J. Appl. Prob. **48,** 527–546 (2011)
Printed in England

A CENTRAL LIMIT THEOREM AND ITS APPLICATIONS TO MULTICOLOR RANDOMLY REINFORCED URNS

PATRIZIA BERTI,* *Università di Modena e Reggio-Emilia*
IRENE CRIMALDI,** *Università di Bologna*
LUCA PRATELLI,*** *Accademia Navale*
PIETRO RIGO,**** *Università di Pavia*

Abstract

Let (X_n) be a sequence of integrable real random variables, adapted to a filtration $(\mathcal{G}_n)$. Define $C_n = \sqrt{n}\{(1/n)\sum_{k=1}^n X_k - \mathrm{E}(X_{n+1} \mid \mathcal{G}_n)\}$ and $D_n = \sqrt{n}\{\mathrm{E}(X_{n+1} \mid \mathcal{G}_n) - Z\}$, where Z is the almost-sure limit of $\mathrm{E}(X_{n+1} \mid \mathcal{G}_n)$ (assumed to exist). Conditions for $(C_n, D_n) \to \mathcal{N}(0, U) \times \mathcal{N}(0, V)$ stably are given, where U and V are certain random variables. In particular, under such conditions, we obtain $\sqrt{n}\{(1/n)\sum_{k=1}^n X_k - Z\} = C_n + D_n \to \mathcal{N}(0, U + V)$ stably. This central limit theorem has natural applications to Bayesian statistics and urn problems. The latter are investigated, by paying special attention to multicolor randomly reinforced urns.

Keywords: Bayesian statistics; central limit theorem; empirical distribution; Poisson–Dirichlet process; predictive distribution; random probability measure; stable convergence; urn model

2010 Mathematics Subject Classification: Primary 60F05; 60G57; 60B10; 62F15

1. Introduction and motivations

As regards asymptotics in urn models, there is not a unique reference framework. Rather, there are many (ingenious) disjoint ideas, one for each class of problems. Well-known examples are martingale methods, exchangeability, branching processes, stochastic approximation, dynamical systems, and so on; see [16].

Those limit theorems which unify various urn problems, thus, look of some interest.

In this paper, we focus on the central limit theorem (CLT). While thought for urn problems, our CLT is stated for an arbitrary sequence of real random variables. Thus, it potentially applies to every urn situation, even if its main application (known to us) is an important special case of *randomly reinforced urns* (RRUs).

Let (X_n) be a sequence of real random variables such that $\mathrm{E}\,|X_n| < \infty$. Define $Z_n = \mathrm{E}(X_{n+1} \mid \mathcal{G}_n)$, where $\mathcal{G} = (\mathcal{G}_n)$ is some filtration which makes (X_n) adapted. Under various assumptions, one obtains $Z_n \to Z$ almost surely (a.s.) and in L_1 for some random variable Z.

Received 4 May 2010; revision received 31 January 2011.
* Postal address: Dipartimento di Matematica Pura ed Applicata 'G. Vitali', Università di Modena e Reggio-Emilia, via Campi 213/B, 41100 Modena, Italy. Email address: patrizia.berti@unimore.it
** Postal address: Dipartimento di Matematica, Università di Bologna, Piazza di Porta San Donato 5, 40126 Bologna, Italy. Email address: crimaldi@dm.unibo.it
*** Postal address: Accademia Navale, viale Italia 72, 57100 Livorno, Italy. Email address: pratel@mail.dm.unipi.it
**** Postal address: Dipartimento di Economia Politica e Metodi Quantitativi, Università di Pavia, via S. Felice 5, 27100 Pavia, Italy. Email address: prigo@eco.unipv.it

https://doi.org/10.1239/jap/1308662642 Published online by Cambridge University Press

Define further $\bar{X}_n = (1/n)\sum_{k=1}^n X_k$ and

$$C_n = \sqrt{n}(\bar{X}_n - Z_n), \qquad D_n = \sqrt{n}(Z_n - Z), \qquad W_n = C_n + D_n = \sqrt{n}(\bar{X}_n - Z).$$

The limit distribution of C_n, D_n, or W_n is a main goal in various fields, including Bayesian statistics, discrete-time filtering, gambling, and urn problems. See [2], [4], [5], [6], [7], [8], [10], and the references therein. In fact, suppose that the next observation X_{n+1} is to be predicted conditionally on the available information $\mathcal{G}_n$. If the predictor Z_n cannot be evaluated in closed form, we need some estimate $\hat{Z}_n$, and C_n reduces to the scaled error when $\hat{Z}_n = \bar{X}_n$. Furthermore, $\bar{X}_n$ is a good estimate of Z_n under some distributional assumptions on (X_n), for instance, when (X_n) is exchangeable, as is usual in Bayesian statistics. Similarly, D_n and W_n are of interest provided Z is regarded as a random parameter. In this case, Z_n is the Bayesian estimate (of Z) under quadratic loss and $\bar{X}_n$ can often be viewed as the maximum likelihood estimate. Note also that, in the trivial case where (X_n) is independent and identically distributed and $\mathcal{G}_n = \sigma(X_1, \dots, X_n)$, we obtain $C_n = W_n = \sqrt{n}(\bar{X}_n - \mathrm{E}\,X_1)$ and $D_n = 0$. As to urn problems, X_n could be the indicator of {black ball at time n} in a multicolor urn. Then, Z_n becomes the proportion of black balls in the urn at time n and $\bar{X}_n$ the observed frequency of black balls at time n.

In Theorem 1 we give conditions for

$$(C_n, D_n) \to \mathcal{N}(0, U) \times \mathcal{N}(0, V) \quad \text{stably}, \tag{1}$$

where U and V are certain random variables, and $\mathcal{N}(0, L)$ denotes the Gaussian kernel with mean 0 and variance L. A nice consequence is that

$$W_n = C_n + D_n \to \mathcal{N}(0, U + V) \quad \text{stably}.$$

Stable convergence, in the sense of Aldous and Rényi, is a strong form of convergence in distribution; the definition is recalled in Section 2.

To check the conditions for (1), it is fundamental to know something about the convergence rates of

$$Z_{n+1} - Z_n \quad \text{and} \quad \mathrm{E}(Z_{n+1} - Z_n \mid \mathcal{G}_n).$$

Hence, such conditions become simpler when (Z_n) is a $\mathcal{G}$-martingale. Since

$$\mathrm{E}(Z_{n+1} \mid \mathcal{G}_n) = \mathrm{E}(\mathrm{E}(X_{n+2} \mid \mathcal{G}_{n+1}) \mid \mathcal{G}_n) = \mathrm{E}(X_{n+2} \mid \mathcal{G}_n) \quad \text{a.s.},$$

(Z_n) is trivially a $\mathcal{G}$-martingale in the case

$$\mathrm{P}(X_k \in \cdot \mid \mathcal{G}_n) = \mathrm{P}(X_{n+1} \in \cdot \mid \mathcal{G}_n) \quad \text{a.s. for all } 0 \le n < k. \tag{2}$$

Those ($\mathcal{G}$-adapted) sequences (X_n) satisfying (2) are investigated in [5] and are called *conditionally identically distributed with respect to* $\mathcal{G}$. Note that (2) holds if (X_n) is exchangeable and $\mathcal{G}_n = \sigma(X_1, \dots, X_n)$.

Together with Theorem 1, the main contribution of this paper is one of its applications, that is, an important special case of RRUs. Two other applications are r-step predictions and Poisson–Dirichlet sequences. We refer the reader to Subsections 4.1 and 4.2 for the latter, and we next describe this type of urn.

An urn contains black and red balls. At each time $n \ge 1$, a ball is drawn and then replaced together with a random number of balls of the same color. Say that B_n black balls or R_n red

https://doi.org/10.1239/jap/1308662642 Published online by Cambridge University Press

balls are added to the urn according to whether $X_n = 1$ or $X_n = 0$, where X_n is the indicator of {black ball at time n}. Define

$$\mathcal{G}_n = \sigma(X_1, B_1, R_1, \dots, X_n, B_n, R_n),$$

and suppose that

$$B_n \geq 0, \qquad R_n \geq 0, \qquad \mathrm{E}\, B_n = \mathrm{E}\, R_n,$$
$$\sup_n \mathrm{E}((B_n + R_n)^u) < \infty \quad \text{for some } u > 2,$$
$$m := \lim_n \mathrm{E}\, B_n > 0, \qquad q := \lim_n \mathrm{E}\, B_n^2, \qquad s := \lim_n \mathrm{E}\, R_n^2,$$
$$(B_n, R_n) \text{ independent of } (X_1, B_1, R_1, \dots, X_{n-1}, B_{n-1}, R_{n-1}, X_n).$$

Then, as shown in Corollary 3 below, condition (1) holds with

$$U = Z(1-Z)\bigg(\frac{(1-Z)q + Zs}{m^2} - 1\bigg) \quad \text{and} \quad V = Z(1-Z)\frac{(1-Z)q + Zs}{m^2}.$$

A remark on the assumption $\mathrm{E}\, B_n = \mathrm{E}\, R_n$ is in order. Such an assumption is technically fundamental for Corollary 3, but it is not required by RRUs, as defined in [9]. Indeed, $\mathrm{E}\, B_n \neq \mathrm{E}\, R_n$ is closer to the spirit of RRUs and those real problems motivating them. However, $\mathrm{E}\, B_n = \mathrm{E}\, R_n$ is an important special case of RRUs. For instance, it might be the null hypothesis in an application.

Corollary 3 improves the existing result on this type of urn, obtained in [2], in two ways. First, Corollary 3 implies convergence of the pairs (C_n, D_n) and not only of D_n. Hence, we also get $W_n \to \mathcal{N}(0, U + V)$ stably. Second, unlike [2], neither the sequence $((B_n, R_n))$ is identically distributed nor the random variables $B_n + R_n$ have compact support.

By the same argument used for two-color urns, multicolor versions of Corollary 3 are easily manufactured. To the authors' knowledge, results of this type are not currently available. Briefly, for a d-color urn, let $X_{n,j}$ be the indicator of {ball of color j at time n}, where $n \geq 1$ and $1 \leq j \leq d$. Suppose that $A_{n,j}$ balls of color j are added in the case $X_{n,j} = 1$. The random variables $A_{n,j}$ satisfy the same type of conditions stated above for B_n and R_n; see Subsection 4.4 for details. Then

$$(\boldsymbol{C}_n, \boldsymbol{D}_n) \to \mathcal{N}_d(0, \boldsymbol{U}) \times \mathcal{N}_d(0, \boldsymbol{V}) \quad \text{stably},$$

where $\boldsymbol{C}_n$ and $\boldsymbol{D}_n$ are the vectorial versions of C_n and D_n, while $\boldsymbol{U}$ and $\boldsymbol{V}$ are certain random covariance matrices; see Corollary 5.

A last note is the following. In the previous urn, the nth matrix of reinforcements is

$$\boldsymbol{A}_n = \mathrm{diag}(A_{n,1}, \dots, A_{n,d}).$$

Since $\mathrm{E}\, A_{n,1} = \dots = \mathrm{E}\, A_{n,d}$, the leading eigenvalue of the mean matrix $\mathrm{E}\, \boldsymbol{A}_n$ has multiplicity greater than 1. Even if significant for applications, this particular case (the leading eigenvalue of $\mathrm{E}\, \boldsymbol{A}_n$ is not simple) is typically neglected; see [3], [12], [13], and [16, p. 20]. Our result, and indeed the result in [2], contribute to (partially) filling this gap.

2. Stable convergence

Stable convergence has been introduced by Rényi in [18] and subsequently investigated by various authors. In a sense, it is intermediate between convergence in distribution and convergence in probability. We recall here basic definitions. For more information, we refer the reader to [1], [8], [11], and the references therein.

Let $(\Omega, \mathcal{A}, \mathrm{P})$ be a probability space, and let S be a metric space. A *kernel* on S, or a *random probability measure* on S, is a measurable collection $N = \{N(\omega) : \omega \in \Omega\}$ of probability measures on the Borel σ-field on S. Measurability means that

$$N(\omega)(f) = \int f(x) N(\omega)(\mathrm{d}x)$$

is $\mathcal{A}$-measurable, as a function of $\omega \in \Omega$, for each bounded Borel map $f : S \to \mathbb{R}$.

Let (Y_n) be a sequence of S-valued random variables, and let N be a kernel on S. Both (Y_n) and N are defined on $(\Omega, \mathcal{A}, \mathrm{P})$. Say that Y_n converges *stably* to N in the case

$$\mathrm{P}(Y_n \in \cdot \mid H) \to \mathrm{E}(N(\cdot) \mid H) \quad \text{weakly for all } H \in \mathcal{A} \text{ such that } \mathrm{P}(H) > 0.$$

Clearly, if $Y_n \to N$ stably then Y_n converges in distribution to the probability law $\mathrm{E}(N(\cdot))$ (just let $H = \Omega$). Moreover, when S is separable, it is not hard to see that $Y_n \xrightarrow{\mathrm{P}} Y$ if and only if Y_n converges stably to the kernel $N = \delta_Y$.

We next mention a strong form of stable convergence, introduced in [8], to be used later on. Let $\mathcal{F}_n \subset \mathcal{A}$ be a sub-σ-field, $n \geq 1$. Say that Y_n converges to N *stably in the strong sense*, with respect to the sequence $(\mathcal{F}_n)$, in the case

$$\mathrm{E}(f(Y_n) \mid \mathcal{F}_n) \xrightarrow{\mathrm{P}} N(f) \quad \text{for each } f \in C_b(S),$$

where $C_b(S)$ denotes the set of real bounded continuous functions on S.

Finally, we state a simple but useful fact as a lemma.

Lemma 1. *Suppose that S is a separable metric space, that C_n and D_n are S-valued random variables on $(\Omega, \mathcal{A}, \mathrm{P})$, $n \geq 1$, that M and N are kernels on S defined on $(\Omega, \mathcal{A}, \mathrm{P})$, and that $\mathcal{G} = (\mathcal{G}_n : n \geq 1)$ is an (increasing) filtration satisfying*

$$\sigma(C_n) \subset \mathcal{G}_n \quad \text{and} \quad \sigma(D_n) \subset \mathcal{G}_\infty \quad \text{for all } n,$$

where $\mathcal{G}_\infty = \sigma(\bigcup_n \mathcal{G}_n)$. If $C_n \to M$ stably and $D_n \to N$ stably in the strong sense, with respect to $\mathcal{G}$, then

$$(C_n, D_n) \to M \times N \quad \text{stably.}$$

(Here, $M \times N$ is the kernel on $S \times S$ such that $(M \times N)(\omega) = M(\omega) \times N(\omega)$ for all ω.)

Proof. By standard arguments, since S is separable and $\sigma(C_n, D_n) \subset \mathcal{G}_\infty$, it suffices to prove that $\mathrm{E}(I_H f_1(C_n) f_2(D_n)) \to \mathrm{E}(I_H M(f_1) N(f_2))$ whenever $H \in \bigcup_n \mathcal{G}_n$ and $f_1, f_2 \in C_b(S)$. Let $L_n = \mathrm{E}(f_2(D_n) \mid \mathcal{G}_n) - N(f_2)$. Since $H \in \bigcup_n \mathcal{G}_n$, there is k such that $H \in \mathcal{G}_n$ for $n \geq k$. Thus,

$$\begin{aligned}\mathrm{E}(I_H f_1(C_n) f_2(D_n)) &= \mathrm{E}(I_H f_1(C_n)\, \mathrm{E}(f_2(D_n) \mid \mathcal{G}_n)) \\ &= \mathrm{E}(I_H f_1(C_n) N(f_2)) + \mathrm{E}(I_H f_1(C_n) L_n) \quad \text{for all } n \geq k.\end{aligned}$$

Finally, $|\mathrm{E}(I_H f_1(C_n) L_n)| \leq \sup|f_1|\, \mathrm{E}\,|L_n| \to 0$, since $D_n \to N$ stably in the strong sense, and $\mathrm{E}(I_H f_1(C_n) N(f_2)) \to \mathrm{E}(I_H M(f_1) N(f_2))$ as $C_n \to M$ stably.

https://doi.org/10.1239/jap/1308662642 Published online by Cambridge University Press

3. A central limit theorem

In the sequel, $(X_n : n \geq 1)$ is a sequence of real random variables on the probability space $(\Omega, \mathcal{A}, \mathrm{P})$ and $\mathcal{G} = (\mathcal{G}_n : n \geq 0)$ is an (increasing) filtration. We assume that $\mathrm{E}\,|X_n| < \infty$, and we let

$$\bar{X}_n = \frac{1}{n}\sum_{k=1}^{n} X_k, \qquad Z_n = \mathrm{E}(X_{n+1} \mid \mathcal{G}_n), \quad \text{and} \quad C_n = \sqrt{n}(\bar{X}_n - Z_n).$$

In the case $Z_n \to Z$ a.s. for some real random variable Z, we also define $D_n = \sqrt{n}(Z_n - Z)$.

Sufficient conditions for $Z_n \to Z$ a.s. and in L_1 are $\sup_n \mathrm{E}\, X_n^2 < \infty$ and

$$\mathrm{E}((\mathrm{E}(Z_{n+1} \mid \mathcal{G}_n) - Z_n)^2) = o(n^{-3}). \tag{3}$$

In this case, in fact, (Z_n) is a uniformly integrable quasi-martingale.

We recall that a sequence (Y_n) of real integrable random variables is a *quasi-martingale* (with respect to the filtration $\mathcal{G}$) if it is $\mathcal{G}$-adapted and

$$\sum_n \mathrm{E}\,|\mathrm{E}(Y_{n+1} \mid \mathcal{G}_n) - Y_n| < \infty.$$

If (Y_n) is a quasi-martingale and $\sup_n \mathrm{E}\,|Y_n| < \infty$, then Y_n converges a.s.

Let $\mathcal{N}(a, b)$ denote the one-dimensional Gaussian law with mean a and variance $b \geq 0$ (where $\mathcal{N}(a, 0) = \delta_a$). Note that $\mathcal{N}(0, L)$ is a kernel on $\mathbb{R}$ for each real nonnegative random variable L. We are now in a position to state our CLT.

Theorem 1. *Suppose that $\sigma(X_n) \subset \mathcal{G}_n$ for each $n \geq 1$, (X_n^2) is uniformly integrable, and condition (3) holds. Let us consider the following conditions.*

(a) $(1/\sqrt{n})\,\mathrm{E}(\max_{1\leq k\leq n} k|Z_{k-1} - Z_k|) \to 0$,

(b) $(1/n)\sum_{k=1}^{n}\{X_k - Z_{k-1} + k(Z_{k-1} - Z_k)\}^2 \xrightarrow{\mathrm{P}} U$,

(c) $\sqrt{n}\,\mathrm{E}(\sup_{k\geq n} |Z_{k-1} - Z_k|) \to 0$,

(d) $n\sum_{k\geq n}(Z_{k-1} - Z_k)^2 \xrightarrow{\mathrm{P}} V$,

where U and V are real nonnegative random variables. Then, $C_n \to \mathcal{N}(0, U)$ stably under (a)–(b), and $D_n \to \mathcal{N}(0, V)$ stably in the strong sense, with respect to $\mathcal{G}$, under (c)–(d). In particular,

$$(C_n, D_n) \to \mathcal{N}(0, U) \times \mathcal{N}(0, V)$$

stably under (a)–(d).

Proof. Since $\sigma(C_n) \subset \mathcal{G}_n$ and Z can be taken $\mathcal{G}_\infty$-measurable, Lemma 1 applies. Thus, it suffices to prove that $C_n \to \mathcal{N}(0, U)$ stably and $D_n \to \mathcal{N}(0, V)$ stably in the strong sense.

Part 1: $C_n \to \mathcal{N}(0, U)$ stably. Suppose that conditions (a)–(b) hold. First note that

$$\begin{aligned}\sqrt{n}C_n &= n\bar{X}_n - nZ_n\\ &= \sum_{k=1}^{n} X_k + \sum_{k=1}^{n}((k-1)Z_{k-1} - kZ_k)\\ &= \sum_{k=1}^{n}(X_k - Z_{k-1} + k(Z_{k-1} - Z_k)),\end{aligned}$$

https://doi.org/10.1239/jap/1308662642 Published online by Cambridge University Press

where $Z_0 = \mathrm{E}(X_1 \mid \mathcal{G}_0)$. Letting

$$Y_{n,k} = \frac{X_k - Z_{k-1} + k(\mathrm{E}(Z_k \mid \mathcal{G}_{k-1}) - Z_k)}{\sqrt{n}}$$

and

$$Q_n = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} k(Z_{k-1} - \mathrm{E}(Z_k \mid \mathcal{G}_{k-1})),$$

it follows that $C_n = \sum_{k=1}^n Y_{n,k} + Q_n$. By condition (3),

$$\mathrm{E}\,|Q_n| \le \frac{1}{\sqrt{n}} \sum_{k=1}^{n} k\sqrt{\mathrm{E}((Z_{k-1} - \mathrm{E}(Z_k \mid \mathcal{G}_{k-1}))^2)} = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} \mathrm{o}(k^{-1/2}) \to 0.$$

Hence, it suffices to prove that $\sum_{k=1}^n Y_{n,k} \to \mathcal{N}(0, U)$ stably. Letting $\mathcal{F}_{n,k} = \mathcal{G}_k$, $k = 1, \dots, n$, we obtain $\mathrm{E}(Y_{n,k} \mid \mathcal{F}_{n,k-1}) = 0$ a.s. Thus, by Corollary 7 of [8], $\sum_{k=1}^n Y_{n,k} \to \mathcal{N}(0, U)$ stably whenever

(i) $\mathrm{E}\{\max_{1\le k\le n} |Y_{n,k}|\} \to 0$,

(ii) $\sum_{k=1}^n Y_{n,k}^2 \xrightarrow{\mathrm{P}} U$.

As to (i), first note that

$$\sqrt{n} \max_{1\le k\le n} |Y_{n,k}| \le \max_{1\le k\le n} |X_k - Z_{k-1}| + \sum_{k=1}^{n} k|\mathrm{E}(Z_k \mid \mathcal{G}_{k-1}) - Z_{k-1}| + \max_{1\le k\le n} k|Z_{k-1} - Z_k|.$$

Since (X_n^2) is uniformly integrable, $((X_n - Z_{n-1})^2)$ is uniformly integrable as well, and this implies that $(1/n)\,\mathrm{E}(\max_{1\le k\le n}(X_k - Z_{k-1})^2) \to 0$. By condition (3),

$$\frac{1}{\sqrt{n}} \sum_{k=1}^{n} k\,\mathrm{E}\,|\mathrm{E}(Z_k \mid \mathcal{G}_{k-1}) - Z_{k-1}| = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} \mathrm{o}(k^{-1/2}) \to 0.$$

Thus, (i) follows from condition (a).

As to (ii), write

$$\begin{aligned}\sum_{k=1}^{n} Y_{n,k}^2 &= \frac{1}{n} \sum_{k=1}^{n} (X_k - Z_{k-1} + k(Z_{k-1} - Z_k))^2 + \frac{1}{n} \sum_{k=1}^{n} k^2(\mathrm{E}(Z_k \mid \mathcal{G}_{k-1}) - Z_{k-1})^2 \\ &\quad + \frac{2}{n} \sum_{k=1}^{n} (X_k - Z_{k-1} + k(Z_{k-1} - Z_k))k(\mathrm{E}(Z_k \mid \mathcal{G}_{k-1}) - Z_{k-1}) \\ &= R_n + S_n + T_n \quad \text{say}.\end{aligned}$$

Then, $R_n \xrightarrow{\mathrm{P}} U$ by (b) and $\mathrm{E}\,|S_n| = \mathrm{E}\,S_n \to 0$ by (3). Furthermore, $T_n \xrightarrow{\mathrm{P}} 0$, since

$$\frac{T_n^2}{4} \le \frac{1}{n} \sum_{k=1}^{n} (X_k - Z_{k-1} + k(Z_{k-1} - Z_k))^2 \frac{1}{n} \sum_{k=1}^{n} k^2(\mathrm{E}(Z_k \mid \mathcal{G}_{k-1}) - Z_{k-1})^2 = R_n S_n.$$

Hence, (ii) holds, concluding the proof of part 1.

Part 2: $D_n \to \mathcal{N}(0, V)$ stably in the strong sense. Suppose that conditions (c)–(d) hold. We first recall a known result; see Example 6 of [8]. Let (L_n) be a $\mathcal{G}$-martingale such that $L_n \to L$ a.s. and in L_1 for some real random variable L. Then

$$\sqrt{n}(L_n - L) \to \mathcal{N}(0, V) \quad \text{stably in the strong sense with respect to } \mathcal{G},$$

provided that

(c*) $\sqrt{n}\,\mathrm{E}(\sup_{k\geq n} |L_{k-1} - L_k|) \to 0$,

(d*) $n \sum_{k\geq n} (L_{k-1} - L_k)^2 \xrightarrow{\mathrm{P}} V$.

Next, define $L_0 = Z_0$ and

$$L_n = Z_n - \sum_{k=0}^{n-1} (\mathrm{E}(Z_{k+1} \mid \mathcal{G}_k) - Z_k).$$

Then, (L_n) is a $\mathcal{G}$-martingale. Also, $L_n \to L$ a.s. and in L_1, for some L, as (Z_n) is a uniformly integrable quasi-martingale. In particular, $L_n - L$ can be written as $L_n - L = \sum_{k\geq n}(L_k - L_{k+1})$ a.s. Similarly, $Z_n - Z = \sum_{k\geq n}(Z_k - Z_{k+1})$ a.s. It follows that

$$\begin{aligned}
\mathrm{E}\,|D_n - \sqrt{n}(L_n - L)| &= \sqrt{n}\,\mathrm{E}\,|(Z_n - Z) - (L_n - L)| \\
&= \sqrt{n}\,\mathrm{E}\left|\sum_{k\geq n}\{(Z_k - L_k) - (Z_{k+1} - L_{k+1})\}\right| \\
&\leq \sqrt{n}\sum_{k\geq n} \mathrm{E}\,|Z_k - \mathrm{E}(Z_{k+1} \mid \mathcal{G}_k)| \\
&= \sqrt{n}\sum_{k\geq n} \mathrm{o}(k^{-3/2}) \\
&\to 0.
\end{aligned}$$

Thus, $D_n \to \mathcal{N}(0, V)$ stably in the strong sense if and only if $\sqrt{n}(L_n - L) \to \mathcal{N}(0, V)$ stably in the strong sense, and to conclude the proof, it suffices to check conditions (c*)–(d*). In turn, (c*)–(d*) are a straightforward consequence of conditions (3), (c), (d), and

$$L_{k-1} - L_k = (Z_{k-1} - Z_k) + (\mathrm{E}(Z_k \mid \mathcal{G}_{k-1}) - Z_{k-1}).$$

This completes the proof of Theorem 1.

Some remarks on Theorem 1 are in order. In real problems, one of the quantities of main interest is

$$W_n = \sqrt{n}(\bar{X}_n - Z).$$

Under the assumptions of Theorem 1, we obtain

$$W_n = C_n + D_n \to \mathcal{N}(0, U + V) \quad \text{stably}.$$

Condition (3) trivially holds when (X_n) is conditionally identically distributed with respect to $\mathcal{G}$; see [5] and Section 1. In particular, (3) holds if (X_n) is exchangeable and $\mathcal{G}_n = \sigma(X_1, \dots, X_n)$.

Under Theorem 1(c), Theorem 1(a) can be replaced by

(a*) $\sup_n (1/n) \sum_{k=1}^{n} k^2\,\mathrm{E}((Z_{k-1} - Z_k)^2) < \infty$.

https://doi.org/10.1239/jap/1308662642 Published online by Cambridge University Press

Indeed, (a*) and (c) imply (a) (we omit the calculations). Note that, for proving $C_n \to \mathcal{N}(0, U)$ stably under (a*), (b), and (c), we can rely on more classical versions of the martingale CLT, such as Theorem 3.2 of [11].

To check Theorem 1(b) and (d), the following simple lemma can help.

Lemma 2. *Let (Y_n) be a $\mathcal{G}$-adapted sequence of real random variables. If $\sum_{n=1}^{\infty} n^{-2}\,\mathrm{E}\,Y_n^2 < \infty$ and $\mathrm{E}(Y_{n+1} \mid \mathcal{G}_n) \to Y$ a.s. for some random variable Y, then*

$$n\sum_{k\geq n}\frac{Y_k}{k^2} \to Y \quad \textit{a.s.} \quad \textit{and} \quad \frac{1}{n}\sum_{k=1}^{n} Y_k \to Y \quad \textit{a.s.}$$

Proof. Let $L_n = \sum_{k=1}^{n}(Y_k - \mathrm{E}(Y_k \mid \mathcal{G}_{k-1}))/k$. Then, L_n is a $\mathcal{G}$-martingale such that

$$\sup_n \mathrm{E}\,L_n^2 \leq 4\sum_k \frac{\mathrm{E}\,Y_k^2}{k^2} < \infty.$$

Thus, L_n converges a.s. and the Abel summation formula yields

$$n\sum_{k\geq n}\frac{Y_k - \mathrm{E}(Y_k \mid \mathcal{G}_{k-1})}{k^2} \to 0 \quad \text{a.s.}$$

Since $\mathrm{E}(Y_{n+1} \mid \mathcal{G}_n) \to Y$ a.s. and $n\sum_{k\geq n} 1/k^2 \to 1$, it follows that

$$n\sum_{k\geq n}\frac{Y_k}{k^2} = n\sum_{k\geq n}\frac{Y_k - \mathrm{E}(Y_k \mid \mathcal{G}_{k-1})}{k^2} + n\sum_{k\geq n}\frac{\mathrm{E}(Y_k \mid \mathcal{G}_{k-1})}{k^2} \to Y \quad \text{a.s.}$$

Similarly, the Kroneker lemma and $\mathrm{E}(Y_{n+1} \mid \mathcal{G}_n) \to Y$ a.s. yield

$$\frac{1}{n}\sum_{k=1}^{n} Y_k = \frac{1}{n}\sum_{k=1}^{n}\mathrm{E}(Y_k \mid \mathcal{G}_{k-1}) + \frac{1}{n}\sum_{k=1}^{n} k\,\frac{Y_k - \mathrm{E}(Y_k \mid \mathcal{G}_{k-1})}{k} \to Y \quad \text{a.s.}$$

This completes the proof.

Finally, as regards D_n, a natural question to ask is whether

$$\mathrm{E}(f(D_n) \mid \mathcal{G}_n) \to \mathcal{N}(0, V)(f) \quad \text{a.s. for each } f \in C_b(\mathbb{R}). \tag{4}$$

This is a strengthening of $D_n \to \mathcal{N}(0, V)$ stably in the strong sense, as $\mathrm{E}(f(D_n) \mid \mathcal{G}_n)$ is requested to converge a.s. and not only in probability. Conditions for (4) are given by the next proposition.

Proposition 1. *Let (X_n) be a (nonnecessarily $\mathcal{G}$-adapted) sequence of integrable random variables. Condition (4) holds whenever (Z_n) is uniformly integrable and*

$$\sum_{k\geq 1}\sqrt{k}\,\mathrm{E}\,|\mathrm{E}(Z_k \mid \mathcal{G}_{k-1}) - Z_{k-1}| < \infty,$$

$$\mathrm{E}\Big(\sup_{k\geq 1}\sqrt{k}|Z_{k-1} - Z_k|\Big) < \infty, \qquad n\sum_{k\geq n}(Z_{k-1} - Z_k)^2 \to V \quad \textit{a.s.}$$

Proof. Just repeat (the second part of) the proof of Theorem 1, but use Theorem 2.2 of [7] instead of Example 6 of [8].

4. Applications

4.1. *r*-step predictions

Suppose that we want to make conditional forecasts on a sequence of events $A_n \in \mathcal{G}_n$. To fix ideas, for each n, we aim to predict

$$A_n^* = \Big(\bigcap_{j\in J} A_{n+j}\Big) \cap \Big(\bigcap_{j\in J^c} A_{n+j}^c\Big)$$

conditionally on $\mathcal{G}_n$, where J is a given subset of $\{1, \dots, r\}$ and $J^c = \{1, \dots, r\} \setminus J$. Letting $X_n = I_{A_n}$, the predictor can be written as

$$Z_n^* = \mathrm{E}\bigg(\prod_{j\in J} X_{n+j} \prod_{j\in J^c} (1 - X_{n+j}) \;\bigg|\; \mathcal{G}_n\bigg).$$

In the spirit of Section 1, when Z_n^* cannot be evaluated in closed form, we need to estimate it. Under some assumptions, in particular when (X_n) is exchangeable and $\mathcal{G}_n = \sigma(X_1, \dots, X_n)$, a reasonable estimate of Z_n^* is $\bar{X}_n^h(1 - \bar{X}_n)^{r-h}$, where $h = \mathrm{card}(J)$. Usually, under such assumptions, we also have $Z_n \to Z$ a.s. and $Z_n^* \to Z^h(1 - Z)^{r-h}$ a.s. for some random variable Z. So, it makes sense to define

$$C_n^* = \sqrt{n}\{\bar{X}_n^h(1 - \bar{X}_n)^{r-h} - Z_n^*\}, \qquad D_n^* = \sqrt{n}\{Z_n^* - Z^h(1 - Z)^{r-h}\}.$$

The following result is a straightforward consequence of Theorem 1.

Corollary 1. *Let (X_n) be a $\mathcal{G}$-adapted sequence of indicators satisfying (3). If conditions (a)–(d) of Theorem 1 hold then*

$$(C_n^*, D_n^*) \to \mathcal{N}(0, \sigma^2 U) \times \mathcal{N}(0, \sigma^2 V) \quad \textit{stably},$$

where

$$\sigma^2 = \{hZ^{h-1}(1 - Z)^{r-h} - (r - h)Z^h(1 - Z)^{r-h-1}\}^2.$$

Proof. We just give a sketch of the proof. Let $f(x) = x^h(1-x)^{r-h}$. Based on Theorem 1(c), it can be shown that $\sqrt{n}\,\mathrm{E}\,|Z_n^* - f(Z_n)| \to 0$. Thus, C_n^* can be replaced by $\sqrt{n}\{f(\bar{X}_n) - f(Z_n)\}$ and D_n^* can be replaced by $\sqrt{n}\{f(Z_n) - f(Z)\}$. By the mean value theorem,

$$\sqrt{n}\{f(\bar{X}_n) - f(Z_n)\} = \sqrt{n}\, f'(M_n)(\bar{X}_n - Z_n) = f'(M_n)C_n,$$

where M_n is between $\bar{X}_n$ and Z_n. By (3), $Z_n \to Z$ a.s. and $\bar{X}_n \to Z$ a.s. Hence, $f'(M_n) \to f'(Z)$ a.s. as f' is continuous. By Theorem 1, $C_n \to \mathcal{N}(0, U)$ stably. Thus,

$$\sqrt{n}\{f(\bar{X}_n) - f(Z_n)\} \to f'(Z)\mathcal{N}(0, U) = \mathcal{N}(0, \sigma^2 U) \quad \text{stably}.$$

By a similar argument, it can be seen that $\sqrt{n}\{f(Z_n) - f(Z)\} \to \mathcal{N}(0, \sigma^2 V)$ stably in the strong sense. An application of Lemma 1 concludes the proof.

Roughly speaking, Corollary 1 states that if 1-step predictions behave nicely then r-step predictions behave nicely as well. In fact, (C_n^*, D_n^*) converges stably under the same conditions which imply convergence of (C_n, D_n), and the respective limits are connected in a simple way. Forthcoming Subsections 4.2 and 4.3 provide examples of indicators satisfying the assumptions of Corollary 1.

https://doi.org/10.1239/jap/1308662642 Published online by Cambridge University Press

4.2. Poisson–Dirichlet sequences

Let $\mathcal{Y}$ be a finite set, and let (Y_n) be a sequence of $\mathcal{Y}$-valued random variables satisfying

$$P(Y_{n+1} \in A \mid Y_1, \dots, Y_n) = \frac{\sum_{y \in A} (S_{n,y} - \alpha) I_{\{S_{n,y} \neq 0\}} + (\theta + \alpha \sum_{y \in \mathcal{Y}} I_{\{S_{n,y} \neq 0\}}) \nu(A)}{\theta + n}$$

a.s. for all $A \subset \mathcal{Y}$ and $n \geq 1$. Here, $0 \leq \alpha < 1$ and $\theta > -\alpha$ are constants, ν is the probability distribution of Y_1, and $S_{n,y} = \sum_{k=1}^{n} I_{\{Y_k = y\}}$.

Sequences (Y_n) of this type play a role in various frameworks, mainly in population genetics. They can be regarded as a generalization of those exchangeable sequences directed by a two-parameter Poisson–Dirichlet process; see [17]. For $\alpha = 0$, (Y_n) reduces to a classical Dirichlet sequence (i.e. an exchangeable sequence directed by a Dirichlet process). But, for $\alpha \neq 0$, (Y_n) may even fail to be exchangeable.

From the point of view of Theorem 1, however, it is only important that $P(Y_{n+1} \in \cdot \mid Y_1, \dots, Y_n)$ can be written down explicitly. Indeed, the following result is available.

Corollary 2. *Let $\mathcal{G}_n = \sigma(Y_1, \dots, Y_n)$ and $X_n = I_A(Y_n)$, where $A \subset \mathcal{Y}$. Then, condition (3) holds (so that $Z_n \to Z$ a.s.) and*

$$(C_n, D_n) \to \delta_0 \times \mathcal{N}(0, Z(1-Z)) \quad \textit{stably}.$$

Proof. Let $Q_n = -\alpha \sum_{y \in A} I_{\{S_{n,y} \neq 0\}} + (\theta + \alpha \sum_{y \in \mathcal{Y}} I_{\{S_{n,y} \neq 0\}}) \nu(A)$. Since

$$Z_n = P(Y_{n+1} \in A \mid Y_1, \dots, Y_n) = \frac{n \bar{X}_n + Q_n}{\theta + n} \quad \text{and} \quad |Q_n| \leq c$$

for some constant c, then $C_n \to 0$ a.s. By Lemma 1 and Theorem 1, it thus suffices to check (3), and Theorem 1(c) and (d) with $V = Z(1-Z)$. On noting that

$$Z_{n+1} - Z_n = \frac{X_{n+1} - Z_n}{\theta + n + 1} + \frac{Q_{n+1} - Q_n}{\theta + n + 1},$$

condition (c) trivially holds. Since $S_{n+1,y} = S_{n,y} + I_{\{Y_{n+1} = y\}}$, we obtain

$$Q_{n+1} - Q_n = -\alpha \nu(A^c) \sum_{y \in A} I_{\{S_{n,y} = 0\}} I_{\{Y_{n+1} = y\}} + \alpha \nu(A) \sum_{y \in A^c} I_{\{S_{n,y} = 0\}} I_{\{Y_{n+1} = y\}}.$$

It follows that

$$E(|Q_{n+1} - Q_n| \mid \mathcal{G}_n) \leq 2 \sum_{y \in \mathcal{Y}} I_{\{S_{n,y} = 0\}} P(Y_{n+1} = y \mid \mathcal{G}_n) \leq \frac{d}{\theta + n} \quad \text{a.s.}$$

for some constant d, and this implies that

$$|E(Z_{n+1} \mid \mathcal{G}_n) - Z_n| = \frac{|E(Q_{n+1} - Q_n \mid \mathcal{G}_n)|}{\theta + n + 1} \leq \frac{d}{(\theta + n)^2} \quad \text{a.s.}$$

Hence, condition (3) holds. To check (d), note that $\sum_k k^2 E((Z_{k-1} - Z_k)^4) < \infty$. Since $Z_k \to Z$ a.s. (by (3)), we also obtain

$$E((X_k - Z_{k-1})^2 \mid \mathcal{G}_{k-1}) = Z_{k-1} - Z_{k-1}^2 \to Z(1-Z) \quad \text{a.s.},$$
$$E((Q_k - Q_{k-1})^2 \mid \mathcal{G}_{k-1}) + 2 E((X_k - Z_{k-1})(Q_k - Q_{k-1}) \mid \mathcal{G}_{k-1}) \to 0 \quad \text{a.s.}$$

Thus,

$$k^2 \operatorname{E}((Z_{k-1} - Z_k)^2 \mid \mathcal{G}_{k-1}) \to Z(1 - Z) \quad \text{a.s.}$$

Letting $Y_k = k^2(Z_{k-1} - Z_k)^2$ and $Y = Z(1 - Z)$, Lemma 2 implies that

$$n \sum_{k \geq n} (Z_{k-1} - Z_k)^2 = n \sum_{k \geq n} \frac{Y_k}{k^2} \to Z(1 - Z) \quad \text{a.s.}$$

This completes the proof.

As is clear from the previous proof, all assumptions of Proposition 1 are satisfied. Therefore, D_n meets condition (4) with $V = Z(1 - Z)$.

A result analogous to Corollary 2 is Theorem 4.1 of [4]. The main tool for proving the latter, indeed, is Theorem 1.

4.3. Two-color randomly reinforced urns

An urn contains $b > 0$ black balls and $r > 0$ red balls. At each time $n \geq 1$, a ball is drawn and then replaced together with a random number of balls of the same color. Say that B_n black balls or R_n red balls are added to the urn according to whether $X_n = 1$ or $X_n = 0$, where X_n is the indicator of {black ball at time n}.

Urns of this type have some history starting with [9]. See also [2], [4], [5], [7], [15], [16], and the references therein.

To model such urns, we assume that X_n, B_n, and R_n are random variables on the probability space $(\Omega, \mathcal{A}, \mathrm{P})$ such that

$$X_n \in \{0, 1\}, \qquad B_n \geq 0, \qquad R_n \geq 0, \tag{5}$$

$$(B_n, R_n) \text{ independent of } (X_1, B_1, R_1, \dots, X_{n-1}, B_{n-1}, R_{n-1}, X_n),$$

$$Z_n = \mathrm{P}(X_{n+1} = 1 \mid \mathcal{G}_n) = \frac{b + \sum_{k=1}^n B_k X_k}{b + r + \sum_{k=1}^n (B_k X_k + R_k(1 - X_k))} \quad \text{a.s.}$$

for each $n \geq 1$, where

$$\mathcal{G}_0 = \{\varnothing, \Omega\}, \qquad \mathcal{G}_n = \sigma(X_1, B_1, R_1, \dots, X_n, B_n, R_n).$$

In the particular case $B_n = R_n$, in Example 3.5 of [5], it was shown that C_n converges stably to a Gaussian kernel whenever $\operatorname{E} B_1^2 < \infty$ and the sequence $(B_n : n \geq 1)$ is identically distributed. Furthermore, in Corollary 4.1 of [7], D_n is shown to satisfy condition (4). The latter result on D_n is extended to $B_n \neq R_n$ in [2], under the assumptions that $B_1 + R_1$ has compact support, $\operatorname{E} B_1 = \operatorname{E} R_1$, and $((B_n, R_n) : n \geq 1)$ is identically distributed.

Based on the results in Section 3, condition (4) can be shown to hold more generally. Indeed, to get condition (4), it is fundamental that $\operatorname{E} B_n = \operatorname{E} R_n$ for all n and the three sequences $(\operatorname{E} B_n)$, $(\operatorname{E} B_n^2)$, and $(\operatorname{E} R_n^2)$ approach a limit. But the identity assumption for distributions of (B_n, R_n) can be dropped, and compact support of $B_n + R_n$ can be replaced by a moment condition such as

$$\sup_n \operatorname{E}((B_n + R_n)^u) < \infty \quad \text{for some } u > 2. \tag{6}$$

Under these conditions, not only does D_n meet (4), but the pairs (C_n, D_n) converge stably as well. In particular, we obtain stable convergence of $W_n = C_n + D_n$, which is of potential interest in urn problems.

Corollary 3. *In addition to (5) and (6), suppose that* $\mathrm{E}\,B_n = \mathrm{E}\,R_n$ *for all* n *and*

$$m := \lim_n \mathrm{E}\,B_n > 0, \qquad q := \lim_n \mathrm{E}\,B_n^2, \qquad s := \lim_n \mathrm{E}\,R_n^2.$$

Then, condition (3) holds (so that $Z_n \to Z$ *a.s.) and*

$$(C_n, D_n) \to \mathcal{N}(0, U) \times \mathcal{N}(0, V) \quad \textit{stably},$$

where

$$U = Z(1-Z)\Big(\frac{(1-Z)q + Zs}{m^2} - 1\Big) \quad \textit{and} \quad V = Z(1-Z)\frac{(1-Z)q + Zs}{m^2}.$$

In particular, $W_n = C_n + D_n \to \mathcal{N}(0, U+V)$ *stably. Moreover,* D_n *meets condition (4), that is,* $\mathrm{E}(f(D_n) \mid \mathcal{G}_n) \to \mathcal{N}(0, V)(f)$ *a.s. for each* $f \in C_b(\mathbb{R})$.

It is worth noting that, arguing as in [2] and [15], we obtain $\mathrm{P}(Z = z) = 0$ for all z. Thus, $\mathcal{N}(0, V)$ is a nondegenerate kernel. In turn, $\mathcal{N}(0, U)$ is nondegenerate unless $q = s = m^2$, and this happens if and only if both B_n and R_n converge in probability (necessarily to m). In the latter case ($q = s = m^2$), $C_n \xrightarrow{\mathrm{P}} 0$ and condition (4) holds with $V = Z(1-Z)$. Thus, in a sense, RRUs behave as classical Polya urns (i.e. those urns with $B_n = R_n = m$) whenever the reinforcements converge in probability.

The proof of Corollary 3 is deferred to Appendix A as it needs some work. Here, to point out the underlying argument, we sketch such a proof under the superfluous but simplifying assumption that $B_n \vee R_n \le c$ for all n and some constant c. Let

$$S_n = b + r + \sum_{k=1}^{n} (B_k X_k + R_k(1 - X_k)).$$

After some algebra, $Z_{n+1} - Z_n$ can be written as

$$\begin{aligned} Z_{n+1} - Z_n &= \frac{(1-Z_n)X_{n+1}B_{n+1} - Z_n(1-X_{n+1})R_{n+1}}{S_{n+1}} \\ &= \frac{(1-Z_n)X_{n+1}B_{n+1}}{S_n + B_{n+1}} - \frac{Z_n(1-X_{n+1})R_{n+1}}{S_n + R_{n+1}}. \end{aligned}$$

By (5) and $\mathrm{E}\,B_{n+1} = \mathrm{E}\,R_{n+1}$,

$$\begin{aligned} \mathrm{E}(Z_{n+1} - Z_n \mid \mathcal{G}_n) &= Z_n(1-Z_n)\,\mathrm{E}\Big(\frac{B_{n+1}}{S_n + B_{n+1}} - \frac{R_{n+1}}{S_n + R_{n+1}} \;\Big|\; \mathcal{G}_n\Big) \\ &= Z_n(1-Z_n)\,\mathrm{E}\Big(\frac{B_{n+1}}{S_n + B_{n+1}} - \frac{B_{n+1}}{S_n} - \frac{R_{n+1}}{S_n + R_{n+1}} + \frac{R_{n+1}}{S_n} \;\Big|\; \mathcal{G}_n\Big) \\ &= Z_n(1-Z_n)\,\mathrm{E}\Big(-\frac{B_{n+1}^2}{S_n(S_n + B_{n+1})} + \frac{R_{n+1}^2}{S_n(S_n + R_{n+1})} \;\Big|\; \mathcal{G}_n\Big) \quad \text{a.s.} \end{aligned}$$

Thus,

$$|\mathrm{E}(Z_{n+1} \mid \mathcal{G}_n) - Z_n| \le \frac{\mathrm{E}\,B_{n+1}^2 + \mathrm{E}\,R_{n+1}^2}{S_n^2} \quad \text{a.s.}$$

Since $\sup_n(\mathrm{E}\, B_n^2 + \mathrm{E}\, R_n^2) < \infty$ and $\mathrm{E}(S_n^{-p}) = \mathrm{O}(n^{-p})$ for all $p > 0$ (as shown in Lemma 3), then

$$\mathrm{E}(|\mathrm{E}(Z_{n+1} \mid \mathcal{G}_n) - Z_n|^p) = \mathrm{O}(n^{-2p}) \quad \text{for all } p > 0.$$

In particular, condition (3) holds and $\sum_k \sqrt{k}\, \mathrm{E}\, |\mathrm{E}(Z_k \mid \mathcal{G}_{k-1}) - Z_{k-1}| < \infty$.

To conclude the proof, in view of Lemma 1, Theorem 1, and Proposition 1, it suffices to check Theorem 1(a) and (b), and

(i) $\mathrm{E}(\sup_{k\geq 1} \sqrt{k}|Z_{k-1} - Z_k|) < \infty$,

(ii) $n \sum_{k\geq n}(Z_{k-1} - Z_k)^2 \to V$ a.s.

Conditions (a) and (i) are straightforward consequences of $|Z_{n+1} - Z_n| \leq c/S_n$ and $\mathrm{E}(S_n^{-p}) = \mathrm{O}(n^{-p})$ for all $p > 0$. Condition (b) follows from the same argument as (ii). To prove (ii), it suffices to show that $\mathrm{E}(Y_{n+1} \mid \mathcal{G}_n) \to V$ a.s., where $Y_n = n^2(Z_{n-1} - Z_n)^2$; see Lemma 2. Write $(n+1)^{-2}\, \mathrm{E}(Y_{n+1} \mid \mathcal{G}_n)$ as

$$Z_n(1 - Z_n)^2\, \mathrm{E}\biggl(\frac{B_{n+1}^2}{(S_n + B_{n+1})^2} \biggm| \mathcal{G}_n\biggr) + Z_n^2(1 - Z_n)\, \mathrm{E}\biggl(\frac{R_{n+1}^2}{(S_n + R_{n+1})^2} \biggm| \mathcal{G}_n\biggr).$$

Since $S_n/n \to m$ a.s. (by Lemma 3) and $B_{n+1} \leq c$, then

$$n^2\, \mathrm{E}\biggl(\frac{B_{n+1}^2}{(S_n + B_{n+1})^2} \biggm| \mathcal{G}_n\biggr) \leq n^2\, \mathrm{E}\biggl(\frac{B_{n+1}^2}{S_n^2} \biggm| \mathcal{G}_n\biggr) = n^2 \frac{\mathrm{E}\, B_{n+1}^2}{S_n^2} \to \frac{q}{m^2} \quad \text{a.s.}$$

and

$$n^2\, \mathrm{E}\biggl(\frac{B_{n+1}^2}{(S_n + B_{n+1})^2} \biggm| \mathcal{G}_n\biggr) \geq n^2\, \mathrm{E}\biggl(\frac{B_{n+1}^2}{(S_n + c)^2} \biggm| \mathcal{G}_n\biggr) = n^2 \frac{\mathrm{E}\, B_{n+1}^2}{(S_n + c)^2} \to \frac{q}{m^2} \quad \text{a.s.}$$

Similarly,

$$n^2\, \mathrm{E}\biggl(\frac{R_{n+1}^2}{(S_n + R_{n+1})^2} \biggm| \mathcal{G}_n\biggr) \to \frac{s}{m^2} \quad \text{a.s.}$$

Since $Z_n \to Z$ a.s., it follows that

$$\mathrm{E}(Y_{n+1} \mid \mathcal{G}_n) \xrightarrow{\text{a.s.}} Z(1 - Z)^2 \frac{q}{m^2} + Z^2(1 - Z) \frac{s}{m^2} = V.$$

This concludes the (sketch of the) proof.

Remark 1. In order for $(C_n, D_n) \to \mathcal{N}(0, U) \times \mathcal{N}(0, V)$ stably, some of the assumptions of Corollary 3 can be stated in a different form. We mention two (independent) facts.

First, condition (6) can be weakened into uniform integrability of $(B_n + R_n)^2$.

Second, (B_n, R_n) independent of $(X_1, B_1, R_1, \ldots, X_{n-1}, B_{n-1}, R_{n-1}, X_n)$ can be replaced by the following four conditions.

(C1) (B_n, R_n) conditionally independent of X_n given $\mathcal{G}_{n-1}$.

(C2) Condition (6) holds for some $u > 4$.

https://doi.org/10.1239/jap/1308662642 Published online by Cambridge University Press

(C3) There exist an integer n_0 and a constant $l > 0$ such that

$$\mathrm{E}(B_n \wedge n^{1/4} \mid \mathcal{G}_{n-1}) \geq l \quad \text{and} \quad \mathrm{E}(R_n \wedge n^{1/4} \mid \mathcal{G}_{n-1}) \geq l \quad \text{a.s.}$$

whenever $n \geq n_0$.

(C4) There exist random variables m, q, and s such that

$$\mathrm{E}(B_n \mid \mathcal{G}_{n-1}) = \mathrm{E}(R_n \mid \mathcal{G}_{n-1}) \xrightarrow{\mathrm{P}} m,$$
$$\mathrm{E}(B_n^2 \mid \mathcal{G}_{n-1}) \xrightarrow{\mathrm{P}} q, \qquad \mathrm{E}(R_n^2 \mid \mathcal{G}_{n-1}) \xrightarrow{\mathrm{P}} s.$$

Even if in a different framework, conditions similar to (C1)–(C4) are in [3].

4.4. The multicolor case

To avoid technicalities, we firstly investigated two-color urns, but Theorem 1 applies to the multicolor case as well.

An urn contains $a_j > 0$ balls of color $j \in \{1, \dots, d\}$, where $d \geq 2$. Let $X_{n,j}$ denote the indicator of {ball of color j at time n}. In the case $X_{n,j} = 1$, the ball which has been drawn is replaced together with $A_{n,j}$ more balls of color j. Formally, we assume that $\{X_{n,j}, A_{n,j} : n \geq 1, 1 \leq j \leq d\}$ random variables on the probability space $(\Omega, \mathcal{A}, \mathrm{P})$ satisfying

$$X_{n,j} \in \{0, 1\}, \qquad \sum_{j=1}^{d} X_{n,j} = 1, \qquad A_{n,j} \geq 0, \tag{7}$$
$$(A_{n,1}, \dots, A_{n,d}) \text{ independent of } (A_{k,j}, X_{k,j}, X_{n,j} : 1 \leq k < n,\ 1 \leq j \leq d),$$
$$Z_{n,j} = \mathrm{P}(X_{n+1,j} = 1 \mid \mathcal{G}_n) = \frac{a_j + \sum_{k=1}^{n} A_{k,j} X_{k,j}}{\sum_{i=1}^{d} a_i + \sum_{k=1}^{n} \sum_{i=1}^{d} A_{k,i} X_{k,i}} \quad \text{a.s.,}$$

where

$$\mathcal{G}_0 = \{\varnothing, \Omega\}, \qquad \mathcal{G}_n = \sigma(A_{k,j}, X_{k,j} : 1 \leq k \leq n,\ 1 \leq j \leq d).$$

Note that

$$Z_{n+1,j} - Z_{n,j} = (1 - Z_{n,j}) \frac{A_{n+1,j} X_{n+1,j}}{S_n + A_{n+1,j}} - Z_{n,j} \sum_{i \neq j} \frac{A_{n+1,i} X_{n+1,i}}{S_n + A_{n+1,i}},$$

where

$$S_n = \sum_{i=1}^{d} a_i + \sum_{k=1}^{n} \sum_{i=1}^{d} A_{k,i} X_{k,i}.$$

In addition to (7), as in Subsection 4.3, we assume the following moment condition:

$$\sup_n \mathrm{E}\bigg(\bigg(\sum_{j=1}^{d} A_{n,j}\bigg)^u\bigg) < \infty \quad \text{for some } u > 2. \tag{8}$$

Furthermore, it is assumed that

$$\mathrm{E}\, A_{n,j} = \mathrm{E}\, A_{n,1} \quad \text{for each } n \geq 1 \text{ and } 1 \leq j \leq d, \tag{9}$$

https://doi.org/10.1239/jap/1308662642 Published online by Cambridge University Press

and

$$m := \lim_n \mathrm{E}\, A_{n,1} > 0, \qquad q_j := \lim_n \mathrm{E}\, A_{n,j}^2 \quad \text{for each } 1 \le j \le d.$$

Fix $1 \le j \le d$. Since $\mathrm{E}\, A_{n,i} = \mathrm{E}\, A_{n,1}$ for all n and i, the same calculation as in Subsection 4.3 yields

$$|\mathrm{E}(Z_{n+1,j} \mid \mathcal{G}_n) - Z_{n,j}| \le \frac{\sum_{i=1}^d \mathrm{E}\, A_{n+1,i}^2}{S_n^2} \quad \text{a.s.}$$

Also, $\mathrm{E}(S_n^{-p}) = \mathrm{O}(n^{-p})$ for all $p > 0$; see Remark 2. Thus,

$$\mathrm{E}(|\mathrm{E}(Z_{n+1,j} \mid \mathcal{G}_n) - Z_{n,j}|^p) = \mathrm{O}(n^{-2p}) \quad \text{for all } p > 0.$$

In particular, $Z_{n,j}$ meets condition (3) so that $Z_{n,j} \to Z_{(j)}$ a.s. for some random variable $Z_{(j)}$. Define

$$C_{n,j} = \sqrt{n}\biggl(\frac{1}{n}\sum_{k=1}^n X_{k,j} - Z_{n,j}\biggr) \quad \text{and} \quad D_{n,j} = \sqrt{n}(Z_{n,j} - Z_{(j)}).$$

The next result is quite expected at this point.

Corollary 4. *Suppose that conditions (7), (8), and (9) hold, and fix $1 \le j \le d$. Then,*

$$(C_{n,j}, D_{n,j}) \to \mathcal{N}(0, U_j) \times \mathcal{N}(0, V_j) \quad \textit{stably},$$

where

$$U_j = V_j - Z_{(j)}(1 - Z_{(j)}) \quad \textit{and} \quad V_j = \frac{Z_{(j)}}{m^2}\biggl\{q_j(1 - Z_{(j)})^2 + Z_{(j)} \sum_{i \ne j} q_i Z_{(i)}\biggr\}.$$

Moreover, $\mathrm{E}(f(D_{n,j}) \mid \mathcal{G}_n) \to \mathcal{N}(0, V_j)(f)$ *a.s. for each* $f \in C_b(\mathbb{R})$*, that is,* $D_{n,j}$ *meets condition (4).*

Proof. Just repeat the proof of Corollary 3 with $X_{n,j}$ in the place of X_n.

A vectorial version of Corollary 4 can be obtained with slight effort. Let $\mathcal{N}_d(0, \Sigma)$ denote the d-dimensional Gaussian law with mean vector 0 and covariance matrix Σ, and let

$$\boldsymbol{C}_n = (C_{n,1}, \ldots, C_{n,d}), \qquad \boldsymbol{D}_n = (D_{n,1}, \ldots, D_{n,d}).$$

Corollary 5. *Suppose that conditions (7), (8), and (9) hold. Then,*

$$(\boldsymbol{C}_n, \boldsymbol{D}_n) \to \mathcal{N}_d(0, \boldsymbol{U}) \times \mathcal{N}_d(0, \boldsymbol{V}) \quad \textit{stably},$$

where $\boldsymbol{U}$ *and* $\boldsymbol{V}$ *are the* $d \times d$ *matrices with entries* $U_{j,j} = U_j$, $V_{j,j} = V_j$, *and*

$$U_{i,j} = V_{i,j} + Z_{(i)}Z_{(j)}, \qquad V_{i,j} = \frac{Z_{(i)}Z_{(j)}}{m^2}\biggl\{\sum_{h=1}^d q_h Z_{(h)} - q_i - q_j\biggr\}, \quad \textit{for } i \ne j.$$

Moreover, $\mathrm{E}(f(\boldsymbol{D}_n) \mid \mathcal{G}_n) \to \mathcal{N}_d(0, \boldsymbol{V})(f)$ *a.s. for each* $f \in C_b(\mathbb{R}^d)$.

Proof. Given a linear functional $\phi\colon \mathbb{R}^d \to \mathbb{R}$, it suffices to see that

$$\phi(\boldsymbol{C}_n) \to \mathcal{N}_d(0, \boldsymbol{U}) \circ \phi^{-1} \quad \text{stably},$$

and

$$\mathrm{E}(g \circ \phi(\boldsymbol{D}_n) \mid \mathcal{G}_n) \to \mathcal{N}_d(0, \boldsymbol{V})(g \circ \phi) \quad \text{a.s. for each } g \in C_b(\mathbb{R}).$$

To this purpose, note that

$$\phi(\boldsymbol{C}_n) = \sqrt{n}\biggl\{\frac{1}{n}\sum_{k=1}^{n}\phi(X_{k,1}, \dots, X_{k,d}) - \mathrm{E}(\phi(X_{n+1,1}, \dots, X_{n+1,d}) \mid \mathcal{G}_n)\biggr\},$$

$$\phi(\boldsymbol{D}_n) = \sqrt{n}\{\mathrm{E}(\phi(X_{n+1,1}, \dots, X_{n+1,d}) \mid \mathcal{G}_n) - \phi(Z_{(1)}, \dots, Z_{(d)})\},$$

and again repeat the proof of Corollary 3 with $\phi(X_{n,1}, \dots, X_{n,d})$ in the place of X_n.

A nice consequence of Corollary 5 is that

$$\boldsymbol{W}_n = \boldsymbol{C}_n + \boldsymbol{D}_n \to \mathcal{N}_d(0, \boldsymbol{U} + \boldsymbol{V}) \quad \text{stably}$$

provided that conditions (7), (8), and (9) hold, where $\boldsymbol{W}_n = (W_{n,1}, \dots, W_{n,d})$ and $W_{n,j} = \sqrt{n}((1/n)\sum_{k=1}^{n} X_{k,j} - Z_{(j)})$.

Appendix A

In the notation of Subsection 4.3, let $S_n = b + r + \sum_{k=1}^{n}(B_k X_k + R_k(1 - X_k))$.

Lemma 3. *Under the assumptions of Corollary 3,*

$$\frac{n}{S_n} \to \frac{1}{m} \quad \textit{a.s. and in } L_p \textit{ for all } p > 0.$$

Proof. Let $Y_n = B_n X_n + R_n(1 - X_n)$. By (5) and $\mathrm{E}\,B_{n+1} = \mathrm{E}\,R_{n+1}$,

$$\begin{aligned}
\mathrm{E}(Y_{n+1} \mid \mathcal{G}_n) &= \mathrm{E}\,B_{n+1}\,\mathrm{E}(X_{n+1} \mid \mathcal{G}_n) + \mathrm{E}\,R_{n+1}\,\mathrm{E}(1 - X_{n+1} \mid \mathcal{G}_n)\\
&= Z_n\,\mathrm{E}\,B_{n+1} + (1 - Z_n)\,\mathrm{E}\,B_{n+1}\\
&= \mathrm{E}\,B_{n+1}\\
&\to m \quad \text{a.s.}
\end{aligned}$$

Since $m > 0$, Lemma 2 implies that $n/S_n = 1/(S_n/n) \to 1/m$ a.s. To conclude the proof, it suffices to see that $\mathrm{E}(S_n^{-p}) = \mathrm{O}(n^{-p})$ for all $p > 0$. Given $c > 0$, define

$$S_n^{(c)} = \sum_{k=1}^{n}\{X_k(B_k \wedge c - \mathrm{E}(B_k \wedge c)) + (1 - X_k)(R_k \wedge c - \mathrm{E}(R_k \wedge c))\}.$$

By a classical martingale inequality (see, e.g. Lemma 1.5 of [14]),

$$\mathrm{P}(|S_n^{(c)}| > x) \le 2\exp\biggl(-\frac{x^2}{2c^2 n}\biggr) \quad \text{for all } x > 0.$$

Since $\mathrm{E}\,B_n = \mathrm{E}\,R_n \to m$, and both (B_n) and (R_n) are uniformly integrable (as $\sup_n(\mathrm{E}\,B_n^2 + \mathrm{E}\,R_n^2) < \infty$), there exist $c > 0$ and an integer n_0 such that

$$m_n := \sum_{k=1}^{n}\min\{\mathrm{E}(B_k \wedge c), \mathrm{E}(R_k \wedge c)\} > n\frac{m}{2} \quad \text{for all } n \ge n_0.$$

https://doi.org/10.1239/jap/1308662642 Published online by Cambridge University Press

Fix one such $c > 0$, and let $l = m/4 > 0$. For every $p > 0$, we can write

$$\begin{aligned} \mathrm{E}(S_n^{-p}) &= p \int_{b+r}^{\infty} t^{-p-1}\, \mathrm{P}(S_n < t)\, \mathrm{d}t \\ &\le \frac{p}{(b+r)^{p+1}} \int_{b+r}^{b+r+nl} \mathrm{P}(S_n < t)\, \mathrm{d}t + p \int_{b+r+nl}^{\infty} t^{-p-1}\, \mathrm{d}t. \end{aligned}$$

Clearly, $p \int_{b+r+nl}^{\infty} t^{-p-1}\, \mathrm{d}t = (b+r+nl)^{-p} = \mathrm{O}(n^{-p})$. Furthermore, for each $n \ge n_0$ and $t < b+r+nl$, since $m_n > n2l$, we obtain

$$\begin{aligned} \mathrm{P}(S_n < t) &\le \mathrm{P}(S_n^{(c)} < t - b - r - m_n) \\ &\le \mathrm{P}(S_n^{(c)} < t - b - r - n2l) \\ &\le \mathrm{P}(|S_n^{(c)}| > b + r + n2l - t) \\ &\le 2 \exp\biggl(-\frac{(b+r+n2l-t)^2}{2c^2 n}\biggr). \end{aligned}$$

Hence, $\int_{b+r}^{b+r+nl} \mathrm{P}(S_n < t)\, \mathrm{d}t \le n2l \exp(-nl^2/2c^2)$ for every $n \ge n_0$, so that $\mathrm{E}(S_n^{-p}) = \mathrm{O}(n^{-p})$.

Remark 2. As in Subsection 4.4, let $S_n = \sum_{i=1}^{d} a_i + \sum_{k=1}^{n} \sum_{i=1}^{d} A_{k,i} X_{k,i}$. Under conditions (7), (8), and (9), the previous proof still applies to such S_n. Thus, $n/S_n \to 1/m$ a.s. and in L_p for all $p > 0$.

Proof of Corollary 3. By Lemma 1, it is enough to prove that $C_n \to \mathcal{N}(0, U)$ stably and D_n meets condition (4). Recall from Subsection 4.3 that

$$Z_{n+1} - Z_n = \frac{(1 - Z_n) X_{n+1} B_{n+1} - Z_n (1 - X_{n+1}) R_{n+1}}{S_{n+1}}$$

and

$$\mathrm{E}(|\mathrm{E}(Z_{n+1} \mid \mathcal{G}_n) - Z_n|^p) = \mathrm{O}(n^{-2p}) \quad \text{for all } p > 0.$$

In particular, condition (3) holds and $\sum_k \sqrt{k}\, \mathrm{E}\, |\mathrm{E}(Z_k \mid \mathcal{G}_{k-1}) - Z_{k-1}| < \infty$.

Part 1: D_n meets condition (4). By (6) and Lemma 3,

$$\mathrm{E}(|Z_{k-1} - Z_k|^u) \le \mathrm{E}\biggl(\frac{(B_k + R_k)^u}{S_{k-1}^u}\biggr) = \mathrm{E}((B_k + R_k)^u)\, \mathrm{E}(S_{k-1}^{-u}) = \mathrm{O}(k^{-u}).$$

Thus, $\mathrm{E}(\sup_k \sqrt{k} |Z_{k-1} - Z_k|)^u \le \sum_k k^{u/2}\, \mathrm{E}(|Z_{k-1} - Z_k|^u) < \infty$ as $u > 2$. In view of Proposition 1, it remains to only prove that

$$\begin{aligned} n \sum_{k \ge n} (Z_{k-1} - Z_k)^2 &= n \sum_{k \ge n} \biggl(\frac{(1 - Z_{k-1}) X_k B_k}{S_k} - \frac{Z_{k-1}(1 - X_k) R_k}{S_k}\biggr)^2 \\ &= n \sum_{k \ge n} \frac{(1 - Z_{k-1})^2 X_k B_k^2}{(S_{k-1} + B_k)^2} + n \sum_{k \ge n} \frac{Z_{k-1}^2 (1 - X_k) R_k^2}{(S_{k-1} + R_k)^2} \end{aligned}$$

converges a.s. to $V = Z(1 - Z)((1 - Z)q + Zs)/m^2$. It is enough to show that

$$n \sum_{k \ge n} \frac{(1 - Z_{k-1})^2 X_k B_k^2}{(S_{k-1} + B_k)^2} \to Z(1 - Z)^2 \frac{q}{m^2} \quad \text{a.s.}$$

and

$$n\sum_{k\geq n}\frac{Z_{k-1}^2(1-X_k)R_k^2}{(S_{k-1}+R_k)^2}\to Z^2(1-Z)\frac{s}{m^2}\quad\text{a.s.}$$

These two limit relations can be proved by exactly the same argument, and, thus, we just prove the first one. Let $U_n=B_nI_{\{B_n\leq\sqrt{n}\}}$. Since $\mathrm{P}(B_n>\sqrt{n})\leq n^{-u/2}\,\mathrm{E}\,B_n^u$, condition (6) yields $\mathrm{P}(B_n\neq U_n$ infinitely often$)=0$. Hence, it suffices to show that

$$n\sum_{k\geq n}\frac{(1-Z_{k-1})^2X_kU_k^2}{(S_{k-1}+U_k)^2}\to Z(1-Z)^2\frac{q}{m^2}\quad\text{a.s.}\tag{10}$$

Let

$$Y_n=n^2\frac{(1-Z_{n-1})^2X_nU_n^2}{(S_{n-1}+U_n)^2}.$$

Since (B_n^2) is uniformly integrable, $\mathrm{E}\,U_n^2\to q$. Furthermore, $S_n/n\to m$ a.s. and $Z_n\to Z$ a.s. Thus,

$$\begin{aligned}\mathrm{E}(Y_{n+1}\mid\mathcal{G}_n)&\leq(1-Z_n)^2(n+1)^2\,\mathrm{E}\bigg(\frac{X_{n+1}U_{n+1}^2}{S_n^2}\;\bigg|\;\mathcal{G}_n\bigg)\\&=Z_n(1-Z_n)^2\frac{(n+1)^2}{S_n^2}\,\mathrm{E}\,U_{n+1}^2\to Z(1-Z)^2\frac{q}{m^2}\quad\text{a.s.}\end{aligned}$$

and

$$\begin{aligned}\mathrm{E}(Y_{n+1}\mid\mathcal{G}_n)&\geq(1-Z_n)^2(n+1)^2\,\mathrm{E}\bigg(\frac{X_{n+1}U_{n+1}^2}{(S_n+\sqrt{n+1})^2}\;\bigg|\;\mathcal{G}_n\bigg)\\&=Z_n(1-Z_n)^2\frac{(n+1)^2}{(S_n+\sqrt{n+1})^2}\,\mathrm{E}\,U_{n+1}^2\\&\to Z(1-Z)^2\frac{q}{m^2}\quad\text{a.s.}\end{aligned}$$

By Lemma 2, to obtain relation (10), it suffices that $\sum_n\mathrm{E}\,Y_n^2/n^2<\infty$. Since

$$\frac{\mathrm{E}\,U_n^4}{n^2}\leq\frac{\mathrm{E}(B_n^2I_{\{B_n^2\leq\sqrt{n}\}})}{n^{3/2}}+\frac{\mathrm{E}(B_n^2I_{\{B_n^2>\sqrt{n}\}})}{n}\leq\frac{\mathrm{E}\,B_n^2}{n^{3/2}}+\frac{\mathrm{E}\,B_n^u}{n^{1+(u-2)/4}},$$

condition (6) implies that $\sum_n\mathrm{E}\,U_n^4/n^2<\infty$. By Lemma 3, $\mathrm{E}(S_{n-1}^{-4})=\mathrm{O}(n^{-4})$. Then,

$$\sum_n\frac{\mathrm{E}\,Y_n^2}{n^2}\leq\sum_n n^2\,\mathrm{E}\bigg(\frac{U_n^4}{S_{n-1}^4}\bigg)=\sum_n n^2\,\mathrm{E}(S_{n-1}^{-4})\,\mathrm{E}\,U_n^4\leq c\sum_n\frac{\mathrm{E}\,U_n^4}{n^2}<\infty$$

for some constant c. Hence, condition (10) holds.

Part 2: $C_n\to\mathcal{N}(0,U)$ stably. It suffices to check Theorem 1(a) and (b) with $U=Z(1-Z)(((1-Z)q+Zs)/m^2-1)$. As to (a), since $\mathrm{E}(|Z_{k-1}-Z_k|^u)=\mathrm{O}(k^{-u})$,

$$\Big(n^{-1/2}\,\mathrm{E}\Big(\max_{1\leq k\leq n}k|Z_{k-1}-Z_k|\Big)\Big)^u\leq n^{-u/2}\sum_{k=1}^n k^u\,\mathrm{E}(|Z_{k-1}-Z_k|^u)\to 0.$$

We now prove condition (b). After some algebra, we obtain

$$\begin{aligned}\mathrm{E}((X_n - Z_{n-1})(Z_{n-1} - Z_n) \mid \mathcal{G}_{n-1}) \\ = -Z_{n-1}(1 - Z_{n-1})\,\mathrm{E}\bigg(\frac{B_n}{S_{n-1} + B_n} \,\bigg|\, \mathcal{G}_{n-1}\bigg) \\ + Z_{n-1}^2(1 - Z_{n-1})\,\mathrm{E}\bigg(\frac{B_n}{S_{n-1} + B_n} - \frac{R_n}{S_{n-1} + R_n} \,\bigg|\, \mathcal{G}_{n-1}\bigg) \quad \text{a.s.}\end{aligned}$$

Arguing as in part 1,

$$n\,\mathrm{E}\bigg(\frac{B_n}{S_{n-1} + B_n} \,\bigg|\, \mathcal{G}_{n-1}\bigg) \to 1 \quad \text{a.s.} \quad \text{and} \quad n\,\mathrm{E}\bigg(\frac{R_n}{S_{n-1} + R_n} \,\bigg|\, \mathcal{G}_{n-1}\bigg) \to 1 \quad \text{a.s.}$$

Thus, $n\,\mathrm{E}((X_n - Z_{n-1})(Z_{n-1} - Z_n) \mid \mathcal{G}_{n-1}) \to -Z(1 - Z)$ a.s. Furthermore,

$$\mathrm{E}((X_n - Z_{n-1})^2 \mid \mathcal{G}_{n-1}) = Z_{n-1} - Z_{n-1}^2 \to Z(1 - Z) \quad \text{a.s.}$$

Thus, Lemma 2 implies that

$$\frac{1}{n}\sum_{k=1}^{n}(X_k - Z_{k-1})^2 + \frac{2}{n}\sum_{k=1}^{n}k(X_k - Z_{k-1})(Z_{k-1} - Z_k) \to -Z(1 - Z) \quad \text{a.s.}$$

Finally, write

$$\frac{1}{n}\sum_{k=1}^{n}k^2(Z_{k-1} - Z_k)^2 = \frac{1}{n}\sum_{k=1}^{n}k^2\bigg\{\frac{(1 - Z_{k-1})^2 X_k B_k^2}{(S_{k-1} + B_k)^2} + \frac{Z_{k-1}^2(1 - X_k)R_k^2}{(S_{k-1} + R_k)^2}\bigg\}.$$

By Lemma 2 and the same truncation technique as used in part 1, $(1/n)\sum_{k=1}^{n}k^2(Z_{k-1} - Z_k)^2 \to V$ a.s. Squaring,

$$\frac{1}{n}\sum_{k=1}^{n}\{X_k - Z_{k-1} + k(Z_{k-1} - Z_k)\}^2 \to V - Z(1 - Z) = U \quad \text{a.s.,}$$

that is, condition (b) holds. This concludes the proof.

References

[1] Aldous, D. J. and Eagleson, G. K. (1978). On mixing and stability of limit theorems. *Ann. Prob.* **6,** 325–331.

[2] Aletti, G., May, C. and Secchi, P. (2009). A central limit theorem, and related results, for a two-color randomly reinforced urn. *Adv. Appl. Prob.* **41,** 829–844.

[3] Bai, Z. D. and Hu, F. (2005). Asymptotics in randomized URN models. *Ann. Appl. Prob.* **15,** 914–940.

[4] Bassetti, F., Crimaldi, I. and Leisen, F. (2010). Conditionally identically distributed species sampling sequences. *Adv. Appl. Prob.* **42,** 433–459.

[5] Berti, P., Pratelli, L. and Rigo, P. (2004). Limit theorems for a class of identically distributed random variables. *Ann. Prob.* **32,** 2029–2052.

[6] Berti, P., Crimaldi, I., Pratelli, L. and Rigo, P. (2009). Rate of convergence of predictive distributions for dependent data. *Bernoulli* **15,** 1351–1367.

[7] Crimaldi, I. (2009). An almost sure conditional convergence result and an application to a generalized Pólya urn. *Internat. Math. Forum* **4,** 1139–1156.

[8] Crimaldi, I., Letta, G. and Pratelli, L. (2007). A strong form of stable convergence. In *Séminaire de Probabilités XL* (Lecture Notes Math. **1899**), Springer, Berlin, pp. 203–225.

[9] DURHAM, S. D., FLOURNOY, N. AND LI, W. (1998). A sequential design for maximizing the probability of a favourable response. *Canad. J. Statist.* **26,** 479–495.
[10] GHOSH, J. K. AND RAMAMOORTHI, R. V. (2003). *Bayesian Nonparametrics*. Springer, New York.
[11] HALL, P. AND HEYDE, C. C. (1980). *Martingale Limit Theory and Its Applications*. Academic Press, New York.
[12] JANSON, S. (2004). Functional limit theorems for multitype branching processes and generalized Pólya urns. *Stoch. Process. Appl.* **110,** 177–245.
[13] JANSON, S. (2005). Limit theorems for triangular urn schemes. *Prob. Theory Relat. Fields* **134,** 417–452.
[14] LEDOUX, M. AND TALAGRAND, M. (1991). *Probability in Banach Spaces*. Springer, Berlin.
[15] MAY, C. AND FLOURNOY, N. (2009). Asymptotics in response-adaptive designs generated by a two-color, randomly reinforced urn. *Ann. Statist.* **37,** 1058–1078.
[16] PEMANTLE, R. (2007). A survey of random processes with reinforcement. *Prob. Surveys* **4,** 1–79.
[17] PITMAN, J. AND YOR, M. (1997). The two-parameter Poisson–Dirichlet distribution derived from a stable subordinator. *Ann. Prob.* **25,** 855–900.
[18] RÉNYI A. (1963). On stable sequences of events. *Sankhyā* A **25,** 293–302.