RENEWAL-TYPE BEHAVIOR OF ABSORPTION TIMES IN MARKOV CHAINS

BERNARD VAN CUTSEM* AND
BERNARD YCART*, Institut IMAG

Abstract
This paper studies the absorption time of an integer-valued Markov chain with a lower-triangular transition matrix. The main results concern the asymptotic behavior of the absorption time when the starting point tends to infinity (asymptotics of moments and central limit theorem). They are obtained using stochastic comparison for Markov chains and the classical theorems of renewal theory. Applications to the description of large random chains of partitions and large random ordered partitions are given.

RENEWAL PROCESSES; RANDOM CHAINS OF PARTITIONS; RANDOM ORDERED PARTITIONS

AMS 1991 SUBJECT CLASSIFICATION: PRIMARY 60K20
SECONDARY 05A18

1. Introduction
The application that motivated this study relates to stratified hierarchies, commonly used as representations of classified data (cf. Benzczi (1973)). In such a representation a collection of \( n \) objects is associated to a chain of partitions of the \( n \)-set. Such a chain is a strictly monotone sequence of partitions of the objects, from \( \{\{1\}, \{2\}, \ldots, \{n\}\} \) to \( \{\{1,2,\ldots,n\}\} \). The level of each partition is by definition its rank in the sequence. The first partition \( \{\{1\}, \{2\}, \ldots, \{n\}\} \) has level 0 and the partition at any further level is formed by taking unions of subsets in the previous one. The length of the chain is the total number of partitions minus 1. It is easy to see that the number \( z(n) \) of chains on an \( n \)-set (or of stratified hierarchies of \( n \) objects) can be defined recursively by \( z(1) = 1 \) and

\[
z(n) = \sum_{k=1}^{n-1} S(n, k)z(k) \quad \forall n \geq 2,
\]

where \( S(n, k) \) is the Stirling number of the second kind (cf. Lengyel (1984)). Since the Stirling number of the second kind \( S(n, k) \) represents the number of partitions of an \( n \)-set into \( k \) subsets, the meaning of the above recurrence is clear: at level 1 of the chain, the partition coming immediately after \( \{\{1\}, \{2\}, \ldots, \{n\}\} \) has \( k \) subsets

* Postal address: LMC/IMAG B.P. 53, 38041 Grenoble Cedex 9, France. e-mail: ycart@imag.fr.
(1 \leq k \leq n - 1) and the rest of the chain can be viewed as a new chain based on those k elements.

We call the random chain of partitions a random variable uniformly distributed on the set of all chains of partitions on an n-set (each of them has probability 1/z(n)). For such a random chain, let X_i be the cardinality of the partition at level i (X_0 = n). Translated in terms of random chains, the recurrence above says that the sequence of random variables (X_i),_{i=0} is a Markov chain with transition matrix P = (\pi_{ij}) with

\[ \pi_{ij} = S(i, j)z(j)/z(i) \quad \forall i = 1, \cdots, n \quad \forall j = 1, \cdots, n - 1. \]

Any Markov chain with this transition matrix will be eventually absorbed at 1. We denote by H_n the absorption time, starting from n:

\[ H_n = \inf \{i \text{ such that } X_i = 1, \text{ when } X_0 = n\}. \]

The absorption time H_n is exactly the length of the random chain of partitions (total number of partitions in the chain minus 1). (We are aware of the unfortunate collision of denomination between chains of partitions and Markov chains.) The main object of this paper is to describe the asymptotic distribution of H_n as n goes to infinity.

Of course the distribution of H_n can be computed explicitly (cf. for instance Barucha-Reid (1960), p. 17). In our case of a triangular transition matrix, the computation is very simple. One has the following recurrence:

\[ P[H_n = 1] = \pi_{n1} \quad \text{and} \quad P[H_n = k] = \sum_{j=2}^{n-1} \pi_{nj} P[H_j = k - 1] \quad \text{for } k = 2, \cdots, n - 1. \]

Unfortunately this formula is of no practical use for large values of n. In our application to chains of partitions, the number z(100), necessary to compute the 100th line of the transition matrix, already has 300 digits. Thus asymptotic results on H_n have to be used. We shall come back to this application and some others in Section 5.

Let us describe now our general setting. We consider a sequence (p_n)_{n=1} of probability distributions on the set of integers \( \mathbb{N} \) such that:

\[ (1.1) \quad p_1(\{0\}) = 1 \quad \text{and} \quad p_n(\{1, \cdots, n-1\}) = 1 \quad \forall n \geq 1. \]
Let $P = (\pi_{nk})_{n,k \geq 1}$ be the transition matrix defined by

$$
\pi_{nk} = \begin{cases} 
  p_n(n-k), & \text{if } n \geq 1 \quad 1 \leq k \leq n-1 \\
  0, & \text{in any other case.}
\end{cases}
$$

Thus $P$ is the following matrix:

$$
\begin{pmatrix}
1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 \\
p_3(2) & p_3(1) & 0 & 0 & 0 \\
p_4(3) & p_4(2) & p_4(1) & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
p_n(n-1) & p_n(2) & p_n(1) & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots 
\end{pmatrix}
$$

If a Markov chain $(X_i)_{i \geq 0}$ has such a transition matrix, then whatever $X_0$, the chain will be absorbed eventually at 1. We shall still denote by $H_n$ its absorption time, starting from state $n$. By reversing the space of states, one can interpret $H_n$ as the first passage time through level $n$ of a Markov chain starting at 1. Such processes have been considered as a generalization of the classical renewal process by Kijima and Sumita (1986).

Our main tool in this study is stochastic comparison. Stoyan’s book (1983) is a good introduction to the subject and to the wide range of its applications. With integer-valued variables, we are considering a very simple case.

**Definition 1.2.** Let $p$ and $q$ be two probability distributions on $\mathbb{N}$. We say that $p$ is stochastically inferior to $q$ iff

$$
p([0, \cdots, k]) \leq q([0, \cdots, k]) \quad \forall k \in \mathbb{N}.
$$

This property will be denoted by $p \ll q$.

Our main assumption is that each $p_n$ is stochastically dominated by the distribution of a square integrable random variable, and that the sequence $(p_n)_{n \geq 1}$ converges to some probability distribution on $\mathbb{N}$.

**Hypothesis 1.3.** There exist two probability distributions $p$ and $q$ on $\mathbb{N}$ such that

(a) $\sum_{k=1}^{\infty} k^2 q(k) < \infty$,  

(b) $p_n \ll q \quad \forall n \geq 1$,  

(c) $\lim_{n \to \infty} p_n(k) = p(k) \quad \forall k \in \mathbb{N}$.

As we shall see in Section 5, this hypothesis is fulfilled in the application to random chains of partitions, with

$$
p(k) = \log (2)^k / k! \quad \forall k \geq 1.
$$
Our idea is to compare the behavior of $H_n$ to that of a random walk driven by the distribution $p$. Let $(T_i)_{i=1}^\infty$ be a sequence of i.i.d. random variables with common distribution $p$. Let $X^*_t = n - T_1 - \cdots - T_t$ be the backward random walk starting at $n$, with steps $T_i$. Denote by $H^*_n$ the crossing time of level 1 for this random walk.

$$H^*_n = \inf \{ i \text{ such that } n - T_1 - \cdots - T_i \leq 1 \}.$$ 

It is intuitively clear that for large $n$'s most steps of the Markov chain $(X_t)$, starting at $X_0 = n$, will have a distribution close to $p$, hence $(X_t)$ should behave approximately as the random walk $(X^*_t)$ on the major part of its journey. Therefore one can expect that $H_n$ will behave asymptotically as $H^*_n$. We propose to make this intuition precise.

As a renewal counting process, the asymptotic behavior of $H^*_n$ is well known. The main classical results, as found for instance in Feller (1949), will be reviewed briefly in Section 2 (Theorem 2.1). The classical results on stochastic comparison that we shall need are recalled in Section 3. Section 4 is devoted to the main result (Theorem 4.1). The results on renewal counting processes that we extend to the absorption time $H_n$ are the asymptotic behavior of moments and the central limit theorem. The key observation in the proof of Theorem 4.1 will be that all the distributions $p_n$ for $n$ large enough can be bounded stochastically above and below by two probability distributions that are arbitrarily close (cf. Lemma 4.2). In Section 5 we shall present some applications to combinatorics. They concern in particular the asymptotic distribution of the length of a random chain of partitions mentioned above, and that of a random ordered partition. Finally we shall discuss in Section 6 some extensions of Theorem 4.1, in particular to non-triangular matrices. The importance of Hypothesis 1.3 will also be illustrated on a family of examples.

2. Elementary results for renewal counting processes

We shall be concerned here with the simplest case of the renewal counting process for sums of i.i.d. discrete positive random variables. Let $(T_i)_{i=1}^\infty$ be a sequence of i.i.d. random variables. Their common distribution is assumed to be concentrated on the set of positive integers. Their common expectation $\mu$ and variance $\sigma^2$ are finite. For any positive integer $n$ we define

$$N_n = \inf \{ i \text{ such that } T_1 + \cdots + T_i \geq n \}.$$ 

The study of the asymptotic behavior of the sequence $(N_n)_{n=1}^\infty$ began in the 1940s with the papers of Feller (1941), (1949) and Doob (1948). We shall recall here only the most basic limit theorems, essentially those of Feller (1949). Many generalizations and refinements have been proposed, and we shall not review them. The interested reader may consult Gut (1974), or Roginsky (1992) (cf. also Chow and Teicher (1978)). The results stated below concern the asymptotics of moments, the central limit theorem and the law of iterated logarithm.
Theorem 2.1

(a) \( \forall r \leq 2 \lim_{n \to \infty} E[(N_n/n)^r] = (1/\mu)^r. \)

(b) \( \lim_{n \to \infty} \mathcal{L}[(N_n - n/\mu)/(\sigma^2/\mu^3)^{1/2}] = N(0, 1). \)

(c) \( \limsup_{n \to \infty} |N_n - n/\mu|/((2\sigma^2/\mu^3)n \log n)^{1/2} = 1 \) a.s.

Part (a) corresponds to Theorem 9 on p. 111 of Feller (1949) (cf. also Gut (1974), Theorem 2.3 on p. 281, Theorem 2.6 and 2.7 on p. 287). Chow and Robbins (1963) generalized it for \( r = 1 \) to the case of possibly dependent \( X_i \)'s, under the assumption that the conditional expectation of \( X_i \) knowing \( X_1, \ldots, X_{i-1} \) does not depend actually on the past, so that their result cannot be applied to our absorption times.

Part (b) can be deduced straightforwardly from the central limit theorem (cf. Feller (1949), Theorem 5, p. 105). Much finer results are now available, including brownian approximations (cf. Csörgő et al. (1987)) and sharp estimates for the remainder terms (cf. Roginsky (1992)).

The law of iterated logarithm (c) appears in Feller (1949) (Theorem 12, p. 114). Generalizations to the non-identically distributed case were proposed by Hatori (1959) and Niculescu (1988). Similar results for increments of more general renewal processes can be found in Deheuvels and Steinebach (1989).

3. Stochastic comparison of Markov chains

Let us first recall the definition of stochastic ordering for random variables.

Definition 3.1. Let \( Y \) and \( Y^* \) be two integer-valued random variables. We say that \( Y \) is stochastically inferior to \( Y^* \) iff the distribution of \( Y \) is stochastically inferior to that of \( Y^* \), i.e.

\[ P[Y \leq k] \leq P[Y^* \leq k] \quad \forall k \in \mathbb{N}. \]

This property will still be denoted by \( Y \ll Y^* \).

The following two characterizations of stochastic ordering are well known. The first one expresses it in terms of expectations of non-decreasing functions (it is often used as a definition). The second one says that two variables \( Y \) and \( Y^* \) such that \( Y \ll Y^* \) can be coupled so that \( Y \) remains almost surely below \( Y^* \).

Theorem 3.2. Let \( Y \) and \( Y^* \) be two integer-valued random variables. Then \( Y \ll Y^* \) iff any of the following properties holds:

(a) For any non-decreasing function \( f \) from \( \mathbb{N} \) into \( \mathbb{R} \):

\[ E[f(Y)] \leq E[f(Y^*)]. \]
(b) There exists a couple \((\tilde{Y}, \tilde{Y}^*)\) such that \(\tilde{Y}\) and \(\tilde{Y}^*\) have the same distributions as \(Y\) and \(Y^*\) respectively and

\[
P[\tilde{Y} \leq \tilde{Y}^*] = 1.
\]

Property (a) can be checked directly from the definition, expressing any non-decreasing function \(f\) as a linear combination with non-negative coefficients of indicator functions of the sets \(\{k, k+1, \ldots\}\). Property (b) is usually deduced from a general coupling result by Strassen (1965), Theorem 11 on p. 436 (cf. for instance Kamae et al. (1977), Theorem 1 on p. 900).

The stochastic comparison of Markov chains has been treated by (among others) Kamae et al. (1977), Keilson and Kester (1977) or more recently Kijima (1989). The idea is the following. If the transition kernels of two Markov chains can be compared stochastically, then the stochastic order carries through in time and the two chains can be coupled so as to force one of them to remain almost surely below the other. We use this result in the following form.

**Theorem 3.3.** Let \((Y_i)_{i \geq 0}\) and \((Y_i^*)_{i \geq 0}\) be two homogeneous Markov chains with values in \(\mathbb{N}\). Denote by \((q, .)\) and \((q^*, .)\) their transition kernels:

\[
q(m, n) = P[Y_{i+1} = n | Y_i = m],
\]

\[
q^*(m, n) = P[Y_i^* = n | Y_i^* = m] \quad \forall m, n \in \mathbb{N}.
\]

Assume \(Y_0 \ll Y_0^*\) and

\[
q(m, .) \ll q^*(m^*, .) \quad \forall m \leq m^* \in \mathbb{N}.
\]

Then there exists a couple \((\tilde{Y}_i, \tilde{Y}_i^*)_{i \geq 0}\) such that \((\tilde{Y}_i)_{i \geq 0}\) and \((\tilde{Y}_i^*)_{i \geq 0}\) have the same distributions as \((Y_i)_{i \geq 0}\) and \((Y_i^*)_{i \geq 0}\) respectively and

\[
\forall i \geq 0 \quad P[\tilde{Y}_i \leq \tilde{Y}_i^*] = 1.
\]

This is a particular case of Theorem 2 on p. 903 of Kamae et al. (1977) (cf. also Stoyan (1983), p. 65). It has been generalized to other types of stochastic orderings by Doisy (1992). Our idea is to compare stochastically absorption times and renewal counting processes. Let \(H_n\) and \(H_n^*\) be defined as in the introduction, the first one as the absorption time at 1 of the Markov chain with matrix \(P\), the second one as the first crossing time of 1 for a backward random walk driven by \(p\).

**Corollary 3.4.** Assume that

\[
p_n \ll p \quad \text{(respectively } p \ll p_n) \quad \forall n \geq 1.
\]

Then

\[
H_n^* \ll H_n \quad \text{(respectively } H_n \ll H_n^*) \quad \forall n \geq 1.
\]

Indeed, Theorem 3.3 can be applied to the Markov chains \((X_i)\) and \((X_i^*) = (n - T_i - \cdots - T_i)\).
In order to construct a coupled version, \((\tilde{X}_i, \tilde{X}_i^*)_{i \geq 1}\) such that \(\tilde{X}_i \equiv \tilde{X}_i^*\) a.s. for all \(i\). Denote
\[
\bar{H}_n = \inf \{i \text{ such that } \tilde{X}_i = 1\} \quad \text{and} \quad \tilde{H}_n^* = \inf \{i \text{ such that } \tilde{X}_i^* \leq 1\}.
\]
One has \(\bar{H}_n \leq \tilde{H}_n^*\) a.s. Now since \((\tilde{X}_i)_{i \geq 1}\) and \((\tilde{X}_i^*)_{i \geq 1}\) have the same distributions as \((X_i)_{i \geq 1}\) and \((X_i^*)_{i \geq 1}\) respectively, \(\bar{H}_n\) and \(\tilde{H}_n^*\) have the same distribution as \(H_n\) and \(H_n^*\). Hence the result, by Theorem 3.2.

The intuitive meaning is quite clear: if the steps of the random walk are longer on average, it will take less time to reach any level. This result is not new. Generalizations and refinements can be found in Kijima (1989).

4. Asymptotic behavior of the absorption time \(H_n\)

The notation and hypotheses are those of Section 1. Recall that \((p_n)\) is a sequence of probability distributions on \(\mathbb{N}\),
\[
p_1(\{0\}) = 1 \quad \text{and} \quad p_n(\{1, \ldots, n-1\}) = 1 \quad \forall n \geq 1.
\]
\((X_i)_{i \geq 1}\) is a Markov chain with transition matrix \(P = (\pi_{ij})_{i,j \geq 1}\), with
\[
\pi_{ij} = p_i(i-j) \quad \text{if } i \geq 1, \ 1 \leq j \leq i
\]
\[
= 0 \quad \text{in any other case}.
\]
\(H_n\) is the time to absorption at 1 of \((X_i)_{i \geq 1}\) starting from state \(n\):
\[
H_n = \inf \{i \text{ such that } X_i = 1, \ \text{when} \ X_0 = n\}.
\]

Our main result is the following.

**Theorem 4.1.** Under Hypothesis 1.3, the expectation \(\mu\) and standard deviation \(\sigma\) of the limit distribution \(p\) are finite and:

(a) \(\lim_{n \to \infty} E[(H_n/n)^r] = (1/\mu)^r \quad \forall r \leq 2\)

(b) \(\lim_{n \to \infty} \mathcal{L}[(H_n - n/\mu)/(n\sigma^2/\mu^3)^{1/2}] = N(0, 1)\).

For \(r = 1\), (a) has been proved in a different context by Kijima and Sumita (1986). The proof of Theorem 4.1 relies essentially on stochastic bounds for the sequence \((p_n)\). We begin with some notation. For all \(n \geq 1\) and \(k \geq 1\) let
\[
r_n(k) = \sum_{i=k}^{\infty} p_n(i), \quad r(k) = \sum_{i=k}^{\infty} p(i), \quad s(k) = \sum_{i=k}^{\infty} q(i).
\]
\[
r_m^+(k) = \sup \{r_n(k), n \geq m\}, \quad r_m^-(k) = \inf \{r_n(k), n \geq m\},
\]
\[
p_m^+(k) = r_m^+(k) - r_m^+(k+1), \quad p_m^-(k) = r_m^-(k) - r_m^-(k+1).
\]
That $p^+_m$ and $p^-_m$ are well-defined probability distributions is obvious. They should be seen respectively as stochastic upper and lower bounds for the remainder after $m$ of the sequence $(p_n)$. One has for all $m$ and all $n \geq m$:

$$p^-_m \ll p_n \ll p^+_m,$$

but also

$$p^-_m \ll p \ll p^+_m,$$

and

$$p^-_m \ll p^+_m \ll q.$$

Moreover, it can easily be checked that for all $n$ and $k$

$$\lim_{m \to \infty} p^-_m(k) = \lim_{m \to \infty} p^+_m(k) = p(k),$$

so that both sequences $(p^-_m)$ and $(p^+_m)$ satisfy Hypothesis 1.3. Actually (1.3) implies more than a simple convergence in distribution.

**Lemma 4.2.** Let $(p_n)$ be a sequence of probability distributions satisfying Hypothesis 1.3. Then

(a) $\sum_{k=0}^{\infty} k^2p_n(k)$ and $\sum_{k=0}^{\infty} k^2p(k)$ are finite.

(b) $\lim_{n \to \infty} \sum_{k=0}^{\infty} k(p_n(k) - p(k)) = \lim_{n \to \infty} \sum_{k=0}^{\infty} k^2(p_n(k) - p(k)) = 0.$

**Proof of Lemma 4.2.** Since $k^2$ is an increasing function of $k$, (a) comes from Hypothesis 1.3(a) and (b). For (b), let us write

$$\sum_{k=0}^{\infty} k^2(p_n(k) - p(k)) = \sum_{k=0}^{\infty} (2k - 1)(r_n(k) - r(k)).$$

Now for all $k \geq 1$

$$\lim_{n \to \infty} r_n(k) = r(k) \quad \text{and} \quad |r_n(k) - r(k)| \leq 2s(k).$$

Hence the result by Lebesgue's dominated convergence theorem.

It follows in particular from Lemma 4.2 that the expectations and variances of $p$, $p^+_m$ and $p^-_m$ are finite. We shall denote by $\mu^+_m$ and $\mu^-_m$ the expectations of $p^+_m$ and $p^-_m$ and by $\sigma^+_m$ and $\sigma^-_m$ their standard deviations. One has

$$\lim_{m \to \infty} \mu^-_m = \lim_{m \to \infty} \mu^+_m = \mu, \quad \lim_{m \to \infty} \sigma^-_m = \lim_{m \to \infty} \sigma^+_m = \sigma.$$

Our idea is to compare the absorption time $H_n$ to crossing times of backward random walks starting from $n$ and driven by the distributions $p^+_m$ and $p^-_m$ respectively.
Let \((X_i)_{i \geq 0}\) be a Markov chain with transition matrix \(P\), such that \(X_0 = n\). For \(m < n\), denote by \(H_n(m)\) its crossing time of the level \(m\):

\[H_n(m) = \inf \{i \text{ such that } X_i \leq m\}\]

(with this notation, \(H_n = H_n(1)\)). Let \((T^+_i)_{i \geq 1}\) and \((T^-_i)_{i \geq 1}\) be two sequences of i.i.d. random variables with distributions \(p^+_m\) and \(p^-_m\) respectively. Let

\[X^+_i = n - T^+_i - \cdots - T^+_i \quad \text{and} \quad X^-_i = n - T^-_i - \cdots - T^-_i.\]

Finally, let

\[H^+_n(m) = \inf \{i \text{ such that } X^+_i \leq m\} \quad \text{and} \quad H^-_n(m) = \inf \{i \text{ such that } X^-_i \leq m\}.\]

By the reasoning of Corollary 3.4, one has for all \(n > m\)

\[H^+_n(m) \ll H_n(m) \ll H^-_n(m),\]

but obviously

\[H_n(m) \ll H_n \ll H_n(m) + m,\]

and therefore

\[H^+_n(m) \ll H_n \ll H^-_n(m) + m.\]

Now for any non-decreasing function \(f\) from \(\mathbb{N}\) into \(\mathbb{R}\) one has

\[E[f(H^+_n(m))] \leq E[f(H_n)] \leq E[f(H^-_n(m) + m)].\]

In particular, for \(r \equiv 2\) and \(f(k) = (k/n)^r\),

\[E[(H^+_n(m)/n)^r] \leq E[(H_n/n)^r] \leq E[((H^-_n(m) + m)/n)^r].\]

Fixing \(m\) and letting \(n\) tend to infinity, one obtains, thanks to (a) of Theorem 2.1,

\[\forall m, \forall \varepsilon > 0, \exists n_0 \text{ such that } \forall n > n_0, \quad (1/\mu^+_m)^r - \varepsilon \leq E[(H_n/n)^r] \leq (1/\mu^-_m)^r + \varepsilon.\]

Since \(\lim_{m \to \infty} \mu^-_m = \lim_{m \to \infty} \mu^+_m = \mu\), Theorem 4.1 (a) follows.

For the central limit theorem (b), observe that the indicator function of \((k - n/\mu)/\sqrt{n} < t\) is a decreasing function of \(k\). Hence

\[P[(H^+_n(m) - n/\mu^+_m)/\sqrt{n} < t] \geq P[(H_n - n/\mu^+_m)/\sqrt{n} < t],\]

and

\[P[(H^-_n(m) - m - n/\mu^-_m)/\sqrt{n} < t] \geq P[(H^-_n(m) + m - n/\mu^-_m)/\sqrt{n} < t].\]

Now the left-hand side (respectively right-hand side) of the first (second) inequality converges to the distribution function of a normal law with mean 0 and variance \((\sigma^+_m)^2/(\mu^+_m)^3\) (respectively \((\sigma^-_m)^2/(\mu^-_m)^3\)). Part (b) of Theorem 4.1 is then obtained using

\[\lim_{m \to \infty} \mu^-_m = \lim_{m \to \infty} \mu^+_m = \mu \quad \text{and} \quad \lim_{m \to \infty} \sigma^-_m = \lim_{m \to \infty} \sigma^+_m = \sigma.\]
Notice that, unlike the case of the classical renewal process, there is no canonical way to define the sequence \( (H_n) \) on the same probability space. Hence almost sure results, and in particular the law of the iterated logarithm, have no practical meaning for \( (H_n) \).

The reader may wonder why no asymptotic result is given concerning the variance of \( H_n \). In view of Theorem 4.1 it is quite natural to conjecture that

\[
\lim_{n \to \infty} \frac{\text{Var} [H_n]}{n} = \frac{\sigma^2}{\mu^3}.
\]

It seems to us that there is little chance of reaching this result by our comparison technique without further assumptions, for the following reasons. Firstly, \( (k - n/\mu)^2 \) is not a monotone function of \( k \) and therefore the technique of proof that was used above does not work directly. However, one could hope to compute \( \text{Var} [H_n] \) by giving more precise asymptotic developments of \( E[H_n] \) and \( E[\mathcal{H}_n] \), as does Feller (1949) (formulae (6.7) and (6.9) on p. 111). The problem is that in our method one controls the step of the Markov chain \( (X_i) \) only up to some fixed crossing level \( m \), so that no asymptotic development obtained in this way can go up to constant terms. In particular, it is easy to give a development of \( E[\mathcal{H}_n^2] \) up to order 1 (Proposition 4.3 below), but the development of \( E[H_n] \) up to order zero remains out of reach. Notice also that this defect is not only due to our method. Indeed, assume that one adds to the hypotheses

\[
p_n(n - 1) = 1 \quad \forall n \leq m.
\]

(Immediately after crossing level \( m \), the Markov chain gets absorbed.) This does not change the conclusions of Theorem 4.1, but it does change \( E[H_n] \) by a constant factor.

**Proposition 4.3.** \( \lim_{n \to \infty} \frac{(E[\mathcal{H}_n^2] - n^2/\mu^2)}{n} = 2\sigma^2/\mu^3 + 1/\mu^2 - 1/\mu. \)

The proof is the same as that of (a) in Theorem 4.1, using formula (6.9) on p. 111 of Feller (1949).

### 5. Applications to combinatorics

For the notions of combinatorics discussed in this section, our general reference is Comtet (1974). Asymptotic results, and in particular central limit theorems, have attracted a lot of attention in that field (cf. Flajolet and Soria (1993) for a recent review). The results are very often obtained using recurrence formulae and the explicit expressions for generating functions they entail (as for instance in Flajolet and Odlyzko (1984)). Our approach, illustrated in this section by four examples, is also based on recurrence formulae, but they are used to construct transition matrices.
**Example 1.** We come back to the example given in the introduction. Recall that a random chain of partitions is defined as a random variable uniformly distributed over the set of ordered chains of partitions. For such a random chain, $X_i$ is the cardinality of the partition at level $i$ ($X_0 = n$). The sequence of random variables $(X_i)_{i \geq 0}$ is a Markov chain with transition matrix $P = (\pi_{ij})$ with

$$\pi_{ij} = S(i, j)z(j)/z(i) \quad \forall i = 1, \ldots, n, \forall j = 1, \ldots, n - 1.$$ 

This transition matrix is of the type discussed so far with

$$p_n(k) = S(n, n - k)z(n - k)/z(n) \quad \forall n, \forall k = 1, \ldots, n - 1.$$ 

The absorption time $H_n$ of this Markov chain is exactly the length of the random chain.

In order to apply Theorem 4.1, one has to check that Hypothesis 1.3 is fulfilled.

**Proposition 5.1.** The sequence $(p_n)$ satisfies Hypothesis 1.3 with

$$p = (\log (2^k/k!))_{k \geq 1}.$$ 

**Proof.** It follows from the asymptotic expansions given in Lengyel (1984) for $z(n)$ (Theorem 1.2, p. 313) and for $S(n, n - k)$ (Lemma 3.1, p. 316) that for any fixed $k$:

$$\lim_{n \to \infty} p_n(k) = \log (2^k/k!).$$

But also (using Theorem 1.1 on p. 313 and Lemma 3.2 on p. 316 of the same reference)

$$\exists c > 0; \forall n \geq 1 \quad \text{and} \quad \forall k \geq 1, \quad p_n(k) < c \log (2^k).$$

Let us fix $N$ large enough to ensure that

$$r = \sum_{k=1}^{\infty} c \log (2^k) < 1.$$ 

Define the distribution $q$ by

$$q(N - 1) = 1 - r \quad \text{and} \quad q(k) = c \log (2^k) \quad \forall k \geq N.$$ 

One has obviously

$$\forall n \geq 1, \quad p_n \ll q \quad \text{and} \quad \sum_{k=1}^{\infty} k^2 q(k) < \infty.$$ 

The expectation and variance of $p$ are

$$\mu = 2 \log (2) \quad \text{and} \quad \sigma^2 = 2 \log (2)(1 - \log (2)).$$

The results of Theorem 4.1 enable us to give a precise description of the ‘typical’ random chain of partitions on a large $n$-set. Its length is of order $n/\mu \approx 0.721n$. The cardinalities of the partitions in the chain decrease slowly: between one level and its
successor the difference is 1 in 69% of the cases, 2 in 24% of the cases, 3 or more in only 7% of the cases (on average the difference is $2 \log (2) \approx 1.386$). One could expect the first element of the random chain of partitions to be a random partition (uniformly distributed in the set of all partitions of an $n$-set). This is not the case, since a ‘typical’ random partition has $n / \log (n)$ elements (cf. Comtet (1974), p. 296), and not $n - 2 \log (2)$.

Recurrences of the type satisfied by $z(n)$ occur quite often in combinatorics. Here is another example.

**Example 2.** Consider the set of all ordered partitions $(A_1, \cdots, A_h)$ of an $n$-set (distinguishing different orderings of the same partition). Let $y(n)$ be the cardinality of that set. One has obviously:

$$y(n) = \sum_{k=1}^{n} S(n, k)k!.$$ 

(Recall that the Stirling number of the second kind $S(n, k)$ is the number of (unordered) partitions of $n$ elements into $k$ subsets.) But the $y(n)$’s can also be defined by $y(0) = 1$ and

$$y(n) = \sum_{k=0}^{n-1} \binom{n}{k}y(k) \quad \forall n \geq 1.$$ 

Indeed to choose an ordered partition of an $n$-set, one can begin with choosing $n - k$ elements to form the first subset and then go on with an ordered partition of the $k$ elements left. (This recurrence differs slightly from that satisfied by the Bell numbers (cf. Comtet (1974), p. 210)).

Let $(A_1, \cdots, A_h)$ be a random ordered partition (uniformly distributed on the set of all ordered partitions on an $n$-set). Let

$$X_0 = n, \cdots, X_i = \text{card}(A_{i+1} \cup \cdots \cup A_h), \cdots, X_h = 0, \cdots X_{h+j} = 0, \cdots.$$ 

The sequence $(X_i)_{i \geq 0}$ of random variables is a Markov chain on $\mathbb{N}$, with transition matrix $P = (\pi_{nk})_{n,k \geq 0}$ defined by:

$$\pi_{nk} = p_n(n - k), \quad \text{if } n \geq 0, 1 \leq k \leq n$$

$$= 0 \quad \text{in any other case.}$$

$$\forall n, \forall k = 1, \cdots, n, \quad p_n(k) = \binom{n}{n-k}y(n-k)/y(n).$$

Let $H_n$ be the absorption time at 0 of the Markov chain $(X_i)_{i \geq 0}$:

$$H_n = \inf \{i \text{ such that } X_i = 0\}.$$ 

$H_n$ represents the number of coordinates (subsets) in a random ordered partition on an $n$-set. Up to a minor change of index (absorption at 0 instead of 1), the problem of determining the asymptotic distribution of $H_n$ is of the type studied so far. The
limit of the $p_n$'s is the same distribution $p$ as before (which is more than a coincidence in view of the results of Lengyel (1974)).

**Proposition 5.2.** The sequence $(p_n)$ satisfies Hypothesis 1.3 with

$$p = (\log (2)^k / k!)_{k \geq 1}.$$

**Proof.** Let us first remark that $u(n) = y(n)/n!$ is solution of

$$u(n) = \sum_{k=0}^{n-1} (1/k!)u(n-k) \quad \forall n \geq 2.$$

Define $v(n) = (1/2) \log (2)^{-n}$. One has

$$v(n)(1 - \varepsilon_n) = \sum_{k=1}^{n-1} (1/k!)v(n-k) \quad \forall n \geq 2,$$

with

$$\varepsilon_n = \sum_{k=n}^{\infty} (1/2) \log (2)^k / k!.$$

Thus $u(n)$ and $v(n)$ can be compared proceeding as in Babai and Lengyel (1992) and Lengyel (1974). One gets:

$$u_n = c_1 \log (2)^{-n}(1 + o(n)).$$

(Numerical evidence suggests that $c_1 = 1/(2 \log (2))$.) From this it follows that

$$\lim_{n \to \infty} p_n(k) = \lim_{n \to \infty} (1/k!)u(n-k)/u(n) = \log (2)^k / k!,$$

but also

$$\exists c_2 > 0, \forall n \geq 1 \text{ and } \forall k \geq 1, \quad p_n(k) < c_2 \log (2)^k / k!.$$ 

Now fix $N$ large enough to ensure

$$r = \sum_{k=N}^{\infty} c_2 \log (2)^k / k! < 1.$$

Defining the distribution $q$ by

$$q(k) = c_2 \log (2)^k / k! \quad q(N-1) = 1 - r \quad \text{and} \quad \forall k > N,$$

one has obviously

$$p_n \ll q \quad \forall n \geq 1 \text{ and } \sum_{k=1}^{\infty} k^2 q(k) < \infty.$$

Thus the results of Theorem 4.1 are true for $H_n$. In particular a random ordered partition on an $n$-set is formed of $n/2 \log 2$ subsets on average, whereas a random ordinary partition has $n/\log (n)$ subsets. This is not surprising, since distinguishing the different orderings gives a much stronger weight to partitions with many subsets.
Example 3. Consider now the Fibonacci sequence $F_n$ (cf. Comtet (1974), p. 45) defined by

$$F_0 = F_1 = 1, \quad F_n = F_{n-1} + F_{n-2}, \quad n \geq 2.$$  

Among other interpretations, $F_n$ can be seen as the number of permutations of the set $\{1, \cdots, n\}$ such that the image of each element is itself or one of its neighbors.

$$F_n = \text{Card} \{\sigma \in S_n \text{ such that } \sigma(1) \in \{1, 2\}; \cdots; \sigma(i) \in \{i - 1, i, i + 1\}; \cdots; \sigma(n) \in \{n - 1, n\}\}.$$  

Indeed, one constructs such a permutation by iterating the following algorithm. Choose first the image of 1. If it is 1, one has to construct a permutation with the same property on $n - 1$ elements. If the image of 1 is 2, then the image of 2 must be 1 and one is left with $n - 2$ elements. Hence the recurrence. Consider now a random variable uniformly distributed on the set defined above. Let $X_i$ be the random variable defined as the number of elements left after the $i$th step of the algorithm ($X_0 = n$). $(X_i)_{i \geq 0}$ is a Markov chain of the type studied here. One has:

$$p_n(1) = \frac{F_{n-1}}{F_n} \quad \text{and} \quad p_n(2) = \frac{F_{n-2}}{F_n} = 1 - p_n(1).$$  

These distributions are stochastically dominated by the Dirac mass at 2 and they converge to $p$ as $n$ tends to infinity, with

$$p(1) = \frac{(\sqrt{5} - 1)}{2} \quad \text{and} \quad p(2) = \frac{(3 - \sqrt{5})}{2}.$$  

The expectation and variance of $p$ are

$$\mu = \frac{(5 - \sqrt{5})}{2} \quad \text{and} \quad \sigma^2 = \sqrt{5} - 2.$$

Let $N$ denote the random variable equal to the number of invariant elements for the random permutation. This number is related to the time to absorption of the Markov chain by

$$H_n = N + (n - N)/2 \quad \text{or} \quad N = 2H_n - n.$$  

Applying Theorem 4.1 one obtains in particular that the asymptotic number of invariant elements in a random permutation as defined above is $n/\sqrt{5}$.

Actually the method proposed here can be of interest in any combinatorial situation where some given structure can be generated recursively. The recursivity leads to a recurrence formula that can be interpreted in terms of a Markov chain. We want to emphasize that the approximation of the Markov chain by a backward random walk has an application to the random generation of such structures. The idea will be better illustrated by our first example. If one wants to generate a random chain of partitions on an $n$-set, the Markovian interpretation suggests the following algorithm. Start by choosing the number of elements that will be grouped together in the next partition, according to the distribution $p_n$. When this integer $k$
is chosen, construct a uniformly distributed random partition having \( n - k \) subsets. Then replace \( n \) by \( n - k \) and iterate. The problem with this algorithm is that the distribution \( p_n \) is hard to compute for large \( n \)'s. An alternative way is to replace \( p_n \) by the limit distribution \( p \), as long as some fixed level \( m \) is not reached. Of course this is only an approximate method, but it is far less expensive than the exact one and its precision can be controlled. A different use of Markov chains for generating uniform random combinatorial structures has been proposed by Aldous (1987). Notice that the absorption time \( H_n \) is approximately the number of iterations in the algorithm above. Knowing its distribution gives some control over the cost of the algorithm.

The use of renewal processes in probabilistic analysis of algorithms is not new (cf. for instance Hofri (1987), p. 176 or Csirik et al. (1991)). We believe that the study presented here can be of interest to the study of randomized algorithms, even outside the field of random structures in combinatorics.

6. Extensions

In this section we comment firstly on the possibility of obtaining more asymptotic results for the absorption time \( H_n \), secondly on extensions to Markov chains with similar transition matrices.

As we pointed out at the end of Section 4, very precise asymptotic expansions are out of reach. However, with stronger assumptions part (a) of Theorem 4.1 can be extended.

**Proposition 6.1.** Let \( \alpha \) be a positive real. Assume Hypothesis 1.3 with

\[
(a') \quad \sum_{k=1}^{\infty} k^{\alpha} q(k) < \infty.
\]

Then

\[
\lim_{n \to \infty} E[(H_n/n)^r] = (1/\mu)^r \quad \forall r \leq \alpha.
\]

The proof is the same as that of Theorem 4.1(a).

Notice that the uniform stochastic boundedness in Hypothesis 1.3 was used more to ensure good properties of the upper bounds \( p_m^+ \) than for the lower bound \( p_m^- \). Under milder hypotheses of convergence, the reasoning of Section 4 would still hold as far as \( p_m^- \) (and upper bounds for the behavior of \( H_n \)) would be concerned. This would be in our opinion of lesser interest. Nonetheless, Hypothesis 1.3 cannot be considered as necessary, even though we believe it is not unduly strong. Let us illustrate its degree of sharpness on an example.

For each \( n \geq 3 \), take

\[
p_n(1) = 1 - f(n), p_n(2) = \cdots = p_n(n - 2) = 0, p_n(n - 1) = f(n),
\]
where \(f(n)\) is some decreasing function of \(n\) with values in \([0,1]\), tending to 0 at infinity (so that \(p_n\) converges to the Dirac mass at 1). The upper bound \(p_m^+\) is such that
\[
p_m^+(1) = 1 - f(m), \quad p_m^+(k) = 0 \text{ for } k < m \quad \text{and} \quad p_m^+(k) = f(k) - f(k + 1) \text{ for } k \geq m.
\]
It is easy to obtain the explicit distribution of \(H_n\):
\[
P[H_n = 1] = f(n) \\
P[H_n = i] = (1 - f(n))(1 - f(n-1)) \cdots (1 - f(n - i + 2))f(n - i + 1)
\]
for \(i = 2, \cdots, n - 2\)
\[
P[H_n = n - 1] = (1 - f(n))(1 - f(n - 1)) \cdots (1 - f(4))(1 - f(3)).
\]
We can check also that for \(n \geq 3\):
\[
E[H_n] = 1 + (1 - f(n))E[H_{n-1}],
\]
and
\[
E[H_n^2] = 2E[H_n] - 1 + (1 - f(n))E[H_{n-1}^2].
\]
For \(f(n) = 1/n\): \(\sum_{k=1}^n |p_n(k) - p(k)|\) tends to zero, but \(\sum_{k=1}^n k |p_n(k) - p(k)|\) does not. The upper bounds \(p_m^+\) have no expectation. One can check that \(E[H_n/n]\) tends to \(\frac{1}{2}\) and not 1 as would be the case if Theorem 4.1 was true (\(E[H_n^2/n^2]\) tends to \(\frac{1}{2}\) instead of 1).

For \(f(n) = 1/n^2\): \(\sum_{k=1}^n k |p_n(k) - p(k)|\) tends to zero, but \(\sum_{k=1}^n k^2 |p_n(k) - p(k)|\) does not. The upper bounds \(p_m^+\) have an expectation, but no variance. One can check that \(E[H_n/n]\) and \(E[H_n^2/n^2]\) tend to 1 as expected. However, their asymptotic expansions are
\[
E[H_n] = n - \log(n) + O(1),
\]
and
\[
E[H_n^2] = n^2 - 2n \log(n) + O(n),
\]
so that the conclusion of Proposition 4.3 does not hold.

For \(f(n) = 1/n^3\): The upper bounds \(p_m^+\) have a variance, Theorem 4.1 and Proposition 4.3 hold.

Finally, let us comment on the triangular shape of the transition matrix \(P\). It was dictated by the applications to combinatorics described in Section 5. It is by no means an obligation. Indeed, classical renewal theory admits also distributions with a positive mass on the negative axis, provided the mean remains positive (cf. for instance Gut (1974)), so that Theorem 4.1 remains true when allowing the distributions \(p_n\) to charge non-positive integers, provided that the limit distribution \(p\) has a positive expectation, and the Markov chain does absorb almost surely. However, we do not see any application for this situation. Also, one can reverse the
space of states and deal with upper-triangular matrices instead. This gives a more natural situation from the point of view of renewal theory. It has been considered before by Kijima and Sumita (1986) for real-valued random variables: Theorems 6.2 and 6.3 below can be considered as extensions of Theorems 4.1 and 4.2 on pp. 88–89 of that reference.

**Theorem 6.2.** Let \((X_i)_{i\geq 1}\) be a Markov chain on \(\mathbb{N}\) with transition matrix \(P = (\pi_{ij})\). Denote by \(p_n\) its \(n\)th line re-indexed as:

\[
p_n(k) = \pi_n(n + k).
\]

Assume the \(p_n\)'s satisfy Hypothesis 1.3 for some distribution \(p\) (giving possibly a positive mass to negative integers, but having a positive expectation) and \(q\). Then \((X_i)_{i\geq 1}\) tends to \(+\infty\) a.s. Let \(H_n\) denote the crossing time of level \(n\):

\[
H_n = \inf\{i \text{ such that } X_i \geq n\}.
\]

Assume \(X_0\) is bounded almost surely.

Then the conclusions of Theorem 4.1 hold for \(H_n\).

The proof is almost exactly the same as that of Theorem 4.1, even simpler since the upper and lower approximations are ordinary random walks and not backward random walks. Also in this case the first-passage times \(H_n\) are all defined by the same sequence \((X_i)_{i\geq 1}\) and almost sure results make sense. The almost sure coupling of Theorem 3.3 used along the same lines permits us moreover to deduce the law of the iterated logarithm (c) of Theorem 2.1.

**Theorem 6.3.** Under the hypotheses of Theorem 6.2, let \(\mu\) and \(\sigma\) be the mean and standard deviation of the limit distribution \(p\). One has

\[
\limsup_{n \to \infty} |H_n - n/\mu|/((2\sigma^2/\mu^2)n \log \log n)^{1/2} = 1 \text{ a.s.}
\]

**Acknowledgements**

We are pleased to thank the referee for pointing out some important references, and Ph. Flajolet for his interesting comments.

**References**


Renewal-Type Behavior of Absorption Times in Markov Chains
Bernard Van Cutsem; Bernard Ycart
Stable URL:
http://links.jstor.org/sici?sici=0001-8678%28199412%2926%3A4%3C988%3ARBOATI%3E2.0.CO%3B2-R

This article references the following linked citations. If you are trying to access articles from an off-campus location, you may be required to first logon via your library web site to access JSTOR. Please visit your library's website or contact a librarian to learn about options for remote access to JSTOR.

References

A Renewal Theorem for Random Variables which are Dependent or Non-Identically Distributed
Y. S. Chow; Herbert Robbins
Stable URL:
http://links.jstor.org/sici?sici=0003-4851%28196306%2934%3A2%3C390%3AARTFRV%3E2.0.CO%3B2-0

Sharp Rates for Increments of Renewal Processes
Paul Deheuvels; Josef Steinebach
Stable URL:
http://links.jstor.org/sici?sici=0091-1798%28198904%2917%3A2%3C700%3ASRFIOR%3E2.0.CO%3B2-2

On the Integral Equation of Renewal Theory
Willy Feller
Stable URL:
http://links.jstor.org/sici?sici=0003-4851%28194109%2912%3A3%3C243%3AOTIEOR%3E2.0.CO%3B2-S

On the Moments and Limit Distributions of Some First Passage Times
Allan Gut
Stable URL:
http://links.jstor.org/sici?sici=0091-1798%28197404%292%3A2%3C277%3AOTMALD%3E2.0.CO%3B2-K
Stochastic Inequalities on Partially Ordered Spaces
T. Kamae; U. Krengel; G. L. O'Brien
Stable URL:
http://links.jstor.org/sici?sici=0091-1798%28197712%295%3A6%3C899%3ASIOPPOS3E2.0.CO%3B2-3

A Useful Generalization of Renewal Theory: Counting Processes Governed by Non-Negative Markovian Increments
Masaaki Kijima; Ushio Sumita
Stable URL:
http://links.jstor.org/sici?sici=0021-9002%28198603%2923%3A1%3C71%3AAUGORT3E2.0.CO%3B2-T

A Central Limit Theorem with Remainder Term for Renewal Processes
Allen L. Roginsky
Stable URL:
http://links.jstor.org/sici?sici=0001-8678%28199206%2924%3A2%3C267%3AAALTWIR3E2.0.CO%3B2-7

The Existence of Probability Measures with Given Marginals
V. Strassen
Stable URL:
http://links.jstor.org/sici?sici=0003-4851%28196504%2936%3A2%3C423%3ATEOPMW3E2.0.CO%3B2-C