= -\inf_{\phi \in A} J_\infty(\phi) - 2\epsilon I(1). \end{aligned}$$

and so, as J is continuous on $[0,1]$,

$$\lim_{\epsilon\to 0} \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left[\frac{T_n}{n} \in B_\epsilon(a)\right] \geq -aJ\left(\frac{2-a}{a}\right) \qquad \text{for all } a \in [1,2].$$

■

Finally, as subtraction is continuous and $R_n = 2n - T_n$, we have the main result for the residue process:

Proposition 9. *Under assumptions 1 and 2, the residue process* $\{R_n/n\}$ *satisfies* *the LDP in* $[0,1]$ *with good convex rate–function* L *where, if* $m \geq 0$,

$$L(x) = \begin{cases} (2-x)I\left(\dfrac{x}{2-x}\right) & \text{if } x \in [0,1], \\ +\infty & \text{otherwise,} \end{cases}$$

and, if $m < 0$,

$$L(x) = \begin{cases} (2-x) \inf\limits_{a \in [x/(2-x),1]} aI\left(\dfrac{x}{(2-x)a}\right) & \text{if } x \in [0,1], \\ +\infty & \text{otherwise.} \end{cases}$$

Corollary 10. *Under assumptions 1 and 2,* $\{R_n\}$ *satisfies a weak law of large numbers with*

$$\lim_{n \to \infty} \mathbb{E}[R_n/n] = \begin{cases} 2m/(1+m) & \text{if } m > 0, \\ 0 & \text{if } m \leq 0, \end{cases}$$

where $m \in [-1,1]$ *is the unique point with* $I(m) = 0$.

PROOF. By Theorems 2.1 and 2.2 of Lewis, Pfister and Sullivan [5], in order to determine the WLLN from the LDP, it suffices to determine the unique zero of L. Straightforward calculation reveals the values given above. ∎

4. Examples

The first example is the original problem of Knuth, where users arrive in an i.i.d. Bernoulli manner.

Example 1. *Let* $\{X_n : n \in \mathbb{N}\}$ *to be an i.i.d Bernoulli sequence taking values in* $\{-1,1\}$ *with the probability of user* i *being a little–chooser* $\mathbb{P}[X_i = +1] = p \in (0,1)$ *and probability* $q := 1-p$ *of being a big–chooser.* $\{X_n\}$ *satisfies Assumption 1 with rate–function:*

$$I_p(x) := \begin{cases} \dfrac{1-x}{2}\log\left(\dfrac{1-x}{2q}\right) + \dfrac{1+x}{2}\log\left(\dfrac{1+x}{2p}\right) & \text{if } x \in [-1,1], \\ +\infty & \text{otherwise,} \end{cases}$$

which satisfies Assumption 2 with $I_p(m) = 0$ *for* $m = \mathbb{E}[X_i] = p-q$. *By Proposition 9 and Corollary 7, the rate–function for* $\{R_n/n\}$ *is given by: if* $p \geq 1/2$,

$$L_p(x) = \begin{cases} (1-x)\log\left(\dfrac{(1-x)}{q(2-x)}\right) + \log\left(\dfrac{1}{p(2-x)}\right) & \text{if } x \in [0,1], \\ +\infty & \text{otherwise,} \end{cases}$$

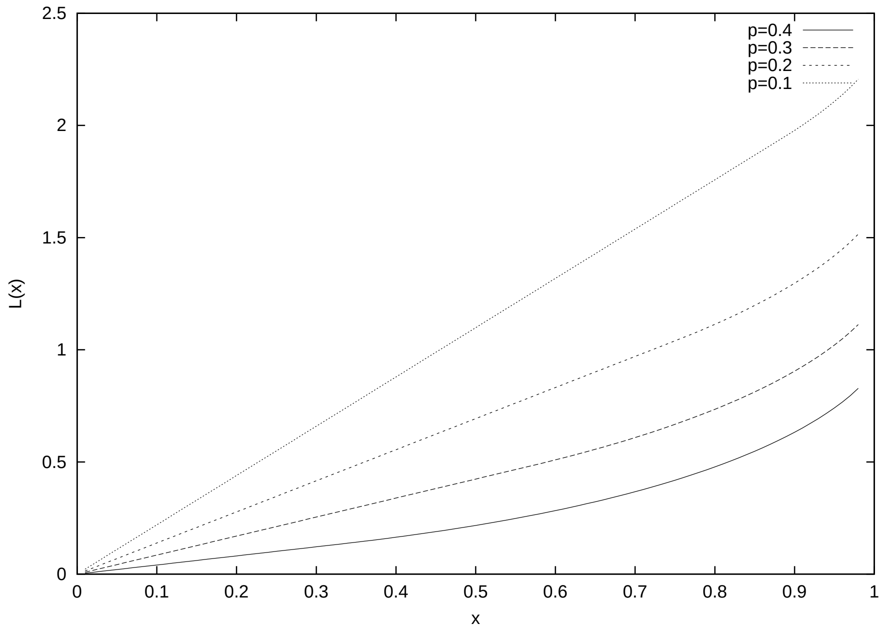

FIG. 2—Bernoulli–chooser rate–functions, $p \leq 1/2$.

and, if $p < 1/2$,

$$L_p(x) = \begin{cases} x \log\left(\frac{q}{p}\right) & \textit{if } x \in [0, (q-p)/q], \\ (1-x)\log\left(\frac{(1-x)}{q(2-x)}\right) + \log\left(\frac{1}{p(2-x)}\right) & \textit{if } x \in ((q-p)/q, 1], \\ +\infty & \textit{otherwise.} \end{cases}$$

Figures 2 and 3 show graphs of $L_p(x)$ *for a range of values of* x *and* p*. The y–range for the Figure 3 has been truncated to make the detail clearer. By Corollary 10, the mean behavior of* R_n/n *is determined by:*

$$\lim_{n\to\infty} \mathbb{E}[R_n/n] = \begin{cases} (p-q)/p & \textit{if } p > 1/2, \\ 0 & \textit{if } p \leq 1/2. \end{cases}$$

The second example is a function of a Markov chain, in which users may use more than one sheet.

Example 2. *Extend the users' behavior so that during a single visit to the cubicle, a big–chooser may use* 1 *to* $N(> 1)$ *sheets and a little–chooser use* 1 *to* $M(> 1)$*. Model this situation by introducing the irreducible recurrent Markov chain* Y_n *on*

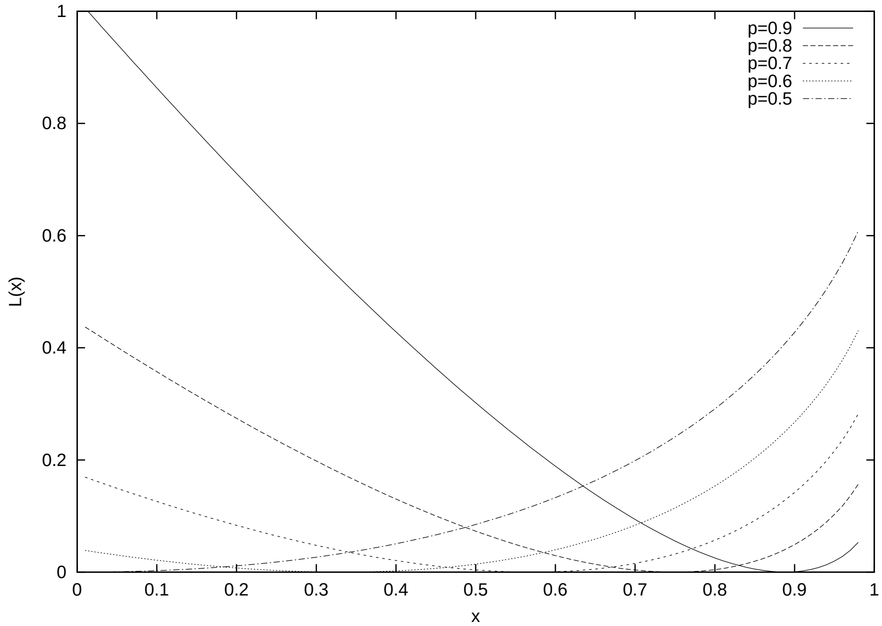

FIG. 3—Bernoulli–chooser rate–functions, $p > 1/2$.

$\mathcal{F} = [-N, -N+1, \ldots, -1, 1, 2, \ldots, M]$ *with transition matrix*

$$\Pi(i,j) = \begin{cases} 1/(N+M) & \textit{if } i \in \{-1,+1\}, \\ 1 & \textit{if } i > +1 \textit{ and } i-j=1, \\ 1 & \textit{if } i < -1 \textit{ and } j-i=1, \\ 0 & \textit{otherwise.} \end{cases} \tag{4.1}$$

After each piece is taken, $|Y_n|$ counts down, remembering how many more pieces a user wishes to take. After a user has taken his last piece, $|Y_n| = 1$ and the next users' behavior is chosen uniformly from the $N + M$ possible behaviors. Let $f : F \to \{-1, +1\}$ be defined by

$$f(k) = \begin{cases} +1 & \textit{if } k > 0, \\ -1 & \textit{if } k < 0. \end{cases}$$

The piece-wise selection process is defined by $X_i := f(Y_i)$.

By Theorem 3.1.2 of [2], $\{S_n/n\}$ satisfies the LDP in [0,1] with good convex rate–function, I. Although it is not possible to determine a closed form for I, for fixed values of N and M, it is readily calculated numerically (see section 3.1 of [2]). By Theorem 3 of [1], its sample–paths, $\{\tilde{S}_n\}$, satisfy the LDP in $\mathcal{C}[0,2]$ with rate–function, I_∞, defined in Equation (3.4).

The unit left eigenvector, $\vec{e}$, of Π determines the mean behavior of Y_n. Define

$c := (N(N+1) + M(M+1))/2$, *then*

$$\vec{e} = \left(\frac{1}{c}, \frac{2}{c}, \ldots, \frac{N}{c}, \frac{M}{c}, \ldots, \frac{2}{c}, \frac{1}{c}\right).$$

The concentration point, m in Assumption 2, for $\{S_n/n\}$ is given by the expectation of f with respect to $\vec{e}$,

$$m = \frac{M(M+1) - N(N+1)}{M(M+1) + N(N+1)}.$$

Clearly, $m \leq 0$ if and only if $M \leq N$. Defining $\gamma := M(M+1)/(N(N+1))$, $m = (1-\gamma)/(1+\gamma))$ and, using Corollary 10, the mean behavior of R_n/n is determined by:

$$\lim_{n\to\infty} \mathbb{E}[R_n/n] = \begin{cases} 1-\gamma & \text{if } M > N \text{ (equivalently } \gamma > 0\text{)}, \\ 0 & \text{if } M \leq N \text{ (equivalently } \gamma \leq 0\text{)}. \end{cases}$$

Acknowledgement

K.D's work supported by Science Foundation Ireland under the National Development Plan.

References

[1] A. Dembo and T. Zajic, 1995 Large deviations: from empirical mean and measure to partial sums. *Stochastic Processes and their Applications* **57**, 191–224.
[2] A. Dembo and O. Zeitouni, 1998 *Large deviation techniques and applications.* Springer.
[3] D.E. Knuth, 1984 The toilet paper problem. *American Mathematical Monthly* **91**, 465–470.
[4] J. T. Lewis and C. E. Pfister, 1995 Thermodynamic probability theory: some aspects of large deviations. *Russian Mathematical Surveys* **50**, no. 2, 279–317.
[5] J. T. Lewis, C. E. Pfister, and W. G. Sullivan, 1995 Entropy, concentration of probability and conditional limit theorems. *Markov Processes and Related Fields* **1**, 319–386.
[6] R.D. Mauldin (ED.), 1981 *The Scottish Book.* Birkhauser. Boston.
[7] David Stirzaker, 1988 A generalization of the matchbox problem. *Mathematical Scientist* **13**, 104–114.