Adaptive designs for sequential experiments

LIN Zheng-ya, ZHANG Li-xi
Department of Mathematics, Zhejiang University, Hangzhou 310028, China
E-mail address: zlin@mail.lz.jz.cn  lzxhang@mail.lz.jz.cn
Received May 6, 2002; revision accepted June 21, 2002

Abstract Various adaptive designs have been proposed and applied to clinical trials, bioassay, psychophysics etc. Adaptive designs are also useful in high cost engineering trials. More and more people have been paying attention to these design methods. This paper introduces several broad families of designs such as the play-the-winner rule randomized play-the-winner rule and its generalization to the multi-arm case doubly biased coin adaptive design Markov chain model.

Key words Clinical trial, Adaptive design, Randomized Play-the-winner rule, Biased coin design, Markov chain, Asymptotic properties, Um model.

Document code A CLC number 0211.4

INTRODUCTION

Traditional designs for clinical trials use the balanced or 50-50 allocation of patients to treatments. For example, in a trial to compare experimental therapy drug to placebo control a standard feature of most designs is to distribute half of patients to each arm. It is reasonable that one may want to reduce the total number of failure outcomes in a trial and keep the capability of making a comparison between experimental therapy and placebo as well. Hence the idea of adaptive designs has been proposed to serve the purpose.

Adaptive design is an important subdivision of experimental designs are designs in which the probability a treatment assigned to the coming patient depends upon the results of the previous patients in the study. The goal is to show assignment probabilities to favor better treatment performance. This kind of design has also been applied to bioassay psychophysics etc. They are also useful in high cost engineering trials.

PLAY-THE-WINNER RULE AND RANDOMIZED PLAY-THE-WINNER RULE

Consider a two-arm clinical trial two treatments with dichotomous response success and failure. Patients are recruited into the clinical trial sequentially and respond immediately to treatments. Zelen 1969 propose the following design which is well known as the play-the-winner PW rule. A success on a particular treatment generates a future trial on the same treatment with a new patient. A failure on a treatment generates a future trial on the alternate treatment. Let \( N_{n1} \) and \( N_{n2} \) be the number of the patients assigned to the treatment 1 and 2 respectively in the first n stages. And let \( p_i = P(\text{success} | \text{treatment } i) \) be the success probability of a patient on the treatment \( i \) \( q_i = 1 - p_i \) \( i = 1, 2 \). Then

\[
\frac{N_{n1}}{n} \rightarrow \frac{q_2}{q_1 + q_2} \quad a.s.
\]

and

\[
\sqrt{n} \left( \frac{N_{n1}}{n} - \frac{q_2}{q_1 + q_2} \right) \rightarrow N(0, \sigma^2_{PW})
\]

where \( \sigma^2_{PW} = q_1 q_2 p_1 + p_2 q_1 + q_2 q_2^3 \). This is first discussed in Zelen 1969.

As pointed out in Wei et al. 1978 and Wee 1979 the PW rule is too deterministic and is not applicable when we have delayed responses from patients to treatments. Motivated as

\* Project supported by the National Natural Science Foundation of China Nos. 19571021 and 10071072.

\[
\text{ISSN} 1009-3095 \text{ Journal of Zhejiang University SCIENCE V. 4} \text{ No. 2} \text{ P. 214 - 220 Mar.} - \text{ Apr. 2003} \\
\text{http://www.periodicals.com.cn} \text{ http://www.zju.edu.cn/English} \\
\text{http://www.zjupress.com.cn} \text{ http://lib.zju.edu.cn/eindex.htm} \text{ zjus@zju.edu.cn}
\]
an extension to Zelen’s 1969 idea. Wei et al. 1978 introduced the following randomized play-the-winner RPW rule. We start with $\alpha$ $\alpha$ balls type 1 and 2 respectively in the urn. If a type 1 ball is drawn a patient is assigned to the treatment 1. If a type 2 ball is drawn a patient is assigned to the treatment 2. The ball is replaced and the patient response is observed. A success on the treatment 1 or a failure on the treatment 2 generates a type 1 ball in the urn. A success on the treatment 2 or a failure on the treatment 1 generates a type 2 ball in the urn. The RPW rule may be regarded as a generalized Polya urn GPU model Wei 1979. Further let $Y_{d1}$ $Y_{d2}$ be number of balls of type 1 2 after $n$ stage. From the results of Athreya et al. 1968 we have

$$\frac{Y_{n1}}{Y_{n1} + Y_{n2}} \to q_2 \quad \text{a.s.}$$

$$\frac{N_{n1}}{n} \to q_2 \quad \text{a.s.}$$

When $p_1 + p_2 < 1 \quad \text{or} \quad q_1 + q_2 > 0.5$ we have the following asymptotic normality

$$\sqrt{n} \left( \frac{Y_{n1}}{Y_{n1} + Y_{n2}} - \frac{q_2}{q_1 + q_2} \right) \to D \quad 1\text{D}$$

$$\frac{N_{n1}}{n} - \frac{q_2}{q_1 + q_2} \to D \quad 2\text{D}$$

where

$$\sigma^2_{\text{RPW}} = \frac{q_1 q_2 5 - q_1^2 - q_2^2}{q_1 + q_2 - 1}.$$  

The asymptotic normality was first given in Smythe et al. 1995. When $q_1 + q_2 < 0.5$ the limiting distributions of both the urn composition and the proportions of patients assigned to each treatments are unknown. The RPW rule is not deterministic and allows delayed responses by the patients.

If treatment 1 is "doing better" both PW rule and RPW rule are shown to favor treatment 1.

MULTI-ARM CASE RPW RULE AND GENERALIZED POLYA URN

One large family of randomized adaptive designs can be developed from the GPU model. Consider an urn containing balls of $K$ types. Initially the urn contains $Y_0 = Y_{01} \cdots Y_{0K}$ balls where $Y_{0i}$ denotes the number of balls of type $Y_{0i}$ $i = 1 \cdots K$. A ball is drawn at random from the urn. Its type is observed and the ball is then replaced. At the $n$th stage following a type $i$ draw $D_1 n$ balls of type $j$ for $j = 1 \cdots K$ are added to the urn. In the most general sense $D_1 n$ can be random and can be some function of a random process outside the urn process in the case of adaptive designs $D_1 n$ will be a random function of patient response. A ball must always be added at each stage in addition to the replacement and the expectation of the total numbers of balls added in each stage is the same say $Y_{n}$ so

$$P[D_1 n = 0] \quad \text{for all} \quad j = 1 \cdots K \quad \text{so}$$

$$\sum_{j=1}^{K} D_{1j} n = \gamma \quad \text{for} \quad \gamma$$

$$\gamma = \frac{Y_{n1} \cdots Y_{nk}}{n} \quad \text{where} \quad Y_{ni}$$

represents the number of balls in the urn of type $i$ after $n$ stage. and

$$N_n = Y_{n1} \cdots Y_{nk} \quad \text{where} \quad N_{ni}$$

represents the number of times a ball i drawn in the first $n$ draws. In the clinical trials $N_{ni}$ is the number of patients assigned to treatment $i$ in the first $n$ stages. Let $v_i = v_{i1} \cdots v_{ik}$ be the left eigenvector corresponding to the largest eigenvalue of $H$ with $v_{i1} + \cdots + v_{ik}$ is 1. Then $v_i$ is just the limiting proportion of both the patients assigned to treatment $i$ the type $i$ balls in the urn i.e.

$$\frac{N_{ni}}{n} \to v_i \quad \text{a.s.}$$

$$\sum_{j=1}^{K} Y_{nj} \to v_i \quad \text{a.s.}$$

Athreya et al. 1968 Smyth 1996 and Bai et al. 1999 2000 showed the normality of $Y_n$
and $N_n$ let $\lambda_1 = \gamma = 1 \lambda_2 \ldots \lambda_K$ be the eigenvalues of $H^*$ and $\lambda = \max \{ R \} \lambda_2 \ldots \lambda_K$ if $\lambda < 1/2$ and
\[
\sum_{n=1}^{\infty} \frac{\| H_n - H \|}{\sqrt{n}} < \infty
\]
then
\[
\sqrt{n} \left( \frac{Y_n}{\sqrt{n}} - v \right) \xrightarrow{D} N(0, \Sigma^*)
\]
and
\[
\sqrt{n} \left( \frac{N_n}{n} - v \right) \xrightarrow{D} N(0, \Sigma^*). \tag{6}
\]

Under similar conditions we showed that $Y_n - nv$ and $N_n - nv$ can be approximated by Gaussian processes
\[
G_{n1} H + G_{n2} \quad \text{and} \quad G_{n1} + \int_0^1 x \frac{G_{n2}}{x} \, dx I - 1'v^2
\]
respectively. Bai et al. 2002c 2001'105 where $G_{n2}$ are independent Gaussian processes which are solutions of the following type equation.
\[
G_t = W_t + \int_0^t \frac{G_s}{s} \, H - 1'v^2 \, ds \quad t > 0
\]
with $G_0 = 0$.

Here $W_t$ is a 3-dimensional Brownian motion. In particular the limiting combining distribution of $Y_n$ and $N_n$ was obtained. For more results one can refer to Bai et al. 1999a, 1999b and Rosenberger 1996.

Define $X_n = i$ if a ball of type $i$ is drawn at $n$th stage $T_n = 1$ if the response of the $n$th patient is a "success" otherwise. Assume that
\[
P \quad T_n = 1 \mid X_n = i \quad \text{i} \quad \text{is independent of } n \quad \text{hence we can write}
\]
\[
p_i = \mathbb{E} \quad T_n = 1 \mid X_n = i \quad q_i = 1 - p_i. \tag{7}
\]

Now we give some important special cases.

**Case 1.** Let $\mathcal{H}^{10} = \{ H_{ij} \}_{i,j=1}^K$ where $\mathcal{H}_{ij}^{10} = p_i$ and $\mathcal{H}_{ij}^{10} = q_j \quad K - 1 \quad i \neq j$. Then $v_i$ in Eq. 3 is
\[
\mathcal{H}_{ij}^{10} = \frac{1/q_k}{\sum_{j=1}^{n} q_j}.
\]

This model was proposed by We\[ 1979 \] which means that at the $n$th stage if a patient is assigned to treatment $i$ and cured then a type $i$ ball is added to the urn otherwise if treatment $i$ for a patient fails then $1/K - 1$ balls are added to each of the other $K - 1$ treatments.

**Case 2.** Let $\mathcal{H}^{20} = \{ H_{ij}^{20} \}_{i,j=1}^K$ where $\mathcal{H}_{ij}^{20} = p_i$ and $\mathcal{H}_{ij}^{20} = q_j \quad M - p_i \quad i \neq j$ and $M$
\[
= \sum_{j=1}^{K} p_i. \quad \text{Then } v_i \quad \text{in Eq. 3 is}
\]
\[
\mathcal{H}_{ij}^{20} = \frac{1/q_k}{\sum_{j=1}^{n} q_j}.
\]

This model was proposed by Bai et al. 2002b which means that at the $n$th stage if a patient is assigned to treatment $i$ and cured then a type $i$ ball is added into the urn otherwise if treatment $i$ for the patients fails then a number of balls proportional to the success rates of the other $K - 1$ treatments are added into the urn. But this design is not practicable since the success probabilities $p_j$ are unknown. They should be replaced by their sample estimators. This leads to the following case.

**Case 3.** Let $S_n = S_1 \ldots S_n$ where $S_n$ denotes the number of successes of the $i$th treatment in the $N_n$ trials $r_{ni} = \sum_{j=1}^{K} N_{nj}$ and $M_{n-1} = \sum_{j=1}^{K} n_{j-1}$ and let
\[
\mathcal{H}^{30} = \mathcal{H}^{30} = \{ p_1 \quad \frac{r_{n-1}q_1}{M_{n-1} - r_{n-1}q_1} \quad \frac{r_{n-1}q_2}{M_{n-1} - r_{n-1}q_2} \quad \cdots \quad \frac{r_{n-1}q_k}{M_{n-1} - r_{n-1}q_k} \}
\]

**Case 4.** The success probabilities defined in Eq. 7 are homogeneous for different stage patients. But this assumption is not always realistic in certain situations where patients may exhibit arbitrary drift in characteristics over time i.e.
\[
\mathcal{H}^{40} = \{ p_1 \quad q_1 \quad \cdots \quad p_k \} \quad n \quad n = \mathbb{E} \quad T_n = 1 \mid X_n = i \quad q_i \quad \frac{n_{j-1}}{n-1 - j} \quad n. \tag{8}
\]

It is natural to assume that the treatments are statistically equivalent.
ble in such case i.e. it is assumed that \( p_i \) \( n_i \to p_i \, n_i = 1 \ldots K \). In such case the design in Case 1 becomes \( \hat{H}^\Pi = \hat{H}^\Pi_{n_i} \hat{H}^\Pi_{y_j} K = 1 \) where \( \hat{H}^\Pi_{n_i} = p_i \, n_i \) and \( \hat{H}^\Pi_{y_j} = q_j \, n_j \). 

Both \( \hat{H}^\Pi \) and \( \hat{H}^\Xi \) are independent of \( n_i \) the stage number and so the models are homogeneous. But \( \hat{H}^\Xi \) and \( \hat{H}^\Pi \) are dependent on \( n_i \) and so the models are non-homogeneous. 

Remark 1. When \( K = 2 \) the above three adaptive designs are all governed by the RPW rule. So all of these designs are extensions of the RPW rule for \( K \) treatments. 

Remark 2. For designs \( \hat{H}^\Pi \) and \( \hat{H}^\Xi \) when \( K > 2 \) notice that for any \( i \neq j \) \( 1 \leq i \leq K \) if \( p_i > p_j \) we have 

\[
\frac{\hat{p}_i}{\hat{p}_j} - \frac{\hat{p}_j}{\hat{p}_i} = \frac{p_i - p_j}{p_i \hat{M} - p_i \sqrt{q_i} - 1/q_i} = \frac{p_i - p_j}{p_i \sum_{k \neq j} p_k / q_j} > 0.
\]

The design \( \hat{H}^\Xi \) is more reasonable than the design \( \hat{H}^\Pi \) because more patients will be assigned to a better treatment. 

Remark 3. Let \( a \geq 0 \). An extension of the design \( \hat{H}^\Xi \) has \( r_{n_i} \) and \( M_{n_i} \) \( a = \sum_{j=1}^{k} r_{n_i} - r_{n_j} \) instead of \( r_{n_i} - r_{n_j} \) and \( M_{n_i} - 1 \) respectively. Taking \( a = 0 \) and 1 yields designs \( \hat{H}^\Pi \) and \( \hat{H}^\Xi \) respectively. Designs dependent on estimated unknown parameters The asymptotic normality of Case 1 and Case 2 follows from Eqs 5 and 6 immediately. In the Case 4 if 

\[
\sum_{n_i=1}^{\infty} \left| \frac{1}{\sqrt{n}} \right| < \infty \quad i = 1 \ldots K \]

then the condition 4 and so asymptotic properties also follow. But the model in Case 3 can not satisfy the condition 4 since the fastest convergence rate of \( H_n \) is that 

\[
\| H_n - H \| = \sqrt{n} \] in probability. The asymptotic normality of such case were studied in Hu and Zhang*. We considered the general model with design matrices of the type \( \hat{H}_n = \hat{H} \hat{\Theta}_n \) dependent on the estimated unknown parameter \( \hat{\Theta}_n \). Strong consistency and the asymptotic normalities are established for both the urn composition and the number of patients assigned to each treatment. It should be mentioned that the condition 4 cannot be reduced to \( \| H_n - H \| = \emptyset \) \( \sqrt{n} \) in general. 

Designs with delay responses Typically clinical trials do not result in immediate outcomes i.e. individual patient outcomes may not be immediately available prior to the randomization of the next patient. Consequently the urn cannot be updated immediately but can be updated only when the outcomes become available. Fortunately it is verified that stochastic staggered entry and delay mechanisms do not affect the asymptotic properties of both the urn composition \( Y_n \) and the sample fractions \( N_n \) for a wide class of designs defined by GPL Bai et al. 2002a. 

DOUBLY BIASED COIN ADAPTIVE DESIGNS 

We come back to the PW rule and RPW rule. As it is known the PW rule is too deterministic and is not applicable when we have delayed responses from patients of treatments. The RPW rule and its generalizations seem solve this problem. However in using the RPW rule when \( q_1 + q_2 < 0.5 \) the limiting distributions of the proportions of patients assigned to each treatments are unknown. But in practice both \( q_1 \) and \( q_2 \) are usually very small. So the RPW rule is not practical in such cases. The asymptotic variation of the proportion becomes a big problem in using the RPW rule even the adaptive designs based on the GPU model for \( q_1 + q_2 < 0.5 \). Even in the case that \( q_1 + q_2 > 0.5 \) if \( q_1 + q_2 \) is near 0.5 \( \sigma_{RPW} \) is much larger than \( \sigma_{PW} \). That is to say the RPW rule is too random so that the asymptotic variance of proportion of patients assigned to each treatment is very large when the cure rates are large and so it is much less stable than the PW rule. Also in using the multi-arm RPW the condition that \( \lambda < 1/2 \) is very hard to check. Even in the 3-arm case it is very difficult to check this condition. 

Now with keeping the desired allocation pro-

portions \( v_1 = q \frac{\alpha}{m} \) \( q_1 + q_2 \) and \( v_2 = q \frac{\alpha}{m} q_1 + q_2 \) just as in the case of the PW rule and RPW rule our goal is to reduce the asymptotic variance. A natural way is as follows. At the \( n + 1 \) th stage we assign a patient to a certain treatment by comparing the value \( N_{n+1}/n \) with \( v_1 \) or \( N_{n+2}/n \) with \( v_2 \). If \( N_{n+1}/n \) is larger than \( v_1 \) then we assign a patient to the treatment 1 with a probability less than \( v_1 \). If \( N_{n+1}/n \) is less than \( v_1 \) then we assign a patient to the treatment 1 with a probability larger than \( v_1 \). If \( N_{n+1}/n \) equals \( v_1 \) then we assign a patient to the treatment 1 with probability \( v_1 \) and to the treatment 2 with probability \( v_2 \). By choosing suitable function we may minimize the asymptotic variance. However \( v_1 \) and \( v_2 \) are unknown so they should be replaced by their estimators based on the sample of the previous \( n \) stages. So the following adaptive design of clinical trial is considered and proposed.

At the first stage a patient is assigned to each treatment with the same probability 1/2. After \( m \) assignments we let \( S_{m} \) be the number of successes of all the \( N_{m} \) patients on the treatment \( k \) in the first \( m \) assignments \( k = 1 \) as usual.

And let \( \hat{\rho}_{m} = \frac{S_{m} + 1/2}{N_{m} + 1} \) be the sample estimation of \( \rho_{m} \) and write \( \hat{\rho}_{m} = \hat{\rho}_{m} \) \( k = 1 \). At the \( m + 1 \) th stage the \( m + 1 \) th patient is assigned to the treatment 1 with probability \( \frac{N_{m+1}}{m} \) \( \hat{\rho}_{m+1} \) and to the treatment 2 with probability \( 1 - \frac{N_{m+1}}{m} \) \( \hat{\rho}_{m+1} \) where \( \hat{\rho}_{m+1} = \frac{\hat{q}_{m+2}}{\hat{q}_{m+1} + \hat{q}_{m+2}} \) is the sample estimation of \( v_1 = q \frac{\alpha}{m} q_1 + q_2 \). The functions \( x \rho \) is called allocation rule. A large class of functions can be chosen as an allocation rule. If it is one of the following forms

\[
\begin{align*}
\text{For } \rho 
&= 0 \sqrt{\left( \frac{\exp(k) \frac{\rho}{x} - 1}{1 + 1 - 1 - \rho} \left( \frac{1 - \rho}{1 - x} \right) \right)^2 + 1} \end{align*}
\]

\[
\begin{align*}
\text{For } \rho 
&= 0 \sqrt{\left( \frac{\rho}{x} + 1 - 1 - \rho \left( \frac{1 - \rho}{1 - x} \right) \right)^2 + 1} \end{align*}
\]

\[
\begin{align*}
\text{For } \rho 
&= \left( \frac{\rho}{x} \right)^v + 1 - \rho \left( \frac{1 - \rho}{1 - x} \right) \end{align*}
\]

\[
\begin{align*}
\text{For } \rho 
&= \left( \frac{\rho}{x} \right)^v + 1 - \rho \left( \frac{1 - \rho}{1 - x} \right) \end{align*}
\]

where \( \alpha \geq 0 \) we have that

\[
\frac{N_{n+1}}{n} - v_1 = O\left( \sqrt{\frac{\log \log n}{n}} \right) \quad a.s. \quad \text{and} \quad \sqrt{n} \left( \frac{N_{n+1}}{n} - v_1 \right) \xrightarrow{D} 0 \sigma_{DAD}^2
\]

where

\[
\sigma_{DAD}^2 = \sigma_a^2 = \frac{q_1 q_2 p_1 + p_2}{q_1 + q_2} + \frac{2 q_1 q_2}{1 + 2 \alpha q_1 + q_2}
\]

It should be noted that the asymptotic normality holds for all \( 0 < p_1 < 1 \) and \( 0 < p_2 < 1 \). Hu et al. 2003. It is easily seen that \( \sigma_a^2 \) is a strictly monotonous decreasing function of \( \alpha \) \( \geq 0 \) \( \sigma_a^2 \rightarrow \sigma_{PW} \) as \( \alpha 

Also \( \sigma_a^2 < \sigma_{RPW}^2 \) for all \( \alpha > 1 \) whenever \( q_1 + q_2 > 1/2 \). Furthermore \( \sigma_a \) is more stable than \( \sigma_{RPW} \). So this adaptive design is more stable than the RPW rule. The larger \( \alpha \) is the more stable and less random is the design. So this design is a compromise between the stability in the PW rule and the randomization in the RPW rule. Even when \( \alpha \) is very large it keeps enough randomization to avoid determinism. Such design can keep the spirit of the RPW rule in that it assigns more patients to the better treatment and allows delayed responses by the patients.

In such kind of designs the assignments are adapted by both the results of responses and the current proportions of patients assigned and its original idea came from Efron’ 1971 biased coin design. So they are called doubly adaptive biased coin design first introduced by Eisele 1994 and Eisele et al. 1995 in the two-arm case with the responses which are from standard exponential families. However the results of Eisele 1994 and Eisele et al. 1995 are not applicable since their condition \( \rho \) is impossible to be satisfied because if \( q_1 \) and \( q_2 \) are small both \( v_1 \) and \( v_2 = 1 - v_1 \) should be very large by their \( \rho \) \( v \). But this condition is a key in their proving methods.

Hu et al. 2003 studied the general multi-arm case where the condition that \( \lambda < 1/2 \) in using GPU model is no longer a problem.
MARKOV CHAIN ADAPTIVE DESIGNS

In this section we propose another class of adaptive designs the so-called Markov chain adaptive design. Consider the two-arm case first. Suppose that at the stage \( n \) the treatment 1 is assigned to the \( n \)th patient. Then \( n + 1 \)th patient will be assigned either treatment 1 or treatment 2 according certain probabilities which depend on the response on the \( n \)th patient. Let \( \alpha \), be the probability of assigning the \( n + 1 \)th patient to treatment 1 when the response of the \( n \)th patient to treatment 1 is "success" and let \( \alpha_f \) be the probability of assigning the \( n + 1 \)th patient to treatment 1 when the response of the \( n \)th patient to treatment 1 is "failure". Similarly define \( \beta \) and \( \beta_f \) with treatment 2 instead of treatment 1 in the definitions of \( \alpha \) and \( \alpha_f \). When \( \alpha_\beta = 1 \) and \( \alpha_f = 0 \) \( \beta_\beta = 1 \) and \( \beta_f = 0 \) we get Zelen’s PW rule. We may choose the parameters \( \alpha_\beta, \alpha_f, \beta_\beta, \beta_f \) for different goals.

Define \( X_n = 1 \) or 0 if \( n \)th patient is assigned to the treatment 1 or 2 respectively. Then \( N_{n1} = \sum_{i=1}^{n} X_i \) is the number of patients assigned to treatment 1 in the first \( n \) stages. Let \( p_{\alpha} \) \( n = P \) success \( | \ X_n = 1 \) and \( p_{\beta} \) \( n = P \) success \( | \ X_n = 0 \) and \( \alpha_\beta = p_{\alpha} \) \( n \alpha_\beta + 1 - p_{\alpha} \) \( n \alpha_\beta + \beta_\beta - 1 + p_{\beta} \) \( n \beta_\beta + \beta_f \). Then \( X_n \) is a Markov chain with the transition probability matrix

\[
P_n = \begin{pmatrix} \alpha_n & 1 - \beta_n \\ 1 - \alpha_n & \beta_n \end{pmatrix}
\]

When \( p_{\alpha} \) \( n = p_1 \) and \( p_{\beta} \) \( n = p_2 \) for all \( n \) \( X_n \) is homogeneous. For this case we can write \( \alpha_\beta = \alpha_\beta \) and \( \beta_\beta = \beta_\beta \) and we have that

\[
N_{n1} / n \overset{p}{\rightarrow} \mu \quad \text{and} \quad \sqrt{n} \left( N_{n1} / n - \mu \right) \overset{d}{\rightarrow} 0 \sigma^2 \quad 10
\]

where

\[
\mu = \frac{1 - \beta - \beta_\beta \sigma^2}{2 - \alpha - \beta \beta_\beta} = \frac{1 - \alpha_\beta 1 - \beta_\beta \alpha + \beta_\beta}{2 - \alpha - \beta_\beta^2}
\]

\[
N_{n1} / n \overset{p}{\rightarrow} \mu \quad \text{and} \quad \sqrt{n} \left( N_{n1} / n - \mu \right) \overset{d}{\rightarrow} 0 \sigma^2 \quad 10
\]

Bai et al. 2001. For non-homogeneous case we have similar results under the condition of type Eq. 10. Furthermore we established the strong consistency and strong approximation of

\[
N_{n1} / n \overset{p}{\rightarrow} \mu \quad \text{and} \quad \sqrt{n} \left( N_{n1} / n - \mu \right) \overset{d}{\rightarrow} 0 \sigma^2 \quad 10
\]

where \( \square \) is a standard Brownian motion \( \square \) Lin et al. 2001. For the multi-arm case we let \( X_n = X_1 \) \( \ldots X_n \) where \( X_n = 1 \) if the \( n \)th patient is assigned to treatment \( i \) and let \( N_n = \sum_{k=1}^{n} X_k \) be the number of patients assigned to each treatment at the first \( n \) stages as usual. Let \( p_{\alpha} \) \( n = P \) success \( | \ X_n = 1 \). Assume the transition probability matrix of the Markov chain \( X_n \) is \( H_n = \sum_{i=1}^{K} H_{ii} \) where \( K \) is a function of \( p_{\alpha} \) \( n = P \) \( i = 1 \ldots K \). Let \( H_n = X_{n+1} \) \( X_n \) \( = 0 \). We showed that \( N_n \) can be approximated by a multi-dimensional Brownian motion

\[
N_n = n \nu - W_n \square = \sum_{k=1}^{K} \sum_{i=1}^{K} H_k \square - \nu \square n^{1/2} \quad \nu \square n^{1/2} \quad 11
\]

Condition \( \square 11 \) is usually satisfied if \( \sum_{i=1}^{K} \left| H_{ii} \right| \overset{p}{\rightarrow} 1 \). And \( n \square n^{1/2} \) \( i = 1, \ldots, K \) is implied by \( \square 9 \).

If \( H_n \) depends on estimated unknown parameters the problem becomes more complicated. Zhang studied such case. Remark 4. In Section 2 we mentioned the normality in Eqs. 1 and 2 but the condition \( p_1 + p_2 < 3/2 \) was imposed. But for normality in Eq. 10 this restriction is cancelled. In the multi-arm case the restriction that \( \lambda < 1/2 \) is also cancelled.
Remark 5. Keeping $\mu$, we can choose $\alpha$ and $\beta$ to minimize $\sigma^2$ so that we get more stable Markov chain adaptive designs. Also if $\alpha$ and $\beta$ was chosen suitably the Markov chain adaptive designs can also keep both randomization and stability.

References

International Peer Review System for Scientific Papers Published in the English-language Journal of Zhejiang University SCIENCE

The characteristics and requirements of world-class journal were considered in establishing a rigid peer review system for scientific papers submitted for publication in the English-language Journal of Zhejiang University SCIENCE from 2002 onward. We sent the over 408 contributions received from January to December in 2002 to the U.S.A. Ireland France Canada Australia Austria Germany New Zealand the Netherlands Finland Poland Portugal Italy Israel Spain Belgium Sweden Switzerland Denmark Japan Singapore Slovak India Greece Czech Mexico Hong Kong Macao Taiwan etc. for pre-publication review by top-notch international scientists in their respective specialties. Experience in scientific papers publication has shown that an international peer review system plays an important part in ensuring the high quality of a journal’s contents and helping it to be known worldwide.

Every little improvement of this journal depends on the strong support from the reviewers. We take this opportunity to give our heartfelt thanks to reviewers in China and abroad for their help since the establishment of the journal.