Statistica Sinica **17**(2007), 387-409

A GENERALIZED DROP-THE-LOSER URN FOR CLINICAL TRIALS WITH DELAYED RESPONSES

Li-Xin Zhang[1], Wai Sum Chan[2], Siu Hung Cheung[2] and Feifang Hu[3]

[1]*Zhejiang University,* [2]*Chinese University of Hong Kong and* [3]*University of Virginia*

Abstract: Urn models are popular and useful for adaptive designs in clinical studies. Among various urn models, the drop-the-loser rule is an efficient adaptive treatment allocation scheme, recently proposed for comparing different treatments in a clinical trial. This rule is superior to other randomization schemes in terms of variability and power. In this paper, the drop-the-loser rule is generalized to cope with more popular and practical circumstances, including (1) delayed responses when test results cannot be obtained immediately, (2) continuous responses, and (3) a pre-specified target of allocation proportion. In addition, our proposed procedure has several favorable asymptotic properties such as strong consistency and asymptotic normality of the allocation proportions.

Key words and phrases: Asymptotic normality, asymptotic power, clinical trial, delayed response, randomized play-the-winner rule, strong consistency.

1. Introduction

In clinical trials, patients usually accrue sequentially. One of the fundamental concerns treatment allocation. Which treatment should be assigned to the next patient? The general consensus is that a randomization scheme should be adopted to minimize selection bias and to provide a solid basis for statistical inference. Adaptive designs can be valuable and ethical randomization schemes that formulate treatment allocation as a function of previous responses. One major objective of research in adaptive design is to develop treatment allocation schemes, so that more patients receive the better treatment.

Pioneering works in the area of adaptive design can be traced to Thompson (1933) and Robbins (1952). Since then, an unremitting generation of research products in this area offers various approaches to treatment allocation schemes applicable to clinical studies. For a discussion of recent developments in this area, refer to Rosenberger (1996), Rosenberger and Lachin (2002), and references therein.

Among different classes of adaptive designs, the one based on urn models receives the most attention. Early works include Athreya and Karlin (1968),

Wei and Durham (1978) and Wei (1979). The basic idea is as follows: there are various types of balls representing particular treatments; patients accrue sequentially; at each stage, the probability of allocating a particular treatment to a patient depends on the numbers of various types of balls in the urn. The response of each patient after treatment plays an essential role in the determination of subsequent urn compositions. The basic strategy is to "reward" more balls to successful treatments. The multi-treatment randomized play-the-winner rule (Andersen, Faries and Tamura (1994)) is an illustrative example. An urn contains K different types of balls, representing K different treatments. When a patient arrives, a ball is drawn at random with replacement. If it is a type i ball, the patient receives treatment i. A successful response to the treatment brings an addition of a type i ball to the urn. If the response is a failure, a ball is added to the urn. This ball is partitioned according to the existing proportion of balls for other treatments in the urn.

A sophisticated formulation of the urn model was given by Durham, Flournoy and Li (1998). They derived a valuable randomized version of the generalized Pólya urn that does not satisfy the regularity conditions of those studied by Athreya and Ney (1972). One major feature of the randomized Pólya urn scheme is to reward only successful treatments, balls are not added to the urn if the treatment is a failure. Parallel ideas can be useful in other areas besides clinical applications. For example, Beggs (2005) and Hopkins and Posch (2005) use related urn concepts to model reinforcement learning in their study of economic behaviors.

The importance of the randomized Pólya urn scheme is that it can be embedded in the family of continuous-time pure birth processes with linear birth rate (Yule processes). This enables the formulation of important limiting behaviors of the urn process (Ivanova and Flournoy (2001)). With the framework of embedding the urn scheme in a continuous-time birth and death process (Ivanova, Rosenberger, Durham and Flournoy (2000), Ivanova and Flournoy (2001), Ivanova (2003)) constructed the drop-the-loser (DL) urn.

The DL rule differs from the randomized Pólya urn of Durham, Flournoy and Li (1998). Instead of adding balls to reward successes, balls are removed when failures are observed. In the urn, besides treatment balls, there are immigration balls. When an immigration ball is selected, balls will be added to all types (except immigration), preventing extinction of types of treatment balls. The mechanism, and other properties of the DL rule, will be outlined in Section 2.

The DL rule was reported to have small variability and high statistical power (Ivanova (2003)). One sensible objective of clinical studies is to increase the power of treatment comparisons. Power depends heavily on the variability of

the treatment allocation scheme. Simulation evidence indicating the strong association between power and variability can be found in Melfi and Page (1998), as well as in Rosenberger, Stallard, Ivanova, Harper and Ricks (2001). A proof in Hu and Rosenberger (2003) confirmed that average power of a randomization procedure is a decreasing function of the variability of the randomization procedure. Therefore, adaptive designs with smaller variability are much preferred.

Recently Hu and Rosenberger (2003) launched a comparative study of several recent adaptive randomization procedures for binary responses: the sequential maximum likelihood procedure (SMLP) (Melfi and Page (2000)), the doubly adaptive biased coin design (DBCD) (Eisele (1994)), the generalized DBCD (Hu and Zhang (2004a)), the randomized play-the-winner (RPW) rule (Wei and Durham (1978)), and the drop-the-loser (DL) rule (Ivanova (2003)). Their study yielded results favoring the adoption of the DL rule due to its variability. For details, one can refer to Hu and Rosenberger (2003) and Hu, Rosenberger and Zhang (2006).

The DL rule has been shown to yield satisfactory results in terms of reducing the number of failures and variability (Rosenberger and Hu (2004)). Nevertheless, it has limitations. First, there is a lack of clear methodology to cope with delayed test responses which are common in clinical studies. Second, the application of the rule is limited to clinical trials with binary responses. Third, it can only be applied to target one particular allocation proportion (Ivanova (2003)) while different targets might be of interest in clinical studies (Rosenberger and Lachin (2002)). In fact, there is a growing interest in target-based designs which are derived with a pre-specified allocation target (see for example Eisele (1994), Eisele and Woodroofe (1995), Melfi and Page (1998, 2000)).

We derive a generalized DL (GDL) rule that differs from other popular urn models in its capability to handle delayed responses and to include pre-specified targets. In Section 2, the DL rule and its major properties will be outlined. Then the GDL rule is defined. Simulation results indicate that with delayed responses, our proposed scheme performs reasonably well. In Section 3, asymptotic properties and variability comparisons are presented. Some general comments and remarks are given in Section 4. When the responses are dichotomous, the GDL rule is shown to be asymptotically most powerful. Proofs are given in the last section. The main technique used in this paper involves the strong approximation of a martingale, and is different from the techniques employed in Ivanova (2003) and Ivanova et al. (2000). Furthermore, we show that the allocation process can be approximated by a standard Wiener process. The asymptotic normality, the rate of convergence and a law of the iterated logarithm are directly obtained from this approximation.

2. The Generalized Drop-the-Loser Rule

In this section we first describe the drop-the-loser rule (Ivanova (2003)) and its major statistical properties. Then our proposed generalized drop-the-loser rule will be introduced.

2.1. Drop-the-loser rule

For explanatory purpose, assume that we have two treatments even though the DL rule can be applied to multi-treatments. The DL rule is as follows.

Consider an urn containing three types of balls. Balls of types 1 and 2 represent treatments. Balls of type 0 are termed immigration balls. We start with $Z_{0,i}$ balls of type i, $i = 0, 1, 2$. Let $\mathbf{Z}_0 = (Z_{0,0}, Z_{0,1}, Z_{0,2})$ be the initial urn composition. After m draws, the urn composition becomes $\mathbf{Z}_m = (Z_{m,0}, Z_{m,1}, Z_{m,2})$. When a subject arrives, one ball is drawn at random. If a treatment ball of type k (1 or 2) is selected, the kth treatment is given to the subject and the response is observed. If it is a failure, the ball is not replaced, $Z_{m+1,k} = Z_{m,k} - 1$, $Z_{m+1,j} = Z_{m,j}$, $j \neq k$. If the treatment is a success, the ball is replaced and consequently, the urn composition remains unchanged, $\mathbf{Z}_{m+1} = \mathbf{Z}_m$. If an immigration ball (type 0) is selected, no subject is treated, and the ball is returned to the urn together with two additional treatment balls, one of each treatment type. Therefore, $Z_{m+1,0} = Z_{m,0}$ and $Z_{m+1,k} = Z_{m,k} + 1$, $k = 1, 2$. This procedure is repeated until a treatment ball is drawn and the subject treated accordingly. The function of the immigration ball is to avoid the extinction of a type of treatment ball.

Let P_k be the probability of success on treatment k, and $Q_k = 1 - P_k$, $k = 1, 2$. Ivanova (2003) studied the properties of the DL rule by embedding the urn composition process $\mathbf{Z}_m$ in an immigration-death process. She defined a two-dimensional process $\mathbf{Z}^*(t) = (Z_1^*(t), Z_2^*(t))$, which is a collection of two continuous-time linear immigration-death processes having common immigration processes with immigration rate $Z_{0,0}$ and independent death processes with death rates Q_1, Q_2, such that $Z_{m,k} = Z_k^*(t_m)$, $k = 1, 2$. Here t_m is the "time" of the mth draw and it is the partial sum of a sequence of independent exponentially distributed random variables with rate parameter 1. Note that t represents a "virtual" time instead of the real time. The embedding technique was developed by Athreya and Karlin (1968) and Athreya and Ney (1972) for the study of the Pólya urn model. Later it was adopted by Durham, Flournoy and Li (1998), Ivanova et al. (2000), Ivanova and Flournoy (2001) for studying sequential clinical trials.

Now, let us state a couple of important asymptotic results of the DL rule. Let $N_k(t)$ be the number of trials on treatment k up to time t, $k = 1, 2$. Ivanova

(2003) showed that

$$\frac{N_1(t)}{N_1(t)+N_2(t)} \xrightarrow{P} v_1 := \frac{\frac{1}{Q_1}}{\frac{1}{Q_1}+\frac{1}{Q_2}} \quad \text{as } t \to \infty, \tag{2.1}$$

$$\sqrt{N_1(t)+N_2(t)}\left(\frac{N_1(t)}{N_1(t)+N_2(t)} - v_1\right) \xrightarrow{D} N(0, \sigma^2_{DL}) \ \text{ as } t \to \infty, \ k = 1, 2, \tag{2.2}$$

where

$$\sigma^2_{DL} = \frac{Q_1 Q_2 (P_1 + P_2)}{(Q_1 + Q_2)^3}, \tag{2.3}$$

is the asymptotic variance. The DL rule has two fundamental properties: (1) it preserves the randomization ingredient of the randomized play-the-winner rule, which yields a non-deterministic scheme; (2) when compared with many other adaptive designs which have the same limit proportions of the DL rule, such as the SMLP, the DBCD and the RPW rules, the DL rule generates an allocation procedure with the minimum asymptotic variance, hence produces higher power for the test of the difference of proportions (Hu and Rosenberger (2003)).

In practice, subjects frequently do not respond immediately. Therefore, the response of an individual may not be available prior to the randomization of the next subject. Delayed response is a scenario in clinical trials that deserves much attention. Besides delayed response, the DL rule is incapable of dealing with non-dichotomous responses. Basically, when the outcomes are delayed or non-dichotomous, it is difficult to embed the sequence of urn compositions in an immigration-birth-death process.

2.2. Generalized drop-the-loser rule

In this section, the GDL is outlined. The treatment allocation scheme is more flexible than the DL rule and accommodates the possibility of delayed responses and pre-assigned allocation proportion targets.

Similar to the DL rule, there are three types of balls in the urn. Balls of types 1 and 2 represent treatments, balls of type 0 are immigration balls. We start with $Z_{0,i}$ (> 0) balls of type i, $i = 0, 1, 2$. Let $\mathbf{Z}_0 = (Z_{0,0}, Z_{0,1}, Z_{0,2})$ be the initial urn composition, and $\mathbf{Z}_m = (Z_{m,0}, Z_{m,1}, Z_{m,2})$ be the urn composition after m draws. Let $Z^+_{m,i} = \max(0, Z_{m,i})$, $i = 0, 1, 2$, and $\mathbf{Z}^+_m = (Z^+_{m,0}, Z^+_{m,1}, Z^+_{m,2})$. When a subject arrives to be allocated to a treatment, a ball is drawn at random according to the urn composition $\mathbf{Z}^+_m$ for the appropriate m. That is, the probability of selecting type i ball is $Z^+_{m,i}/|\mathbf{Z}^+_m|$, with $|\mathbf{Z}^+_m| = Z^+_{m,0} + Z^+_{m,1} + Z^+_{m,2}$.

If an immigration ball (type 0) is drawn, no treatment is assigned and the ball is returned to the urn along with a_k type k treatment balls, $k = 1, 2$. Let $A = a_1 + a_2$ where $a_1, a_2 > 0$. This step is repeated until a treatment ball is drawn.

If a type k ($k = 1, 2$) treatment ball is drawn, the subject is assigned to treatment k and the ball is not replaced immediately. To allow for delayed responses, the addition of balls is made after the subject's response is observed. We denote the outcome of this subject on treatment k by $Y_{m,k}$. The outcome $Y_{m,k}$ may not be available prior to the arrival of the next subject. In fact, the delayed outcome may only be available after several subjects (a random variable) have been allocated to treatments. After the response $Y_{m,k}$ is observed, $D_{m,k}$ (≥ 0) balls of type k are added to the urn.

We allow the urn to have a fractional or negative number of treatment balls. According to the definition of $Z^+_{m,i}$, the treatment balls with negative numbers will never be selected. As a result, the number of treatment balls of each type will not decrease when it is negative. So $Z_{i,m} \geq -1$ for all m and i.

Let $N_{n,k}$ be the number of subjects assigned to treatment k, $k = 1, 2$, after the allocation of treatments to n subjects. It is important to study the statistical behavior of the proportions of patients $N_{n,k}/n$, $k = 1, 2$, assigned to the two treatments.

Let $p_k = \mathsf{E}[D_{m,k}]$, $k = 1, 2$. We assume $0 \leq p_k < 1$ and $q_k = 1 - p_k$, $k = 1, 2$. Thus, after each treated subject, the expected number of balls added according to the outcome observed is not larger than the number of outgoing balls (which is 1).

The DL rule is a particular case of our GDL allocation scheme. For instance, with dichotomous responses and two treatments, the DL rule corresponds to the GDL rule with $a_1 = 1$, $a_2 = 1$ and $D_{m,k} = 1$ if the outcome of treatment k is a success, and 0 otherwise. In addition, $p_k = P_k$, the success probability of a trial on treatment k, $k = 1, 2$.

When the outcomes are not dichotomous, one may choose suitable adding rules $\{D_{m,k}\}$ to define a design. For example, the outcome of a patient after treatment of cancer can be classified as "clinically ineffective", "gradual improvement with extended treatment" or "fully recovered"; one may define $D_{m,k} = 1$ if the outcome is a "fully recovered", $D_{m,k} = \Lambda$ ($0 < \Lambda < 1$) if the outcome is "gradual improvement with extended treatment", and $D_{m,k} = 0$ if the outcome is "clinically ineffective".

Under some suitable conditions (stated in Section 3), we can show that the proportion of subjects assigned to treatment k is

$$\frac{N_{n,k}}{n} \to v_k := \frac{\frac{a_k}{q_k}}{\frac{a_1}{q_1} + \frac{a_2}{q_2}} \quad a.s. \quad k = 1, 2. \tag{2.4}$$

With dichotomous outcomes, if $a_1 = a_2$, $D_{m,k} = 1$ for success and $D_{m,k} = 0$ for failure when type k treatment is assigned, the limiting proportions v_k, $k = 1, 2$, are the same as in (2.1). One can choose a_k's to adjust the allocation proportions. By choosing a_k's suitably, the GDL rule can be used to target any desired allocation.

A more convenient approach to target a pre-specified allocation proportion is to take $D_{m,k} \equiv 0$ for all m and k. Hence, $q_k = 1 - \mathsf{E}D_{m,k} \equiv 1$. If the target allocation proportions is v_k $(k = 1, 2)$, we can simply define a design by choosing $a_k = Cv_k$ where C is a constant and v_k is a function of P_k. For example, Rosenberger et al. (2001) studied the allocation proportions

$$\frac{\sqrt{P_k}}{\sqrt{P_1} + \sqrt{P_2}}, \quad k = 1, 2, \tag{2.5}$$

which minimize the expected number of failures under fixed variance of the estimator of the treatment difference. In this case, we take

$$a_k = C\frac{\sqrt{P_k}}{\sqrt{P_1} + \sqrt{P_2}}, \quad k = 1, 2 \tag{2.6}$$

and the balls are added only through immigration. The superior treatment (the one with larger probability of success) will be rewarded more balls each time an immigration ball is selected. Simulation study in the following section indicates that there is no significant difference among various choices of C.

Remark 2.1. In practice, the P_k are usually unknown. In these cases, simply substitute $\widehat{P}_k$ for P_k, where $\widehat{P}_k$ is the current estimate of P_k, $k = 1, 2$. We propose the estimate

$$\widehat{P}_k = \frac{(\text{number of observed successes on treatment } k) + 1}{(\text{number of observed outcomes on treatment } k) + 2},$$

which is the Bayesian estimate of P_k with a uniform prior distribution, $k = 1, 2$. Variously, one can replace 1 in the numerator by α and 2 in the denominator by $\alpha + \beta$ if the beta distribution $beta(\alpha, \beta)$ is employed as the prior distribution, with the constants α and β estimated from earlier trials.

2.3. Simulation results

In this section, a simulation study is performed to investigate the performance of our allocation scheme. Two different allocation targets, (2.1) and (2.5), are employed as our study cases. Given treatments 1 and 2 with success probabilities P_1 and P_2 respectively, our simulation study is performed with P_1 and P_2 being selected with reference to those choices of Hu and Rosenberger (2003). For the allocation process, $\widehat{P}_k$ given in Remark 2.1 is utilized.

For both the delayed times for the two treatments and the patient entry times, exponential distributions are used. The mean parameters of the delay times for treatments 1 and 2 are λ_1 and λ_2 respectively. For patient entry times, the mean parameter is λ_3. There are three different configurations for the mean parameters. The first one corresponds to the case where there are no delayed responses. The second one corresponds to $(\lambda_1, \lambda_2, \lambda_3) = (1, 1, 1)$, which represents similar delayed times for the responses of the two treatments. Finally, we select $(\lambda_1, \lambda_2, \lambda_3)$ to be $(5, 1, 1)$ to represent a large difference in delayed times for the responses of the two treatments. As explained earlier, for simplicity we pick $D_{m,k} = 0$ for the GDL rules.

The number of subjects n is chosen to be 100 and 500. The number of replications in our simulation study is 10,000. The proportions of subjects being allocated to treatment 1, $N_{n,1}/n$, are tabulated in Tables 1 and 2, since $N_{n,2}/n$ is simply $1 - N_{n,1}/n$.

Table 1. Simulated allocation proportion ($N_{n,1}/n$) of DL rule, GDL rule and DBCD with allocation target v_1 given in (2.1).

		DL		GDL (1)		DBCD	
p_1, p_2	v_1	$n = 100$	$n = 500$	$n = 100$	$n = 500$	$n = 100$	$n = 500$
Immediate Response							
0.8, 0.8	0.50	0.50(0.069)	0.50(0.041)	0.50(0.102)	0.50(0.058)	0.50(0.103)	0.50(0.049)
0.8, 0.6	0.67	0.62(0.060)	0.66(0.031)	0.63(0.079)	0.66(0.042)	0.65(0.076)	0.66(0.037)
0.7, 0.5	0.63	0.60(0.053)	0.62(0.026)	0.60(0.067)	0.62(0.035)	0.62(0.065)	0.62(0.030)
0.5, 0.5	0.50	0.50(0.047)	0.50(0.022)	0.50(0.057)	0.50(0.029)	0.50(0.056)	0.50(0.026)
0.5, 0.2	0.62	0.61(0.035)	0.61(0.016)	0.60(0.042)	0.61(0.021)	0.61(0.043)	0.61(0.020)
0.2, 0.2	0.50	0.50(0.025)	0.50(0.011)	0.50(0.030)	0.50(0.015)	0.50(0.034)	0.50(0.016)
$(\lambda_1, \lambda_2, \lambda_3) = (1, 1, 1)$							
0.8, 0.8	0.50	0.50(0.066)	0.50(0.041)	0.50(0.099)	0.50(0.057)	0.50(0.103)	0.50(0.049)
0.8, 0.6	0.67	0.62(0.058)	0.66(0.031)	0.63(0.078)	0.66(0.042)	0.65(0.076)	0.66(0.037)
0.7, 0.5	0.63	0.60(0.052)	0.62(0.026)	0.60(0.066)	0.62(0.034)	0.62(0.065)	0.62(0.030)
0.5, 0.5	0.50	0.50(0.046)	0.50(0.022)	0.50(0.058)	0.50(0.029)	0.50(0.056)	0.50(0.026)
0.5, 0.2	0.62	0.61(0.035)	0.61(0.016)	0.60(0.041)	0.61(0.021)	0.61(0.043)	0.61(0.020)
0.2, 0.2	0.50	0.50(0.025)	0.50(0.011)	0.50(0.030)	0.50(0.015)	0.50(0.034)	0.50(0.016)
$(\lambda_1, \lambda_2, \lambda_3) = (5, 1, 1)$							
0.8, 0.8	0.50	0.47(0.060)	0.49(0.040)	0.49(0.099)	0.50(0.057)	0.50(0.104)	0.50(0.049)
0.8, 0.6	0.67	0.59(0.055)	0.65(0.030)	0.63(0.077)	0.66(0.042)	0.65(0.078)	0.66(0.037)
0.7, 0.5	0.63	0.58(0.049)	0.62(0.026)	0.60(0.066)	0.62(0.035)	0.61(0.066)	0.63(0.030)
0.5, 0.5	0.50	0.50(0.045)	0.50(0.022)	0.50(0.056)	0.50(0.029)	0.50(0.057)	0.50(0.026)
0.5, 0.2	0.62	0.60(0.033)	0.61(0.016)	0.60(0.042)	0.61(0.021)	0.61(0.044)	0.61(0.020)
0.2, 0.2	0.50	0.50(0.025)	0.50(0.011)	0.50(0.030)	0.50(0.015)	0.50(0.035)	0.50(0.016)

Simulated standard deviations are given in parentheses.
GDL(1):$a_1 = 2v_1, \quad a_2 = 2(1 - v_1)$

Table 2. Simulated allocation proportion ($N_{n,1}/n$) of two GDL rules and DBCD with allocation target v_1 given in (2.5)

		GDL (2)		GDL (3)		DBCD	
p_1, p_2	v_1	$n = 100$	$n = 500$	$n = 100$	$n = 500$	$n = 100$	$n = 500.$
Immediate Response							
0.8, 0.8	0.50	0.50(0.019)	0.50(0.008)	0.50(0.018)	0.50(0.008)	0.50(0.027)	0.50(0.012)
0.8, 0.6	0.54	0.53(0.023)	0.54(0.011)	0.53(0.023)	0.54(0.011)	0.54(0.030)	0.54(0.013)
0.7, 0.5	0.54	0.54(0.028)	0.54(0.013)	0.54(0.028)	0.54(0.013)	0.54(0.033)	0.54(0.014)
0.5, 0.5	0.50	0.50(0.032)	0.50(0.015)	0.50(0.032)	0.50(0.016)	0.50(0.036)	0.50(0.017)
0.5, 0.2	0.61	0.59(0.042)	0.61(0.024)	0.59(0.042)	0.61(0.024)	0.61(0.049)	0.61(0.022)
0.2, 0.2	0.50	0.50(0.051)	0.50(0.029)	0.50(0.051)	0.50(0.029)	0.50(0.058)	0.50(0.026)
$(\lambda_1, \lambda_2, \lambda_3) = (1, 1, 1)$							
0.8, 0.8	0.50	0.50(0.018)	0.50(0.008)	0.50(0.018)	0.50(0.008)	0.50(0.027)	0.50(0.012)
0.8, 0.6	0.54	0.53(0.023)	0.54(0.011)	0.53(0.022)	0.54(0.011)	0.54(0.030)	0.54(0.013)
0.7, 0.5	0.54	0.54(0.028)	0.54(0.013)	0.54(0.027)	0.54(0.013)	0.54(0.033)	0.54(0.015)
0.5, 0.5	0.50	0.50(0.032)	0.50(0.015)	0.50(0.032)	0.50(0.015)	0.50(0.036)	0.50(0.016)
0.5, 0.2	0.61	0.59(0.042)	0.61(0.024)	0.59(0.042)	0.61(0.024)	0.61(0.049)	0.61(0.022)
0.2, 0.2	0.50	0.50(0.051)	0.50(0.029)	0.50(0.051)	0.50(0.029)	0.50(0.058)	0.50(0.026)
$(\lambda_1, \lambda_2, \lambda_3) = (5, 1, 1)$							
0.8, 0.8	0.50	0.50(0.019)	0.50(0.008)	0.50(0.017)	0.50(0.008)	0.50(0.027)	0.50(0.012)
0.8, 0.6	0.54	0.53(0.023)	0.53(0.011)	0.53(0.023)	0.53(0.011)	0.54(0.030)	0.54(0.013)
0.7, 0.5	0.54	0.54(0.028)	0.54(0.013)	0.54(0.028)	0.54(0.013)	0.54(0.033)	0.54(0.015)
0.5, 0.5	0.50	0.50(0.031)	0.50(0.015)	0.50(0.032)	0.50(0.015)	0.50(0.037)	0.50(0.016)
0.5, 0.2	0.61	0.59(0.042)	0.61(0.024)	0.59(0.041)	0.61(0.024)	0.61(0.049)	0.61(0.022)
0.2, 0.2	0.50	0.50(0.050)	0.50(0.029)	0.50(0.052)	0.50(0.029)	0.50(0.058)	0.50(0.026)

Simulated standard deviations are given in parentheses.
GDL (2): $a_1 = 2v_1, \quad a_2 = 2(1 - v_1)$
GDL (3): $a_1 = 2(\sqrt{p_1} + \sqrt{p_2})(v_1) = 2\sqrt{p_1}, \quad a_2 = 2(\sqrt{p_1} + \sqrt{p_2})(1 - v_1) = 2\sqrt{p_2}$

For comparison purposes, the DBCD is also included. The allocation scheme used in this simulation study follows that of Rosenberger and Hu (2004) closely. In addition their suggested value of 2, for the parameter that determines the variability of the allocation proportions arising from the randomized procedure, is adopted.

For Table 1, the allocation target given in (2.1) is used. Even though the DL rule was not designed for delayed responses, for exploratory purposes it is included in the cases with delayed responses. A simplistic approach is adopted. When a treatment ball is chosen, the action of whether to return the ball or not is deferred until the response is observed. We have the following findings. For large sample sizes ($n = 500$) and/or without delayed responses, both the DL rule and the GDL rules are able to provide allocation proportions very close to the target. For smaller sample sizes ($n = 100$) and delayed responses, the

DL rule is outperformed by the GDL rule and the DBCD, especially when both treatments have high success rates (example: $P_1 = 0.8$, $P_2 = 0.6$). Note that the variances of the GDL rule and the DBCD are slightly larger due to the requirement of estimating P_1 and P_2 at each stage when an immigration ball is selected. In return, these estimates provide precise estimates of the efficacies of the treatments, especially when delayed responses are present. This also explains why the GDL rule and the DBCD surpass the DL rule in terms of the convergence of the allocation proportions in such cases.

In Table 2, the optimal allocation target in (2.5) is used. All allocation proportions are quite close to the pre-specified target. In addition, the two choices of C for the immigration rates, a_1 and a_2, which represent the addition of roughly two treatment balls when an immigration ball is selected, do not yield much differences in terms of the allocation proportions. In fact several other possible values of C were tried and, as long as the number of balls added to the urn remained less than 4, similar results were obtained and hence not reported. The immigration ball has two important functions: the first is to prevent the possibility of extinction of a particular type of treatment ball; the second is to add treatment balls to the urn according to the current estimates of P_1 and P_2. Therefore, to allow the immigration ball to play these two roles continuously during the allocation process, the principle is not to add so many treatment balls to the urn that the chance of selecting an immigration ball becomes too small.

Simulation results in Table 2 also reveal that the DBCD's performance is comparable to the GDL rule. The DBCD has an infinitesimal advantage in accuracy in attaining the target allocation, but has slightly larger variances for $n = 100$. However a complete theoretical justification of DBCD with delayed responses, similar to the one provided for the GDL rule in this paper, is still unavailable.

Finally, the use of the Bayesian estimates of P_k $(k = 1, 2)$ works very well. We have also computed the final estimates of the success probabilities, and these are always close to the actual values.

3. Asymptotic Properties of the GDL Rule

In this section several useful asymptotic properties for the GDL rule are given. We consider only the case in which the numbers of the immigrated balls a_k, $k = 1, 2$, are fixed. More complicated scenarios in which a_ks vary from time to time are an interesting topic for future study.

Now, let t_m be the entry time of the mth subject. Assume that $\{t_{m+1} - t_m; m \geq 1\}$ is a sequence of independent and identically distributed random variables. The response time of the mth subject with treatment k is denoted by

$r_m(k)$. Suppose $\{r_m(k); m \geq 1\}$ are sequences of independent random variables, $k = 1, 2$. Further, let the response times be independent of the entry times. We also assume that the draw, removal and addition of balls requires no time, and so the mth subject is randomized at time t_m. For the response time $r_m(k)$, we have the following assumption.

Assumption 3.1. Let $\delta_k(m, n) = I\{r_m(k) > t_{m+n} - t_m\}$ be an indicator function that takes the value 1 if the outcome of the mth subject on treatment k occurs after at least another n subjects are randomized, and 0 otherwise. Suppose for some constants $C > 0$ and $\gamma > 2$, $\mu_k(m, n) = \mathsf{P}\{\delta_k(m, n) = 1\} \leq Cn^{-\gamma}$, $m, n = 1, 2, \ldots$, $k = 1, 2$.

Since the above probability is a decreasing function of n with a power rate, the chance is slim that too many patients arrive before a delayed response is observed.

Remark 3.1. For generalized Friedman's urn models (also known as generalized Pólya urn models) with delayed responses, the assumptions of delayed time have been discussed by Bai, Hu and Rosenberger (2002) and Hu and Zhang (2004b)). Similar to the arguments in Bai, Hu and Rosenberger (2002), we can show that Assumption 3.1 is satisfied if (i) the γth moment of $r_m(k)$ exists and (ii) $E(t_{m+1} - t_m) > 0$ and $E|t_{m+1} - t_m|^{2\gamma} < \infty$. These two conditions can be easily verified in applications.

Assumption 3.2. $\{D_{m,k};\ m \geq 1\}$, $k = 1, 2$, are two sequences of i.i.d. random variables with $0 \leq p_k = \mathsf{E}[D_{m,k}] < 1$ and $\mathsf{E}|D_{m,k}|^p < \infty$ for any $p > 0$, $k = 1, 2$.

Let $\sigma_k^2 = \mathrm{Var}\,(D_{m,k})$ be the variance of the adding rules and $q_k = 1 - p_k$, $k = 1, 2$.

Theorem 3.1. *Suppose Assumptions* 3.1 *and* 3.2 *are satisfied. Let* v_k, $k = 1, 2$, *be defined in* (2.4). *Then there exists a standard Brownian motion* $\{W(t);\ t \geq 0\}$ *such that for any* $\delta > 0$, $N_{n,1} - nv_1 = \sigma W(n) + o(n^{(\gamma+1)/(3\gamma)+\delta})$ *a.s., and* $N_{n,2} - nv_2 = -\sigma W(n) + o(n^{(\gamma+1)/(3\gamma)+\delta})$ *a.s., where*

$$\sigma^2 = \frac{a_1 a_2 (a_2 q_2 \sigma_1^2 + a_1 q_1 \sigma_2^2)}{(a_2 q_1 + a_1 q_2)^3}. \tag{3.1}$$

The proof of the theorem will be given in the last section. By the properties of Wiener processes, the following is an immediate corollary of the theorem.

Corollary 3.1. *Under Assumptions* 3.1 *and* 3.2,

$$\frac{N_{n,k}}{n} - v_k = O\left(\sqrt{\frac{\log\log n}{n}}\right) \quad a.s., \quad k = 1, 2, \tag{3.2}$$

$$\sqrt{n}\Big(\frac{N_{n,k}}{n} - v_k\Big) \xrightarrow{D} N(0, \sigma^2), \quad k = 1, 2, \tag{3.3}$$

where v_k, $k = 1, 2$, *are defined in* (2.4), *and* σ^2 *is defined in* (3.1).

Equation (3.2) gives strong consistency with its rate of convergence for the proportions $N_{n,k}/n$, $k = 1, 2$. Comparing with (2.2), where the result is given through a "virtual" time, (3.3) provides the direct asymptotic distributions of the proportions. The asymptotic distributions and the asymptotic variance can be used to compare with other adaptive designs (Hu and Rosenberger (2003)).

Remark 3.2. From Corollary 3.1, the asymptotic properties of the GDL process does not depend on the delayed mechanism as long as Assumption 3.1 is satisfied. However, in Theorem 3.1, the convergence rate of the error depends on γ which is affected by the degree of the delayed responses.

Example 3.1. *Binary response:* Based on the result of Hu, Rosenberger and Zhang (2006), we can calculate the lower bound of the asymptotic variance for the allocation proportion v_1 given as in (2.4). For the case with dichotomous outcomes, $\sigma_k^2 = p_k q_k$, $k = 1, 2$. Here $p_k = P_k$, $q_k = 1 - P_k$, and P_k is the probability of a success on treatment k, $k = 1, 2$. Let $\mathbf{p} = (p_1, p_2)$ and

$$f(\mathbf{y}) = \frac{\frac{a_1}{1-y_1}}{\frac{a_1}{1-y_1} + \frac{a_2}{1-y_2}}.$$

According to Theorem 1 of Hu, Rosenberger and Zhang (2006), the lower bound of the asymptotic variance is

$$\sigma^2_{min}(\mathbf{p}) := \Big(\frac{\partial f}{\partial \mathbf{y}}\Big|_{\mathbf{p}}\Big)\mathbf{I}^{-1}(\mathbf{p})\Big(\frac{\partial f}{\partial \mathbf{y}}\Big|_{\mathbf{p}}\Big)',$$

where $\mathbf{I}(\mathbf{p}) = diag(\frac{v_1}{p_1 q_1}, \frac{v_2}{p_2 q_2})$ is Fisher's information matrix. Taking derivatives of f we find that

$$\frac{\partial f}{\partial \mathbf{y}}\Big|_{\mathbf{p}} = \Big(\frac{\partial f}{\partial y_1}\Big|_{\mathbf{p}}, \frac{\partial f}{\partial y_2}\Big|_{\mathbf{p}}\Big) = \Big(-\frac{v_1 v_2}{q_1}, \frac{v_1 v_2}{q_2}\Big).$$

It follows that $\sigma^2_{min}(\mathbf{p}) = \sigma^2$ by some elementary calculation, where σ^2 is defined in (3.1). Based on Corollary 3.1, the GDL rule attains this lower bound and hence it is asymptotically the most powerful design.

When $a_1 = a_2$, $D_{m,k} = 1$ for success and $D_{m,k} = 0$ for failure when type k treatment is assigned, the GDL rule becomes the DL rule. The asymptotic variance, $nVar(N_{n,1}/n)$, is

$$\sigma^2_{DL} = \frac{q_1 q_2 (p_1 + p_2)}{(q_1 + q_2)^3} = \frac{Q_1 Q_2 (P_1 + P_2)}{(Q_1 + Q_2)^3},$$

which is the smallest among all the adaptive designs considered in Hu and Rosenberger (2003).

4. Discussion

One important application of the GDL rule in clinical studies is that it can be used for continuous responses. For instance, we may apply the GDL rule to the example studied in Section 8 of Eisele and Woodroofe (1995), where the responses are normally distributed and the desired target proportion is the popular Neyman allocation. Similar to Remark 2.1, we can choose the a_k and $D_{m,k}$ sequentially to target the desired proportion. It would be interesting to compare this design with the doubly adaptive biased coin designs (Hu and Zhang (2004a)) in which the allocation probabilities are functions of sequential estimators of unknown parameters, and the sequential estimation-adjusted urn models (Zhang, Hu and Cheung (2006)). However, when the a_k depend on the process (as indicated in Remark 2.1), the asymptotic properties of the allocation proportion $N_{n,1}/n$ are unknown. This is an interesting future research topic.

For the randomized play-the-winner rule with delayed responses, Wei (1988) suggested updating the urn when responses become available. For a generalized Friedman's urn model (the randomized play-the-winner rule is a special case) with delayed responses, the limiting distribution of the urn composition was derived in Bai, Hu and Rosenberger (2002). Further, Hu and Zhang (2004b) obtained the limiting distribution of the allocation proportion. Both papers showed that the delayed responses do not affect the asymptotic properties of the generalized Friedman's urn model. Here we obtain similar results for the GDL rule. Nevertheless, the arguments are only valid in the context of large samples. In practice, the delayed mechanism is important and should not be ignored, as indicated by our simulation findings.

5. Proofs

Theorem 3.1 is proved in this section. Recall that $\mathbf{Z}_n = (Z_{n,0}, Z_{n,1}, Z_{n,2})$ represents the numbers of balls after n draws and $|\mathbf{Z}_n^+| = Z_{n,0}^+ + Z_{n,1}^+ + Z_{n,2}^+$. Because every immigration ball is replaced, $Z_{n,0}^+ = Z_{n,0} = Z_{0,0}$ for all n. Let $\mathbf{X}_n$ be the result of the nth draw, where $X_{n,k} = 1$ if the selected ball is of type k and $X_{n,k} = 0$ otherwise, $k = 0, 1, 2$. Further, let $\mathbf{N}_n^* = (N_{n,0}^*, N_{n,1}^*, N_{n,2}^*) = \sum_{m=1}^n \mathbf{X}_m$, so $N_{n,k}^*$ is the number of selected type k balls in the first n draws. Let $u_n = \max\{m : N_{m,1}^* + N_{m,2}^* \le n\}$. Then u_n is the total number of draws of treatment type balls in the first n assignments, and $N_{n,k} = N_{u_n,k}^*$, $k = 1, 2$.

Let $I_k(m, n)$ be the indicator function, which takes value 1 if the outcome $Y_{m,k}$ on treatment k of the subject assigned at the mth draw occurs after the $(m + n)$th draw and before the $(m + n + 1)$th draw, $k = 1, 2$. Remember that, when $Y_{m,k}$ occurs, we add $D_{m,k} = D(Y_{m,k})$ balls of type k into the urn. So, for given m and n, if $I_k(m, n) = 1$, we add $X_{m,k}D_{m,k}$ balls of type k to the urn.

Consequently, if $m = 0$, $I_k(n-m, m) = I_k(n, 0)$ and the outcome on treatment k assigned at time n occurs after the nth draw and before the $(n+1)$th draw; ...; if $m = n-1$, $I_k(n-m, m) = I_k(1, n-1)$ and the outcome on treatment k assigned at time 1 occurs after the nth draw and before the $(n+1)$th draw. Hence, after the nth draw and before the $(n+1)$th draw, the numbers of balls of each type added according to the outcomes are

$$\begin{aligned} W_{n,k} &= \sum_{m=0}^{n-1} I_k(n-m, m) X_{n-m,k} D_{n-m,k} \\ &= \sum_{m=1}^{n} I_k(m, n-m) X_{m,k} D_{m,k}, \quad k = 1, 2. \end{aligned}$$

The change in the number of type k balls after n draws from the time of the $(n-1)$th draw is $Z_{n,k} - Z_{n-1,k} = a_k X_{n,0} - X_{n,k} + W_{n,k}$, $k = 1, 2$. Recall that a_k here is the number of added type k balls when an immigration ball is drawn. So, for $k = 1, 2$, the number of type k balls added after n draws is

$$\begin{aligned} Z_{n,k} - Z_{0,k} &= a_k \sum_{j=1}^{n} X_{j,0} - \sum_{j=1}^{n} X_{j,k} + \sum_{j=1}^{n} W_{j,k} \\ &= a_k \sum_{m=1}^{n} X_{m,0} - \sum_{m=1}^{n} X_{m,k} + \sum_{m=1}^{n} \sum_{j=m}^{n} X_{m,k} D_{m,k} I_k(m, j-m) \\ &= a_k \sum_{m=1}^{n} X_{m,0} - \sum_{m=1}^{n} X_{m,k} + \sum_{m=1}^{n} \sum_{j=m}^{\infty} X_{m,k} D_{m,k} I_k(m, j-m) \\ &\quad - \sum_{m=1}^{n} \sum_{j=n+1}^{\infty} X_{m,k} D_{m,k} I_k(m, j-m) \\ &= a_k \sum_{m=1}^{n} X_{m,0} + \sum_{m=1}^{n} X_{m,k}(D_{m,k} - 1) - \sum_{m=1}^{n} \sum_{j=n+1}^{\infty} X_{m,k} D_{m,k} I_k(m, j-m) \\ &=: a_k \sum_{m=1}^{n} X_{m,0} + \sum_{m=1}^{n} X_{m,k}(D_{m,k} - 1) - R_{n,k}. \end{aligned} \tag{5.1}$$

That is

$$\Delta Z_{n,k} = a_k X_{n,0} + X_{n,k}(D_{n,k} - 1) - \Delta R_{n,k}, \quad k = 1, 2, \tag{5.2}$$

where Δ denotes the differencing operand of a sequence $\{z_n\}$. From (5.1), it

follows that

$$\begin{aligned} Z_{n,k} - Z_{0,k} &= a_k N^*_{n,0} - q_k N^*_{n,k} + \sum_{m=1}^{n} X_{m,k}(D_{m,k} - \mathsf{E}[D_{m,k}]) - R_{n,k} \\ &=: a_k N^*_{n,0} - q_k N^*_{n,k} + M_{n,k} - R_{n,k}, \quad k = 1, 2. \end{aligned} \tag{5.3}$$

We prove Theorem 3.1 by showing that $R_{n,k}$ and $Z_{n,k}$ can be neglected, and the major term $M_{n,k}$ can be approximated by a Wiener process. Notice that $Z_{n,k}$ is a function of $\{I_k(m,j)\}$. We show that $Z_{n,k}$ can be neglected by using the fact that $\mathsf{E}[I_k(n,j)]$ decays very rapidly. So we first replace Assumption 3.1 by the following one on $I_k(n,j)$.

Condition A. For some $\varphi > 1$, $\sum_{j=n}^{\infty} \mathsf{E}[I_k(m,j)] \le Cn^{-\varphi}$, for all n, m and $k = 1, 2$.

The summation in Condition A is the probability of the event that the subject who is assigned to treatment k at the mth draw responds after at least another n draws, and it is required that this probability decays with a power rate, similar to Assumption 3.1. The following claim provides the connection.

Claim. Assumption 3.1 implies Condition A with $\varphi = \gamma - 1 - \epsilon$ for any $\epsilon > 0$.

Proof. Let $N^*_m = N^*_{m,1} + N^*_{m,2}$. Notice that $\mathsf{E}[I_k(m,n)]$ is the probability of the event that the N^*_mth subject (who is assigned after the mth ball is drawn) on treatment k responds after the $(m+n)$th draw and before the $(m+n+1)$th draw. So $\mathsf{E}[I_k(m,n)|N^*_m = p] \le \mathsf{P}(E_1|N^*_m = p) + \mathsf{P}(E_2|N^*_m = p)$, where E_1 is the event that the pth subject responds after at least another $n^{1-\epsilon}$ subjects arrive, and E_2 is the event that there at least $n - n^{1-\epsilon}$ draws of type 0 balls from the mth draw to the $(m+n)$th draw. The event E_1 depends only on the response time of the pth subject and the waiting times for future subjects. However, the event $\{N^*_m = p\}$ depends only on past draws and assignments. So, E_1 and $\{N^*_m = p\}$ are independent. It follows that $\mathsf{P}(E_1|N^*_m = p) = \mathsf{P}(E_1) \le Cn^{-\gamma(1-\epsilon)}$ by Assumption 3.1. For $\mathsf{P}(E_2|N^*_m = p)$, we consider the event E_3 that the largest run of "1"s in $X_{m,0}, \ldots, X_{m+n,0}$ is at least n^ϵ. Notice that, for event E_3^C, there are at least n/n^ϵ zeros in $X_{m,0}, \ldots, X_{m+n,0}$, and then at most $n - n/n^\epsilon$ ones. So, E_2 does not occur. It follows that $\mathsf{P}(E_2|N^*_m = p) \le \mathsf{P}(E_3|N^*_m = p)$. Hence we conclude that

$$\begin{aligned} \mathsf{E}[I_k(m,n)] &\le Cn^{-\gamma(1-\epsilon)} + \mathsf{P}(E_3) \\ &\le Cn^{-\gamma(1-\epsilon)} + \sum_{i=m}^{m+n} \mathsf{P}\{X_{i,0} = \cdots = X_{i+[n^\epsilon],0} = 1\}. \end{aligned}$$

On the other hand,

$$\begin{aligned}
&\mathsf{P}\{X_{i,0} = \cdots = X_{i+[n^\epsilon],0} = 1\} \\
&\quad = \mathsf{E}\Big[I\{X_{i,0} = \cdots = X_{i+[n^\epsilon]-1,0} = 1\}\mathsf{P}[X_{i+[n^\epsilon],0} = 1|\mathcal{F}_{i+[n^\epsilon]-1}]\Big] \\
&\quad = \mathsf{E}\Big[I\{X_{i,0} = \cdots = X_{i+[n^\epsilon]-1,0} = 1\}\frac{Z_{0,0}}{|\mathbf{Z}^+_{i+[n^\epsilon]-1}|}\Big] \\
&\quad \le \mathsf{P}\{X_{i,0} = \cdots = X_{i+[n^\epsilon]-1,0} = 1\}\frac{Z_{0,0}}{Z_{0,0} + A([n^\epsilon]-1)},
\end{aligned}$$

since at each stage from stage i to $i + [n^\epsilon] - 1$ at least $A = a_1 + a_2$ balls are added to the run and no ball is removed, where $\mathcal{F}_n = \sigma(\mathbf{X}_1, \ldots, \mathbf{X}_n, \mathbf{Y}_1, \ldots, \mathbf{Y}_n)$ is the history sigma field. Here and in the remainder of this paper, we take $\mathbf{Y}_n = (Y_{n,1}, \ldots, Y_{n,K})$, $n \ge 1$. So,

$$\mathsf{P}\{X_{i,0} = \cdots = X_{i+[n^\epsilon],0} = 1\} \le \prod_{j=1}^{[n^\epsilon]} \frac{Z_{0,0}}{Z_{0,0} + A(j-1)} \le C\exp\{-n^\epsilon\}.$$

It follows that $\mathsf{E}[I_k(m,n)] \le Cn^{-\gamma(1-\epsilon)} + Cn\exp\{-n^\epsilon\} \le Cn^{-\gamma(1-\epsilon)}$. Hence $\sum_{j=n}^\infty \mathsf{E}[I_k(m,j)] \le Cn^{-\gamma(1-\epsilon)+1}$.

The next lemma gives the convergence rate of the remainders $R_{n,k}$, $k = 1, 2$.

Lemma 5.1. *Assume $\mathsf{E}[|D_{m,k}|^p] < \infty$ for any $p > 0$ and Condition* A *is satisfied. Then for any $\delta > 0$, we have*

$$\mathsf{E}\Big[\max_{m\le n}|R_{m,k}|\Big] = o\Big(n^{\frac{1}{(\varphi+1)+\delta}}\Big), \quad k = 1, 2, \tag{5.4}$$

$$|R_{n,k}| = o\Big(n^{\frac{1}{(\varphi+1)+\delta}}\Big) \quad a.s., \quad k = 1, 2. \tag{5.5}$$

Proof. (5.5) is implied by (5.4) if we notice that

$$\sum_{i=1}^\infty \mathsf{P}\Big(\max_{2^i\le n\le 2^{i+1}} \frac{|R_{n,k}|}{n^{\frac{1}{(\varphi+1)+2\delta}}} \ge \epsilon\Big) \le C\sum_{i=1}^\infty 2^{-i\delta} < \infty.$$

Now we need to verify (5.4). Fix k. For any $1 \le i \le n$,

$$\begin{aligned}
|R_{i,k}| &= \Big|\sum_{m=1}^i \sum_{j=i-m+1}^\infty X_{m,k}I_k(m,j)D_{m,k}\Big| \\
&\le \sum_{m=1}^i \sum_{j=i-m+1}^\infty I_k(m,j)|D_{m,k}|I\{|D_{m,k}| \le n^{\frac{\delta}{3}}\} + \sum_{m=1}^i |D_{m,k}|I\{|D_{m,k}| > n^{\frac{\delta}{3}}\} \\
&\le n^{\frac{\delta}{3}} \sum_{m=1}^i \sum_{j=i-m+1}^\infty I_k(m,j) + \sum_{m=1}^n |D_{m,k}|I\{|D_{m,k}| > n^{\frac{\delta}{3}}\}.
\end{aligned}$$

The expectation of the second term does not exceed

$$n\mathsf{E}[|D_{1,k}|I\{|D_{1,k}| > n^{\frac{\delta}{3}}\}] \le n^{1-\frac{\delta p}{3}}\mathsf{E}[|D_{1,k}|^p] \le C$$

whenever $p \ge 3/\delta$. So, it is enough to show that

$$\mathsf{E}\Big[\max_{i\le n}\overline{R}_{i,k}\Big] = O\Big(n^{\frac{1}{(\varphi+1)}+\frac{\delta}{2}}\Big), \tag{5.6}$$

where $\overline{R}_{i,k} = \sum_{m=1}^{i}\sum_{j=i-m+1}^{\infty} I_k(m,j)$. Let $1 \le P \le n$ be an integer whose value will be specified later. Then $\overline{R}_{i,k} \le P$ if $i \le P$. For $P \le i \le n$,

$$\begin{aligned}\overline{R}_{i,k} &= \sum_{m=i-P+1}^{i}\sum_{j=i-m+1}^{\infty} I_k(m,j) + \sum_{m=1}^{i-P}\sum_{j=i-m+1}^{\infty} I_k(m,j)\\ &\le P + \sum_{m=1}^{i-P}\sum_{j=P}^{\infty} I_k(m,j) \le P + \sum_{m=1}^{n}\sum_{j=P}^{\infty} I_k(m,j).\end{aligned}$$

It follows that $\mathsf{E}[\max_{i\le n}\overline{R}_{i,k}] \le P + \sum_{m=1}^{n}\sum_{j=P}^{\infty}\mathsf{E}[I_k(m,j)] \le P + CnP^{-\varphi}$. Choosing $P = [n^{1/(\varphi+1)+\delta/2}]$ yields (5.6).

Lemma 5.2. *Let $\mathcal{F}_n = \sigma(\mathbf{X}_1,\dots,\mathbf{X}_n,\mathbf{Y}_1,\dots,\mathbf{Y}_n)$. Let $V_{n,0} = \sum_{m=1}^{n}(X_{m,0} - \mathsf{E}[X_{m,0}|\mathcal{F}_{m-1}])$ and $V_{n,k} = \sum_{m=1}^{n}\{X_{m,k}(D_{m,k}-1) - \mathsf{E}[X_{m,k}(D_{m,k}-1)|\mathcal{F}_{n-1}]\}$, $k = 1,2$. Assume $\mathsf{E}[|D_{m,k}|^p] < \infty$ for $p \ge 2$. Then there exists a constant $C_p > 0$ such that the martingales $\{Y_{n,k},\mathcal{F}_n;\ n \ge 1\}$, $k = 0,1,2$, satisfy*

$$\mathsf{E}\Big[\max_{i\le n}|V_{m+i,k} - V_{m,k}|^p\Big] \le C_p n^{\frac{p}{2}} \quad \text{for all } m \text{ and } n, \quad k = 0,1,2. \tag{5.7}$$

Proof. Notice that $|\Delta V_{n,0}| \le 1$ and

$$\mathsf{E}\Big[|\Delta V_{n,k}|^p\Big|\mathcal{F}_{n-1}\Big] \le 2^{p-1}(1 + \mathsf{E}[|D_{n,k}|^p]) \le C_p, \quad k = 1,2.$$

(5.7) follows from the Rosenthal type inequality.

Let $U_{n,k} = a_k V_{n,0} + V_{n,k}$, $k = 1,2$. $U_{n,k}$ is the sum of conditionally centered changes in number of type k balls in the first n draws, $k = 1,2$. It can be shown that $\{U_{n,k},\mathcal{F}_n;\ n \ge 1\}$ is a martingale satisfying a similar inequality as (5.7), $k = 1,2$. The next lemma gives the convergence rate of the urn proportions $\mathbf{Z}_n$.

Lemma 5.3. *Under Assumption* 3.2 *and Condition* A, *for each $k = 1,2$ and any $\delta > 0$,*

$$\max_{j\le n} Z_{j,k} \le Z_{0,k} \vee \frac{a_k Z_{0,0}}{q_k} + 2\max_{j\le n}|U_{j,k}| + \max_{j\le n}|R_{j,k}|, \tag{5.8}$$

$$\mathsf{E}|Z_{n,k}| = o\Big(n^{\frac{1}{\varphi+1}+\delta}\Big), \tag{5.9}$$

$$\max_{j\le n}|Z_{j,k}| = o\Big(n^{\frac{1}{3}+\frac{1}{3\varphi+3}+\delta}\Big) \text{ in } L_1, \tag{5.10}$$

$$Z_{n,k} = o\Big(n^{\frac{1}{3}+\frac{1}{3\varphi+3}+\delta}\Big) \quad a.s.. \tag{5.11}$$

Proof. According to (5.2), it is obvious that

$$\begin{aligned} Z_{n,k} &= Z_{n-1,k} + \frac{a_k Z^+_{n-1,0} - q_k Z^+_{n-1,k}}{|\mathbf{Z}^+_{n-1}|} + \Delta U_{n,k} - \Delta R_{n,k} \\ &= Z_{n-1,k} + \frac{a_k Z_{0,0} - q_k Z^+_{n-1,k}}{|\mathbf{Z}^+_{n-1}|} + \Delta U_{n,k} - \Delta R_{n,k}. \end{aligned} \tag{5.12}$$

Then

$$\begin{aligned} &Z_{n,k} \le Z_{n-1,k} + \Delta U_{n,k} - \Delta R_{n,k}, \quad \text{if } Z_{n-1,k} \ge a_k\frac{Z_{0,0}}{q_k}; \\ &Z_{n,k} \le Z_{n-1,k} + a_k + \Delta U_{n,k} - \Delta R_{n,k}, \quad \text{if } Z_{n-1,k} < a_k\frac{Z_{0,0}}{q_k}. \end{aligned} \tag{5.13}$$

Let $S_n = \max\{1 \le j \le n : Z_{j,k} < a_k Z_{0,0}/q_k\}$, where $\max(\emptyset) = 0$. Then, according to (5.13),

$$\begin{aligned} Z_{n,k} &\le Z_{n-1,k} + \Delta U_{n,k} - R_{n,k} + R_{n-1,k} \le \cdots \\ &\le Z_{S_n,k} + \Delta U_{S_n+1,k} + \cdots + \Delta U_{n,k} - R_{n,k} + R_{S_n,k} \\ &\le Z_{0,k} \vee \Big\{a_k\frac{Z_{0,0}}{q_k}\Big\} + U_{n,k} - U_{S_n,k} - R_{n,k} + R_{S_n,k} \\ &\le Z_{0,k} \vee \Big\{a_k\frac{Z_{0,0}}{q_k}\Big\} + U_{n,k} - U_{S_n,k} + \max_{m\le n}|R_{m,k}|. \end{aligned} \tag{5.14}$$

(5.8) is proved. Notice that $S_n \le n$ is a stopping time. It follows that $\mathsf{E}U_{n,k} = \mathsf{E}U_{S_n,k}$. By (5.4) and (5.14) we conclude that $\mathsf{E}Z_{n,k} \le o(n^{1/(\varphi+1)+\delta})$. (5.9) is proved by the fact that $Z_{n,k} \ge -1$ and $|Z_{n,k}| = Z_{n,k} + 2Z^-_{n,k}$.

Next, we verify (5.11). Fix m. By replacing $Z_{j,k}$ with $Z_{m+j,k}$ in the definition of the stopping time S_n, with similar arguments as in showing (5.8) we can show that

$$\max_{0\le i\le n} Z_{i+m,k} \le Z_{m,k} \vee \frac{a_k Z_{0,0}}{q_k} + 2\max_{0\le i\le n}|U_{m+i,k} - U_{m,k}| + \max_{j\le n+m}|R_{j,k}|. \tag{5.15}$$

Now, for each $p \geq 2$ and $0 < t < 1/2$, if $n \geq 1/(4t)$, then by (5.7), (5.9) and (5.15),

$$\begin{aligned}
\mathsf{E}\max_{m\leq n} Z_{m,k} &\leq \mathsf{E}\Big[\max_i \max_{i[nt]\leq m\leq (i+1)[nt]} Z_{m,k}\Big] \\
&\leq a_k\frac{Z_{0,0}}{q_k} + \sum_i \mathsf{E}\Big[|Z_{i[nt],k}|\Big] + \mathsf{E}\Big[\max_{j\leq n}|R_{j,k}|\Big] \\
&\quad + 2\mathsf{E}\Big[\max_i \max_{i[nt]\leq m\leq (i+1)[nt]} |U_{m,k} - U_{i[nt],k}|\Big] \\
&\leq C\Big\{t^{-1}n^{\frac{1}{\varphi+1}+\delta} + n^{\frac{1}{\varphi+1}+\delta}\Big\} + 2\Big(\mathsf{E}\Big[\max_i \max_{i[nt]\leq m\leq (i+1)[nt]} |U_{m,k} - U_{i[nt],k}|^p\Big]\Big)^{\frac{1}{p}} \\
&\leq C\Big\{t^{-1}n^{\frac{1}{\varphi+1}+\delta} + n^{\frac{1}{\varphi+1}+\delta}\Big\} + 2\Big(\sum_i \mathsf{E}\Big[\max_{i[nt]\leq m\leq (i+1)[nt]} |U_{m,k} - U_{i[nt],k}|^p\Big]\Big)^{\frac{1}{p}} \\
&\leq C\Big\{t^{-1}n^{\frac{1}{\varphi+1}+\delta} + n^{\frac{1}{\varphi+1}+\delta} + \Big(\sum_i([nt])^{\frac{p}{2}}\Big)^{\frac{1}{p}}\Big\} \\
&\leq C\Big\{t^{-1}n^{\frac{1}{\varphi+1}+\delta} + t^{\frac{1}{2}-\frac{1}{p}}n^{\frac{1}{2}}\Big\},
\end{aligned}$$

where the sums and maximums are taken over $\{i \geq 0 : i[nt] \leq n\}$. Here $C > 0$ is a constant and does not depend on t and n. Notice that $Z_{m,k} \geq -1$. If $t = n^{-1/3+2/(3\varphi+3)}$, we have

$$\mathsf{E}\Big[\max_{m\leq n}|Z_{m,k}|\Big] \leq 2 + \mathsf{E}\Big[\max_{m\leq n} Z_{m,k}\Big] \leq Cn^{\frac{1}{3}+\frac{1}{3(\varphi+1)}+\delta} + Cn^{\frac{1}{3}+\frac{1}{3(\varphi+1)}+\frac{1}{p}(\frac{1}{3}-\frac{2}{3(\varphi+1)})}.$$

Choosing p such that $1/p(1/3 - 2/[3(\varphi + 1)]) \leq \delta$ yields (5.10) immediately. With the same arguments as in showing (5.4) from (5.7), (5.11) can be derived easily from (5.10) and the Borel-Cantelli Lemma. The proof of the lemma is now complete. (5.11) and (5.4) indicates that the terms $Z_{n,k}$ and $R_{n,k}$ in (5.3) can be neglected.

Now we begin the proof of Theorem 3.1. Let $s = a_1/q_1 + a_2/q_2$. Then $v_k = (a_k/q_k)/s$, $k = 1, 2$. Let $\mathcal{A}_n = \sigma(\mathbf{X}_1, \ldots, \mathbf{X}_n, \mathbf{X}_{n+1}, \mathbf{Y}_1, \ldots, \mathbf{Y}_n)$, $M_{n,k} = \sum_{m=1}^n X_{m,k}(D_{m,k} - \mathsf{E}D_{m,k})$, $k = 1, 2$. Then $\{(M_{n,1}, M_{n,2}), \mathcal{A}_n;\ n \geq 1\}$ is a martingale with

$$\sum_{m=1}^n \mathsf{E}\Big[(\Delta M_{m,k})^2|\mathcal{A}_{m-1}\Big] = \sum_{m=1}^n X_{m,k}\mathrm{Var}\,(D_{m,k}) = N^*_{n,k}\sigma_k^2, \tag{5.16}$$

$\mathsf{E}[\Delta M_{n,k} \cdot \Delta M_{n,j}|\mathcal{A}_{n-1}] = 0$, $j \neq k$, and $\mathsf{E}[|\Delta M_{n,k}|^p|\mathcal{A}_{n-1}] \leq 2^p\mathsf{E}[|D_{1,k}|^p]$. According to the law of the iterated logarithm for martingales, we have

$$M_{n,k} = O(\sqrt{n\log\log n}) \quad a.s.. \tag{5.17}$$

Combining (5.3), (5.5), (5.11) and (5.17) yields that, for any $\delta > 0$,

$$\begin{aligned} a_k N^*_{n,0} - q_k N^*_{n,k} &= -M_{n,k} + o\Big(n^{\frac{1}{3}+\frac{1}{3\varphi+3}+\frac{\delta}{2}}\Big) \\ &= -M_{n,k} + o\Big(n^{\frac{\gamma+1}{3\gamma}+\delta}\Big) \\ &= O(\sqrt{n\log\log n}) \quad a.s., \quad k = 1, 2. \end{aligned} \tag{5.18}$$

Together with the fact that $N^*_{n,0} + N^*_{n,1} + N^*_{n,2} = n$, we have

$$\begin{aligned} N^*_{n,k} &= n\frac{\frac{a_k}{q_k}}{\frac{a_1}{q_1} + \frac{a_2}{q_2} + 1} + O(\sqrt{n\log\log n}) \\ &= n\frac{s}{s+1}v_k + O(\sqrt{n\log\log n}) \quad a.s., \quad k = 1, 2, \end{aligned} \tag{5.19}$$

$$\begin{aligned} a_1 q_2 N^*_{n,2} - a_2 q_1 N^*_{n,1} &= a_2(a_1 N^*_{n,0} - q_1 N^*_{n,1}) - a_1(a_2 N^*_{n,0} - q_2 N^*_{n,2}) \\ &= a_1 M_{n,2} - a_2 M_{n,1} + o\Big(n^{\frac{\gamma+1}{3\gamma}+\delta}\Big) \quad a.s.. \end{aligned} \tag{5.20}$$

We consider the martingale $\{M_n =: a_1 M_{n,2} - a_2 M_{n,1}\}$. From (5.16) and (5.19), it follows that

$$\begin{aligned} &\sum_{m=1}^{n} \mathsf{E}\Big[(\Delta M_m)^2 | \mathcal{A}_{m-1}\Big] \\ &\quad = \sum_{m=1}^{n} a_1^2 \mathsf{E}\Big[(\Delta M_{m,2})^2 | \mathcal{A}_{m-1}\Big] + a_2^2 \sum_{m=1}^{n} \mathsf{E}\Big[(\Delta M_{m,1})^2 | \mathcal{A}_{m-1}\Big] \\ &\quad = n\frac{s}{s+1}(a_1^2 v_2 \sigma_2^2 + a_2^2 v_1 \sigma_1^2) + O(\sqrt{n\log\log n}) \quad a.s.. \end{aligned}$$

By the Skorokhod Embedding Theorem (cf., Hall and Heyde (1980)), there exists an $\mathcal{A}_n$-adapted non-decreasing sequence of random variables $\{T_n\}$ and a standard Brownian motion B, such that

$$\mathsf{E}\Big[\Delta T_n | \mathcal{A}_{n-1}] = \mathsf{E}\Big[(\Delta M_n)^2 | \mathcal{A}_{n-1}\Big], \quad \mathsf{E}|\Delta T_n|^{\frac{p}{2}} \le C_p \mathsf{E}|\Delta M_n|^p \le c_p, \quad \forall p > 2,$$

$$M_n = B(T_n), \quad n = 1, 2, \dots . \tag{5.21}$$

Note that $\{\sum_{m=1}^{n}(\Delta T_m - \mathsf{E}[\Delta T_m | \mathcal{A}_{m-1}])\}$ is also a martingale. According to the Law of the Iterated Logarithm, we have

$$\begin{aligned} T_n &= \sum_{m=1}^{n} \mathsf{E}[\Delta T_m | \mathcal{A}_{m-1}] + O(\sqrt{n\log\log n}) \\ &= n\frac{s}{s+1}(a_1^2 v_2 \sigma_2^2 + a_2^2 v_1 \sigma_1^2) + O(\sqrt{n\log\log n}) \quad a.s.. \end{aligned}$$

On the other hand, by (5.19), we have $N^*_{n,1}+N^*_{n,2} = [s/(s+1)]n+O(\sqrt{n\log\log n})$ a.s.. So, $u_n = \max\{m : N^*_{m,1} + N^*_{m,2} \le n\} = [(s+1)/s]n + O(\sqrt{n\log\log n})$ a.s.. It follows that

$$T_{u_n} = n(a_1^2 v_2\sigma_2^2 + a_2^2 v_1\sigma_1^2) + O(\sqrt{n\log\log n}) \quad a.s.. \tag{5.22}$$

Substituting (5.22), (5.21) into (5.20) and applying the properties of Brownian motion (cf., Theorem 1.2.1 of Csörgő and Révész (1981)), we have

$$\begin{aligned} a_1q_2N_{n,2} - a_2q_1N_{n,1} &= a_1q_2N^*_{u_n,2} - a_2q_1N^*_{u_n,1} = B(T_{u_n}) + o\Big(u_n^{\frac{\gamma+1}{3\gamma}+\delta}\Big) \\ &= B\Big(n(a_1^2v_2\sigma_2^2 + a_2^2v_1\sigma_1^2)\Big) + O\Big((n\log\log n)^{\frac14}(\log n)^{\frac12}\Big) + o\Big(n^{\frac{\gamma+1}{3\gamma}+\delta}\Big) \\ &= B\Big(n(a_1^2v_2\sigma_2^2 + a_2^2v_1\sigma_1^2)\Big) + o\Big(n^{\frac{\gamma+1}{3\gamma}+\delta}\Big) \quad a.s.. \end{aligned}$$

Together with the fact that $N_{n,1} + N_{n,2} = n$, we have

$$N_{n,1} = \frac{a_1q_2}{a_1q_2+a_2q_1}n - \frac{B\Big(n(a_1^2v_2\sigma_2^2 + a_2^2v_1\sigma_1^2)\Big)}{a_1q_2+a_2q_1} + o\Big(n^{\frac{\gamma+1}{3\gamma}+\delta}\Big) \quad a.s..$$

Notice that $\sigma^2 = (a_2^2v_1\sigma_1^2 + a_1^2v_2\sigma_2^2)/(a_1q_2+a_2q_1)^2$, where σ^2 is defined in (3.1). Let

$$W(t) = -\frac{1}{\sigma}\frac{B\Big(t(a_1^2v_2\sigma_2^2 + a_2^2v_1\sigma_1^2)\Big)}{a_1q_2+a_2q_1}.$$

Then $\{W(t); t \ge 0\}$ is a standard Brownian motion. The proof of Theorem 3.1 is complete.

Acknowledgement

The authors are grateful to Professor William F. Rosenberger and two referees for providing several helpful comments which led to a significant improvement of the article. This work was supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region (Project no. CUHK400204). L.X. Zhang's work was partially supported by grants from the National Science Foundation of China (Project no. 10471126).

References

Andersen, J., Faries, D. and Tamura, R. (1994). A randomized Play-the-Winner design for multi-arm clinical trials. *Comm. Statist. Theory Methods* **23**, 309-323.

Athreya, K. B. and Karlin, S. (1968). Embedding urn schemes into continuous time branching processes and related limit theorems. *Ann. Math. Statist.* **39**, 1801-1817.

Athreya, K. B. and Ney, P. E. (1972). *Branching Processes*. Spring-Verlag, Berlin.

Bai, Z. D., Hu, F. and Rosenberger, W. F. (2002). Asymptotic properties of adaptive designs for clinical trials with delayed response. *Ann. Statist.* **30**, 122-139.

Beggs, A. W. (2005). On the convergence of reinforcement learning. *J. Econom. Theory* **122**, 1-26.

Csörgő, M. and Révész, P. (1981). *Strong Approximations in Probability and Statistics*. Academic Press, New York.

Durham, S. D., Flournoy, F. and Li, W. (1998). A sequential design for maximizing the probability for a favorable response. *Canad. J. Statist.* **26**, 479-495.

Eisele, J. (1994). The doubly adaptive biased coin design for sequential clinical trials. *J. Statist. Plann. Inference* **38**, 249-262.

Eisele, J. and Woodroofe, M. (1995). Central limit theorems for doubly adaptive biased coin designs. *Ann. Statist.* **23**, 234-254.

Hall, P. and Heyde, C. C. (1980). *Martingale Limit Theory and its Applications*. Academic Press, London.

Hopkins, E. and Posch, M. (2005). Attainability of boundary points under reinforcement learning. *Games Econom. Behav.* **53**, 110-125.

Hu, F. and Rosenberger, W. F. (2003). Optimality, variability, power: evaluating response-adaptive randomization procedures for treatment comparisons. *J. Amer. Statist. Assoc.* **98**, 671-678.

Hu, F., Rosenberger, W. F. and Zhang, L.-X. (2006). Asymptotically best response-adaptive randomization procedures. *J. Statist. Plann. Inference* **136**, 1911-1922.

Hu, F. and Zhang, L.-X. (2004a). Asymptotic properties of doubly adaptive biased coin designs for multi-treatment clinical trials. *Ann. Statist.* **32**, 268-301.

Hu, F. and Zhang, L.-X. (2004b). Asymptotic normality of urn models for clinical trials with delayed response. *Bernoulli* **10**, 447-463.

Ivanova, A. (2003). A play-the-winner type urn model with reduced variability. *Metrika* **58**, 1-13.

Ivanova, A. and Flournoy, N. (2001). A birth and death urn for ternary outcomes: stochastic processes applied to urn models. In *Probability and Statistical Models with Applications* (Edited by Ch. A. Charalambides, M. V. Koutras and N. Balakrishnan), 583-600. Chapman & Hall/CRC.

Ivanova, A., Rosenberger, W. F., Durham, S. D. and Flournoy, N. (2000). A birth and death urn for randomized clinical trials: asymptotic methods. *Sankhyā Ser. B* **62**, 104-118.

Melfi, V. F. and Page, C. (1998). Variability in adaptive designs for estimation of success probabilities. In *New Developments and Applications in Experimental Design* (Edited by Flournoy, N., Rosenberger, W. F. and Wong, W. K.), 106-114, Hayward: Institute of Mathematical Statistics.

Melfi, V. F. and Page, C. (2000). Estimation after adaptive allocation. *J. Statist. Plann. Inference* **87**, 353-363.

Robbins, H. (1952). Some aspects of the sequential design of experiments. *Bull. Amer. Math. Soc.* **58**, 527-535.

Rosenberger, W. F. (1996). New directions in adaptive designs. *Statist. Sci.* **11**, 137-149.

Rosenberger, W. F. and Hu, F. (2004). Maximizing power and minimizing treatment failures. *Clinical Trials* **1**, 141-147.

Rosenberger, W. F. and Lachin, J. M. (2002). *Randomization in Clinical Trials: Theory and Practice.* Wiley, New York.

Rosenberger, W. F., Stallard, N., Ivanova, A., Harper, C. N. and Ricks, M. L. (2001). Optimal adaptive designs for binary response trials. *Biometrics* **57**, 909-913.

Thompson, W. R. (1933). On the likelihood that one unknown probability exceeds another in the revises of the evidence of the two samples. *Biometrika* **25**, 275-294.

Wei, L. J. (1979). The generalized Polya's urn for sequential medical trials. *Ann. Statist.* **7**, 291-296.

Wei, L. J. (1988). Constructing exact two-sample permutational tests with the randomized play-the-winner rule. *Biometrika* **75**, 603-606.

Wei, L. J. and Durham, S. D. (1978). The randomized play-the-winner rule in medical trials. *Amer. Statist. Assoc.* **73**, 840-843.

Zhang, L. X., Hu, F. and Cheung, S. H. (2006). Asymptotic theorems of sequential estimation-adjusted urn models. *Ann. Appl. Probab.* **16**, 340-369.

Department of Mathematics, Zhejiang University, Hangzhou 310028, PR China.

E-mail: lxzhang@mail.hz.zj.cn

Department of Finance, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong, PR China.

E-mail: chanws@cuhk.edu.hk

Department of statistics, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong, PR China.

E-mail: shcheung@sta.cuhk.edu.hk

Department of Statistics, University of Virginia, Halsey Hall, Charlottesville, Virginia 22904-4135, USA.

E-mail: fh6e@virginia.edu

(Received August 2004; accepted August 2005)