Stochastic Processes and their Applications 80 (1999) 87–101

Asymptotic theorems for urn models with nonhomogeneous generating matrices

Z.D. Bai *, Feifang Hu

Department of Statistics and Applied Probability, National University of Singapore, Singapore 119260, Singapore

Received 11 February 1998; received in revised form 29 October 1998; accepted 30 October 1998

Abstract

The generalized Friedman's urn (GFU) model has been extensively applied to biostatistics. However, in the literature, all the asymptotic results concerning the GFU are established under the assumption of a homogeneous generating matrix, whereas, in practical applications, the generating matrices are often nonhomogeneous. On the other hand, even for the homogeneous case, the generating matrix is assumed in the literature to have a diagonal Jordan form and satisfies $\lambda > 2\,\mathrm{Re}(\lambda_1)$, where λ and λ_1 are the largest eigenvalue and the eigenvalue of the second largest real part of the generating matrix (see Smythe, 1996, Stochastic Process. Appl. 65, 115–137). In this paper, we study the asymptotic properties of the GFU model associated with nonhomogeneous generating matrices. The results are applicable to a variety of settings, such as the adaptive allocation rules with time trends in clinical trials and those with covariates. These results also apply to the case of a homogeneous generating matrix with a general Jordan form as well as the case where $\lambda = 2\,\mathrm{Re}(\lambda_1)$. © 1999 Elsevier Science B.V. All rights reserved.

AMS classification: primary 62E20; 62L05; secondary 62F12

Keywords: Adaptive designs; Asymptotic normality; Consistency; Generalized Friedman's urn model; Non-homogeneous generating matrix

1. Introduction

Adaptive designs have often been proposed as a way sequentially to assign more patients to better treatments, based on outcomes of previous treatments in clinical trials. A very important class of adaptive designs is one based on the generalized Friedman's urn (GFU) model (see Athreya and Karlin (1968); GFU is also named as generalized Pólya urn (GPU) in the literature), which has application in clinical trials, bioassay and psychophysics. References are made to Wei (1979), Rosenberger et al. (1997) and

* Corresponding author.

PII: S0304-4149(98)00094-5

Rosenberger and Grill (1997). A general review on this subject with respect to clinical trials is given in Rosenberger (1996).

Adaptive designs using the GFU model can be formulated as follows. Assume, at the beginning, an urn contains particles of K distinct types, denoted by $\boldsymbol{Y}_0 = (Y_{01}, \ldots, Y_{0K})$, respectively representing K 'treatments' in a clinical trial, where Y_{0k} denotes the number of particles of type k, $k = 1, \ldots, K$. These treatments are to be sequentially allocated in n consecutive stages. At stage i, $i = 1, \ldots, n$ a particle is drawn from the urn with replacement. If a type k particle is drawn at the ith stage, then the treatment k is assigned to the patient i, $k = 1, \ldots, K$, $i = 1, \ldots, n$. Let $\xi(i)$ denote a random variable associated with the ith stage of the clinical trial, which may include measurements on the ith patient and the outcome of the treatment at the ith stage. Afterwards, additional $D_{k,q}(i)$ particles of type q are added to the urn, $q = 1, \ldots, K$, where $D_{k,q}(i)$ is a function of $\xi(i)$. This procedure is repeated to the nth stage. After n splits and generations, the composition of the urn is denoted by the vector $\boldsymbol{Y}_n = (Y_{n1}, \ldots, Y_{nk})$, where Y_{nk} represents the number of type k particles in the urn. Furthermore, we define $\boldsymbol{D}_i = \langle\langle D_{k,q}(i),\ k, q = 1, \ldots, K\rangle\rangle$ and $\boldsymbol{H}_i = \langle\langle E(D_{k,q}(i)),\ k, q = 1, \ldots, K\rangle\rangle$, $i = 1, \ldots, n$. The matrix $\boldsymbol{D}_i$'s are called *rules* and $\boldsymbol{H}_i$'s are the *generating matrices*.

We call the GFU model *homogeneous* if $\boldsymbol{H}_i = \boldsymbol{H}$ for all $i = 1, \ldots, n$. For a homogeneous GFU model, under the assumptions (i) $\Pr\{D_{k,q} = 0,\ q = 1, \ldots, K\} = 0$ for every $k = 1, \ldots, K$ and (ii) $\boldsymbol{H}$ is positive regular, Athreya and Karlin (1968) and Athreya and Ney (1972) show that

$$\frac{N_{nk}}{n} \to v_k \quad \text{and} \quad \frac{Y_{nk}}{\sum_{q=1}^{K} Y_{nq}} \to v_k \tag{1.1}$$

almost surely as $n \to \infty$, where $\boldsymbol{v} = (v_1, \ldots, v_K)$ is the left eigenvector of the largest eigenvalue λ of $\boldsymbol{H}$. Let λ_1 denote the eigenvalue of the second largest real part, with corresponding right eigenvector $\boldsymbol{\eta}$. Furthermore, under the additional assumption that $\lambda > 2\,\mathrm{Re}(\lambda_1)$, Athreya and Karlin (1968) show that

$$n^{-1/2}\boldsymbol{Y}_n\boldsymbol{\eta}' \to \mathrm{N}(0, \sigma^2), \tag{1.2}$$

where σ^2 is a constant. When $\lambda = 2\,\mathrm{Re}(\lambda_1)$ and λ_1 is a simple eigenvalue, then (1.2) holds with the normalization constant $\sqrt{n}$ replaced by of $\sqrt{n \ln n}$.

Smythe (1996) defines the Extended Pólya Urn model (EPU) as a GFU with $\sum_{q=1}^{K} E(D_{k,q}) = c > 0$, $k = 1, \ldots, K$, namely, adding an expected constant total number of balls at each stage. For the EPU, Smythe (1996) established the asymptotic normality of $\boldsymbol{Y}_n$ and $\boldsymbol{N}_n$ under the assumptions: (i) for each nonprincipal eigenvalue λ_j, $\lambda > 2\,\mathrm{Re}(\lambda_j)$; (ii) all eigenvalues are simple, and no two distinct complex eigenvalues have the same real part, except for conjugate pairs; and (iii) the eigenvectors are linearly independent, where $\boldsymbol{N}_n = (N_{n1}, \ldots, N_{nK})$ and N_{nK} is the number of times a type k particle drawn in the first n trials.

In this paper, the asymptotic properties of the urn composition $\boldsymbol{Y}_n = (Y_{n1}, \ldots, Y_{nK})$ are investigated for EPUs with nonhomogeneous generating matrices $\{\boldsymbol{H}_i\}$. Throughout this paper, we assume $\sum_{q=1}^{K} D_{kq}(i) = c > 0$, for all $k = 1, \ldots, K$ and $i = 1, \ldots, n$, i.e., adding a total number of balls at each stage. We also assume that there exists a positive regular matrix $\boldsymbol{H}$ such that

$$\sum_{i=1}^{\infty} \frac{\alpha_i}{i} < \infty, \tag{1.3}$$

where $\alpha_i = \|\boldsymbol{H}_i - \boldsymbol{H}\|_\infty$. Theorems 2.1 and 2.2 show that

$$\left(\sum_{i=1}^{K} Y_{ni}\right)^{-1} (Y_{n1}, \ldots, Y_{nK})^{\mathrm{T}}$$

converges to the left eigenvector corresponding to the maximal eigenvalue of $\boldsymbol{H}$ in probability. Further we show the asymptotic normality of $\boldsymbol{Y}_n$ in Theorem 2.3.

To illustrate that the theory of nonhomogeneous EPUs has wider applications than that for the homogeneous GFU, we present the following three examples.

Example 1 (*Adaptive allocation rules with time trends in clinical trial*). The homogeneous generating matrix $\boldsymbol{H}$ is often not met in clinical trials, when patients may exhibit a drift in characteristics over time. Several examples are given in Altman and Royston (1988). Coad (1991, 1992), studied prestratification and poststratification techniques to deal with time trends associated with the adaptive allocation rules in clinical trials.

Assume that the subject response is dichotomous, e.g., $T_i = 1$ if the response of the subject i is a success and 0 otherwise. Define $p_{ik} = P(T_i = 1 \mid X_i = k)$, for $i = 1, \ldots, n$ and $k = 1, \ldots, K$, where $X_i = k$ indicates that a type k particle is drawn at the ith stage. The case $p_{ik} = p_k$, where the probabilities of success are homogeneous within treatment groups, is well studied in the literature. But this assumption is not always realistic over the course of recruitment. There may be a drift in patient characteristics over time, for example, $\lim_{i\to\infty} p_{ik} = p_k$. In this case, the generating matrices are heterogeneous. As pointed out in Rosenberger (1996), it is usually mathematically difficult to consider the case that the probabilities of success are not homogeneous within treatment groups.

Example 2 (*Adaptive allocation rules associated with covariates*). In the previous example, the probability of success may depend on some observable covariates on the patients, that is $p_{ik} = p_k(\xi_i)$, where ξ_i are covariates observed on the patient i and the result of the treatment at the ith stage. Thus, the kth row of the allocation rule $\boldsymbol{D}_i$ at the ith stage is a function of outcomes of the ith patient when the kth treatment is taken. Thus, the corresponding generating matrix $\boldsymbol{H}_i$ depends on i. In general, we assume that $\boldsymbol{H}_i = \boldsymbol{H}_i(\xi_i)$ or $\boldsymbol{H}_i = \boldsymbol{H}_i(\xi_1, \ldots, \xi_i)$ where $\xi_1, \ldots, \xi_n$ are independent random covariates.

Example 3 (*Homogeneous generating matrix with a general Jordan form*). Smythe (1996) shows the asymptotic normality under the assumption that the homogeneous generating matrix has a diagonal Jordan form. The results in this paper apply to a general Jordan form of $\boldsymbol{H}$. This is because the Assumption 2.1 is trivial here ($\boldsymbol{H}_i = \boldsymbol{H}$). Also Smythe (1996) only considers the case that $\lambda > 2\,\mathrm{Re}(\lambda_1)$, where λ_1 is the eigenvalue of the second largest real part of $\boldsymbol{H}$. Theorems 2.1–2.3 also apply to the case that $\lambda = 2\,\mathrm{Re}(\lambda_1)$.

2. The main results

Suppose that there is a sequence of increasing σ-fields $\{\mathscr{F}_n\}$ and that $\boldsymbol{Y}_n$ is a sequence of random k-vectors of non-negative elements which are adapted to $\{\mathscr{F}_n\}$ and satisfy

$$E(\boldsymbol{Y}_i|\mathscr{F}_{i-1}) = \boldsymbol{Y}_{i-1}\boldsymbol{M}_i, \tag{2.1}$$

where $\boldsymbol{M}_i$ is a matrix that is $\mathscr{F}_{i-1}$-measurable. Then, $\{\boldsymbol{Q}_i = \boldsymbol{Y}_i - E(\boldsymbol{Y}_i|\mathscr{F}_{i-1})\}$ is a sequence of K-dimensional martingale differences with respect to $\{\mathscr{F}_n\}$.

In application to the GFU model, we have the following recursive relation:

$$\boldsymbol{Y}_n = \boldsymbol{Y}_{n-1} + \boldsymbol{X}_n\boldsymbol{D}_n,$$

where $\boldsymbol{D}_n$ is the allocation rule at the stage n and $\boldsymbol{X}_n$ is the result of the nth draw, distributed according to the urn composition at the previous stages, i.e., if $X_n = k$, then the kth component of $\boldsymbol{X}_n$ is 1 and other components are 0. Usually, we assume that $\boldsymbol{D}_n$ is independent of observations at previous stages and

$$E(\boldsymbol{X}_n|\boldsymbol{Y}_{n-1}) = \boldsymbol{Y}_{n-1}/(nc+\beta),$$

where $\beta = \sum_{k=1}^{K} Y_{0k} - c > -c$ and c is the number of particles added to urn at each stage. For brevity, without loss of generality, we usually assume that $c=1$ and $\beta = 0$ in studying the asymptotic properties of the GFUs. In this case, by denoting $\boldsymbol{H}_i = E(\boldsymbol{D}_i)$, in Eq. (2.1), we have $\boldsymbol{M}_i = \boldsymbol{I} + i^{-1}\boldsymbol{H}_i$.

From Eq. (2.1), it is easy to see that

$$E(\boldsymbol{Y}_n) = \boldsymbol{Y}_0\boldsymbol{M}_1\boldsymbol{M}_2\cdots\boldsymbol{M}_n. \tag{2.2}$$

Then, we have

$$\boldsymbol{Y}_n - E(\boldsymbol{Y}_n) = \sum_{i=1}^{n}[E(\boldsymbol{Y}_n|\mathscr{F}_i) - E(\boldsymbol{Y}_n|\mathscr{F}_{i-1})] = \sum_{i=1}^{n}\boldsymbol{Q}_i\boldsymbol{B}_{n,i}, \tag{2.3}$$

where $\boldsymbol{B}_{n,i} = \boldsymbol{M}_{i+1}\cdots\boldsymbol{M}_n$ with the convention that $\boldsymbol{B}_{n,n} = \boldsymbol{I}$ and $\mathscr{F}_0$ denotes the trivial σ-field.

Without loss of generality, we assume $\beta = 0$ in the following discussion. For investigating the asymptotics of $\boldsymbol{Y}_n$ with application in the urn model, we need the following assumptions.

Assumption 2.1. *Suppose that there is a $K\times K$ matrix $\boldsymbol{H}$ of non-negative entries and that $\boldsymbol{H}$ has the Jordan form decomposition*

$$\boldsymbol{T}^{-1}\boldsymbol{H}\boldsymbol{T} = \boldsymbol{J} = \begin{pmatrix} \lambda & 0 & \dots & 0 \\ 0 & \boldsymbol{J}_1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \boldsymbol{J}_s \end{pmatrix} \quad \textit{with } \boldsymbol{J}_t = \begin{pmatrix} \lambda_t & 1 & 0 & \dots & 0 \\ 0 & \lambda_t & 1 & \dots & 0 \\ \vdots & \dots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_t & 1 \\ 0 & 0 & 0 & \dots & \lambda_t \end{pmatrix},$$

where λ is the unique maximum eigenvalue of $\boldsymbol{H}$. Denote the order of $\boldsymbol{J}_t$ by v_t and $\tau = \max\{0, \mathrm{Re}(\lambda_1), \dots, \mathrm{Re}(\lambda_s)\}$. We define $v = \max\{v_t\colon \mathrm{Re}(\lambda_t) = \tau\}$. Assume that $\|\boldsymbol{H}_i - \boldsymbol{H}\|_\infty = \alpha_i$ satisfies

$$\sum_{i=1}^{\infty} \frac{\alpha_i}{i} < \infty. \tag{2.4}$$

Moreover, we assume that the elements of the left eigenvector $\boldsymbol{v} = (v_1, \dots, v_p)^{\mathrm{T}}$ associated with the positive maximal eigenvalue λ are nonnegative and satisfy $\sum_{i=1}^{p} v_i = 1$.

Assumption 2.2. *Suppose that there is a constant C_1 such that $0 \leqslant \|\boldsymbol{Y}_i\| \leqslant C_1 i^{\lambda}$.*

Theorem 2.1. *Under Assumptions* 2.1 *and* 2.2, *$n^{-\lambda}(E\boldsymbol{Y}_n)$ tends to a constant vector of non-negative entries. Furthermore, this constant vector is the left eigenvector $\boldsymbol{v}$ of the matrix $\boldsymbol{H}$ corresponding to λ.*

Proof. Define $\boldsymbol{z}_n = n^{-\lambda} E(\boldsymbol{Y}_n)\boldsymbol{T}$. Then by Eq. (2.1), for any $n > k > 1$, we have

$$\begin{aligned} \boldsymbol{z}_n &= (k/n)^{\lambda} \boldsymbol{z}_k (\boldsymbol{I} + (k+1)^{-1}\boldsymbol{J}) \cdots (\boldsymbol{I} + n^{-1}\boldsymbol{J}) \\ &\quad + \sum_{j=k+1}^{n-1} \frac{(j+1)^{\lambda-1}}{n^{\lambda}} \boldsymbol{z}_j \boldsymbol{W}_{j+1} (\boldsymbol{I} + (j+2)^{-1}\boldsymbol{J}) \cdots (\boldsymbol{I} + n^{-1}\boldsymbol{J}), \end{aligned} \tag{2.5}$$

where $\boldsymbol{W}_j = \boldsymbol{T}^{-1}(\boldsymbol{H}_j - \boldsymbol{H})\boldsymbol{T}$. We consider the elements of the matrix

$$\begin{aligned} &(\boldsymbol{I} + (j+2)^{-1}\boldsymbol{J}) \cdots (\boldsymbol{I} + n^{-1}\boldsymbol{J}) \\ &= \begin{pmatrix} \prod_{i=j+2}^{n} (1 + \frac{\lambda}{i}) & 0 & \dots & 0 \\ 0 & \prod_{i=j+2}^{n} (\boldsymbol{I} + i^{-1}\boldsymbol{J}_1) & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \prod_{i=j+2}^{n} (\boldsymbol{I} + i^{-1}\boldsymbol{J}_s) \end{pmatrix}. \end{aligned}$$

By elementary calculus, one finds that, as $n > j \to \infty$,

$$(1 + \lambda/n) \cdots (1 + \lambda/(j+2)) = (n/j)^{\lambda}(1 + \mathrm{o}(1)),$$

and the $(h, h+i)$-element of the block matrix $\prod_{i=j+2}^{n} (\boldsymbol{I} + i^{-1}\boldsymbol{J}_t)(j/n)^{\lambda}$ has the estimation

$$\frac{1}{i!}(j/n)^{\lambda - \mathrm{Re}(\lambda_t)} \log^i(n/j)(1 + \mathrm{o}(1)) \leqslant \frac{3}{i!} \left(\frac{i}{e(\lambda - |\lambda_t| - \varepsilon)} \right)^i (j/n)^{\varepsilon}, \tag{2.6}$$

where λ_t is the eigenvalues of $\boldsymbol{J}_t$ and $0 < \varepsilon < \lambda - |\lambda_t|$. These imply that

$$(\boldsymbol{I} + (k+1)^{-1}\boldsymbol{J}) \cdots (\boldsymbol{I} + n^{-1}\boldsymbol{J})(k/n)^{\lambda} \to \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \dots & \dots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix} = \boldsymbol{e}_1' \boldsymbol{e}_1, \tag{2.7}$$

where $\boldsymbol{e}_1=(1,0,\ldots,0)$, and for some constant $C_2>0$, by the fact that z_j is bounded,

$$\left|\sum_{j=k+1}^{n-1}\frac{(j+1)^{\lambda-1}}{n^{\lambda}}z_j\boldsymbol{W}_{j+1}(\boldsymbol{I}+(j+2)^{-1}\boldsymbol{J})\cdots(\boldsymbol{I}+n^{-1}\boldsymbol{J})\right|\leqslant C_2\sum_{j=k+1}^{\infty}\frac{\alpha_j}{j}\to 0. \tag{2.8}$$

Consequently, by Eq. (2.5), for any sequences $n>k=k_n\to\infty$, $z_n-z_k\boldsymbol{e}_1'\boldsymbol{e}_1\to 0$. Since z_n is bounded, we conclude that z_n must converge to a limit, say z, satisfying $z=z\boldsymbol{e}_1'\boldsymbol{e}_1$. This implies that $z=\boldsymbol{e}_1$. Then it follows that $n^{-\lambda}E\boldsymbol{Y}_n$ converges to the limit $\boldsymbol{e}_1\boldsymbol{T}^{-1}=\boldsymbol{v}$. The proof of the theorem is complete. □

Assumption 2.3. *For all i, $\boldsymbol{u}'$ is a common right eigenvector of $\boldsymbol{H}_i$ corresponding to the largest eigenvalue and such that $\boldsymbol{Y}_i\boldsymbol{u}'$ is non-random, where $\boldsymbol{u}'$ is a right-eigenvector of $\boldsymbol{H}$ corresponding to λ. Further, we assume that $\|\boldsymbol{R}_i\|\leqslant C_2 i^{2\Delta-1-\varepsilon_1}$ for some constants $\max\{1/2,\tau\}<\Delta\leqslant\lambda$, $C_2>0$ and $\varepsilon_1>0$ where $\boldsymbol{R}_i=E(\boldsymbol{Q}_i^*\boldsymbol{Q}_i)$ and $\boldsymbol{Q}_i^*$ denotes the complex conjugate transpose of $\boldsymbol{Q}_i$.*

Note that Assumption 2.3 implies $\boldsymbol{Q}_i\boldsymbol{u}'=0$ and $\boldsymbol{R}_i\boldsymbol{u}'=0$. Also, under the condition $\sum_{q=1}^{K}D_{kq}(i)=c>0$ mentioned in the introduction, we can take $\boldsymbol{u}=\boldsymbol{1}=(1,\ldots,1)$.

Theorem 2.2. *Under Assumptions 2.1–2.3, $n^{-\Delta}(\boldsymbol{Y}_n-E\boldsymbol{Y}_n)\to 0$, in probability. Especially, when $\Delta=\lambda$, $n^{-\lambda}\boldsymbol{Y}_n$ converges in probability to the same limit of $n^{-\lambda}E\boldsymbol{Y}_n$, as $n\to\infty$.*

Proof. Since $E(\boldsymbol{Q}_i|\mathscr{F}_{i-1})=0$, we have

$$\mathrm{Var}(\boldsymbol{Y}_n)=\sum_{i=1}^{n}\boldsymbol{B}_{n,i}^*\boldsymbol{R}_i\boldsymbol{B}_{n,i}, \tag{2.9}$$

or equivalently,

$$\boldsymbol{T}^*\mathrm{Var}(\boldsymbol{Y}_n)\boldsymbol{T}=\sum_{i=1}^{n}(\boldsymbol{B}_{n,i1}+\boldsymbol{B}_{n,i2})^*\widetilde{\boldsymbol{R}}_i(\boldsymbol{B}_{n,i1}+\boldsymbol{B}_{n,i2}),$$

where $\widetilde{\boldsymbol{R}}_i=\boldsymbol{T}^*\boldsymbol{R}_i\boldsymbol{T}$ and

$$\boldsymbol{T}^{-1}\boldsymbol{B}_{n,i}\boldsymbol{T}=\boldsymbol{B}_{n,i1}+\boldsymbol{B}_{n,i2}, \tag{2.10}$$

and $\boldsymbol{B}_{n,i1}=(\boldsymbol{I}+(i+1)^{-1}\boldsymbol{J})\cdots(\boldsymbol{I}+n^{-1}\boldsymbol{J})$ and $\boldsymbol{B}_{n,i2}=\boldsymbol{B}_{n,i}-\boldsymbol{B}_{n,i1}$.

By Assumption 2.3, the elements of the first column of $\boldsymbol{W}_j$, except the first, are all 0, whereas the elements of the first column and row of $\widetilde{\boldsymbol{R}}_j$ are 0. Substituting $\boldsymbol{T}^{-1}\boldsymbol{H}_j\boldsymbol{T}=\boldsymbol{J}+\boldsymbol{W}_j$ into the expression of $\boldsymbol{B}_{n,i2}$, we find that the general term of $\boldsymbol{B}_{n,i2}$ is a product whose ℓth ($\ell=i+1,\ldots,n$) factor is $(\boldsymbol{I}+\ell^{-1}\boldsymbol{J})$ or $\ell^{-1}\boldsymbol{W}_\ell$, subject to the restriction that there is at least one factor of the form $\ell^{-1}\boldsymbol{W}_\ell$. Therefore, when evaluating $\boldsymbol{B}_{n,i2}^*\widetilde{\boldsymbol{R}}_i\boldsymbol{B}_{n,i2}$, we may change the matrices $\boldsymbol{J}$ and $\boldsymbol{W}_j$ as $\boldsymbol{J}^{(0)}$ and $\boldsymbol{W}_j^{(0)}$, which are obtained by replacing the first rows of $\boldsymbol{J}$ and $\boldsymbol{W}_j$ as zero's. Then, we have

$$\|\boldsymbol{B}_{n,i2}^{*}\widetilde{\boldsymbol{R}}_i\boldsymbol{B}_{n,i2}\| \leqslant \left(\sum_{\ell=1}^{n-i}\sum_{i<j_1<\cdots<j_\ell\leqslant n}(n/i)^{\tau}\frac{\alpha_{j_1}}{j_1}\cdots\frac{\alpha_{j_\ell}}{j_\ell}\right)^2\|\widetilde{\boldsymbol{R}}_i\|$$

$$\leqslant C_2 n^{2\tau} i^{2\Delta-1-\varepsilon_1-2\tau}\left[\exp\left(\sum_{j=i+1}^{\infty}\frac{\alpha_j}{j}\right)-1\right]^2\log^{v}(n/i). \tag{2.11}$$

Therefore,

$$n^{-2\Delta}\left\|\sum_{i=1}^{n}\boldsymbol{B}_{n,i2}^{*}\boldsymbol{R}_i\boldsymbol{B}_{n,i2}\right\|$$

$$\leqslant C_2 n^{-1}\sum_{i=1}^{n}(i/n)^{2\Delta-1-2\tau}\left[\exp\left(\sum_{j=i+1}^{\infty}\frac{\alpha_j}{j}\right)-1\right]^2\log^{v}(n/i)\to 0, \tag{2.12}$$

where, applying Toeplitz Lemma (see Loève, 1984, p. 250), the last limit follows from the facts that $[\exp(\sum_{j=i+1}^{\infty}\alpha_j/j)-1]^2\to 0$ as $i\to\infty$ and

$$n^{-1}\sum_{i=1}^{n}(i/n)^{2\Delta-1-2\tau}\log^{v}(n/i)\to\int_0^1 u^{2\Delta-1-2\tau}\log^{v}(1/u)\,\mathrm{d}u.$$

Similarly, we have

$$n^{-2\Delta}\left\|\sum_{i=1}^{n}\boldsymbol{B}_{n,i1}^{*}\widetilde{\boldsymbol{R}}_i\boldsymbol{B}_{n,i1}\right\|\leqslant C_2 n^{-1-\varepsilon_1}\sum_{i=1}^{n}(i/n)^{2\Delta-1-\varepsilon_1-2\tau}\log^{v}(n/i)\to 0. \tag{2.13}$$

Then, Eqs. (2.9), (2.12) and (2.13) imply

$$\mathrm{Var}(\boldsymbol{Y}_n)\to 0.$$

The proof of the theorem is complete. □

Remark 2.1. From the proof of Theorem 2.2, one sees that if $\Delta\geqslant\max(1/2,\tau)$ and $\|\boldsymbol{R}_i\|\leqslant C_3 i^{2\Delta-1}$, then

$$b_n^{-1}\sum_{i=1}^{n}\boldsymbol{B}_{n,i2}^{*}\widetilde{\boldsymbol{R}}_i\boldsymbol{B}_{n,i2}\to 0, \tag{2.14}$$

where $b_n=n^{2\Delta}$ or $=n^{2\tau}\log^{2v-1}n$ according to $\Delta>\tau$ or $=\tau$, respectively.

Assumption 2.4. *Assume that* $E\|E(\boldsymbol{Q}_i^{*}\boldsymbol{Q}_i|\mathscr{F}_{i-1})-\boldsymbol{R}_i\|\to 0$ *and* $\boldsymbol{R}_i\to\boldsymbol{R}$ *as* $i\to\infty$.

In application to the GFU model, we have $\boldsymbol{Q}_i=\boldsymbol{X}_i\boldsymbol{D}_i-n^{-1}\boldsymbol{Y}_{n-1}\boldsymbol{H}_i$. In the case where $\boldsymbol{X}_i$ is conditionally distributed as Multinomial $(1,\boldsymbol{Y}_{i-1}/i)$ given $\boldsymbol{Y}_{i-1}$ and $\boldsymbol{D}_i$ is independent of $\boldsymbol{Y}_{i-1}$ and $\boldsymbol{X}_i$, we have

$$E(\boldsymbol{Q}_i^*\boldsymbol{Q}_i|\mathscr{F}_{i-1}) = E[(\boldsymbol{D}_i-\boldsymbol{H}_i)^*\boldsymbol{X}_i^*\boldsymbol{X}_i(\boldsymbol{D}_i-\boldsymbol{H}_i)|\mathscr{F}_{i-1})$$
$$+\boldsymbol{H}_i^*E(\boldsymbol{X}_i-i^{-1}\boldsymbol{Y}_{i-1})^*(\boldsymbol{X}_i-i^{-1}\boldsymbol{Y}_{i-1})|\mathscr{F}_{i-1})\boldsymbol{H}_i.$$

Note that

$$E(\boldsymbol{X}_i^*\boldsymbol{X}_i|\mathscr{F}_{i-1}) = \mathrm{diag}(i^{-1}\boldsymbol{Y}_{i-1}) \to \mathrm{diag}(\boldsymbol{v})$$

and

$$E(\boldsymbol{X}_i-i^{-1}\boldsymbol{Y}_{i-1})^*(\boldsymbol{X}_i-i^{-1}\boldsymbol{Y}_{i-1})|\mathscr{F}_{i-1})$$
$$=\mathrm{diag}(i^{-1}\boldsymbol{Y}_{i-1})-(i^{-1}\boldsymbol{Y}_{i-1}')(i^{-1}\boldsymbol{Y}_{i-1}) \to \mathrm{diag}(\boldsymbol{v})-\boldsymbol{v}^*\boldsymbol{v}.$$

Therefore, Assumption 2.4 holds with

$$\boldsymbol{R}_i-\sum_{j=1}^{K} v_j \begin{pmatrix} cov(D_{j1}^{(i)},D_{j1}^{(i)}) & \cdots & cov(D_{j1}^{(i)},D_{jK}^{(i)}) \\ \vdots & \cdots & \vdots \\ cov(D_{jK}^{(i)},D_{j1}^{(i)}) & \cdots & cov(D_{jK}^{(i)},D_{jK}^{(i)}) \end{pmatrix} + \boldsymbol{H}_i^*(\mathrm{diag}(\boldsymbol{v})-\boldsymbol{v}^*\boldsymbol{v})\boldsymbol{H}_i \to 0,$$

where $D_{\ell,j}^{(i)}$ is the (ℓ,j)th element of $\boldsymbol{D}_i$.

Remark 2.2. Assumption 2.1 is true if the non-homogeneous generating matrices $\boldsymbol{H}_i$ converges to a generating matrix $\boldsymbol{H}$ with a rate of $\log^{-1-c} i$ for some $c>0$. The Assumptions 2.2–2.4 are easy to verify and to be satisfied for most matrices involved in adaptive designs. In Section 3, we shall give three different examples in which these assumptions are satisfied.

Theorem 2.3. *Under the Assumptions* 2.1–2.4 *and assuming* $\tau\leqslant\frac{1}{2}$, *then* $V_n^{-1}(\boldsymbol{Y}_n-E\boldsymbol{Y}_n)$ *is asymptotically normal with mean vector* 0 *and variance–covariance matrix* Σ, *where* Σ *is specified later*, $V_n^2=n$ *if* $\tau<1/2$, *and* $V_n^2=n\log^{2\nu-1} n$ *if* $\tau=1/2$. *Here* τ *is defined in Assumption* 2.1.

Proof. We first get a simple approximation of $\boldsymbol{Y}_n-E(\boldsymbol{Y}_n)$. By Eq. (2.3), we have

$$(\boldsymbol{Y}_n-E(\boldsymbol{Y}_n))\boldsymbol{T}=\sum_{i=1}^{n}\widetilde{\boldsymbol{Q}}_i\widetilde{\boldsymbol{B}}_{ni}, \tag{2.15}$$

where $\widetilde{\boldsymbol{Q}}_i=\boldsymbol{Q}_i\boldsymbol{T}$. Substituting Eq. (2.10) into Eq. (2.15), it follows that

$$(\boldsymbol{Y}_n-E(\boldsymbol{Y}_n))\boldsymbol{T}=\sum_{i=1}^{n}\widetilde{\boldsymbol{Q}}_i\boldsymbol{B}_{n,i1}+\sum_{i=1}^{n}\widetilde{\boldsymbol{Q}}_i\boldsymbol{B}_{n,i2}=\boldsymbol{U}_1+\boldsymbol{U}_2. \tag{2.16}$$

By Remark 2.1 with $\Delta=1/2$ and noting that $\mathrm{Var}(V_n^{-1}\boldsymbol{U}_1)$ equals the left hand side of (2.14), it follows that $V_n^{-1}\boldsymbol{U}_2\to 0$ in prob.

Thus, the limiting distribution of $V_n^{-1}(\boldsymbol{Y}_n-E\boldsymbol{Y}_n)\boldsymbol{T}$ is the same as that of $V_n^{-1}\boldsymbol{U}_1$ and now we begin to find the limiting distribution of $V_n^{-1}\boldsymbol{U}_1$. We first find the limiting variance of $V_n^{-1}\boldsymbol{U}_1$. By Assumption 2.3, we know that the first element of $\boldsymbol{U}_1$ is 0. Write $\boldsymbol{T}=(\boldsymbol{u}',\boldsymbol{T}_1,\dots,\boldsymbol{T}_s)=(\boldsymbol{u}',\boldsymbol{T}_-)$ and $\boldsymbol{T}_j=(\boldsymbol{t}'_{j1},\dots,\boldsymbol{t}'_{jv_j})$. Since the first element of $\widetilde{\boldsymbol{Q}}_i$ is 0, we obtain

$$\begin{aligned}
&\mathrm{Var}(V_n^{-1}\boldsymbol{U}_1)\\
&\quad= V_n^{-2}\sum_{i=1}^{n}\prod_{j=i+1}^{n}(\boldsymbol{I}+j^{-1}\boldsymbol{J}^*)\widetilde{\boldsymbol{R}}_i\prod_{j=i+1}^{n}(\boldsymbol{I}+j^{-1}\boldsymbol{J})\\
&\quad=\begin{pmatrix}0 & 0\\ 0 & [V_n^{-2}\sum_{i=1}^{n}\prod_{j=i+1}^{n}(\boldsymbol{I}+j^{-1}\boldsymbol{J}_g^*)\boldsymbol{T}_g^*\boldsymbol{R}_i\boldsymbol{T}_h\prod_{j=i+1}^{n}(\boldsymbol{I}+j^{-1}\boldsymbol{J}_h)]\end{pmatrix}\\
&\quad=\begin{pmatrix}0 & 0\\ 0 & [V_n^{-2}\sum_{i=1}^{n}\prod_{j=i+1}^{n}(\boldsymbol{I}+j^{-1}\boldsymbol{J}_g^*)\boldsymbol{T}_g^*\boldsymbol{R}\boldsymbol{T}_h\prod_{j=i+1}^{n}(\boldsymbol{I}+j^{-1}\boldsymbol{J}_h)]+\mathrm{o}(1)\end{pmatrix}.
\end{aligned}$$

We first consider the case $\tau<\frac{1}{2}$. We employ the identity

$$\int_0^1 x^a\log^b(1/x)\,\mathrm{d}x=\int_0^\infty y^b\mathrm{e}^{-(a+1)y}\,\mathrm{d}y=\frac{\Gamma(b+1)}{(a+1)^{b+1}},\quad \mathrm{Re}(a)>-1,\ b>-1$$

and the limit

$$\frac{1}{n}\sum_{i=1}^{n}(i/n)^a\log^b(n/i)\to\int_0^1 x^a\log^b(1/x)\,\mathrm{d}x.$$

For given g and h, the (s,t) element of the matrix

$$\left[V_n^{-2}\sum_{i=1}^{n}\prod_{j=i+1}^{n}(\boldsymbol{I}+j^{-1}\boldsymbol{J}_g^*)\boldsymbol{T}_g^*\boldsymbol{R}\boldsymbol{T}_h\prod_{j=i+1}^{n}(\boldsymbol{I}+j^{-1}\boldsymbol{J}_h)\right]$$

can be approximated by

$$\begin{aligned}
&\sum_{s'=0}^{s-1}\sum_{t'=0}^{t-1}n^{-1}\sum_{i=1}^{n}(n/i)^{\lambda_g+\bar{\lambda}_h}\frac{\log^{s'+t'}(n/i)}{s'!t'!}[\boldsymbol{T}_g^*\boldsymbol{R}\boldsymbol{T}_h]_{(s-s',t-t')}\\
&\quad\to\sum_{s'=0}^{s-1}\sum_{t'=0}^{t-1}\frac{(s'+t')!}{s'!t'!(1-\lambda_g-\bar{\lambda}_h)^{s'+t'+1}}[\boldsymbol{T}_g^*\boldsymbol{R}\boldsymbol{T}_h]_{(s-s',t-t')},
\end{aligned}\tag{2.17}$$

where $[\boldsymbol{T}_g^*\boldsymbol{R}\boldsymbol{T}_h]_{(s',t')}$ is the (s',t')-element of the matrix $[\boldsymbol{T}_g^*\boldsymbol{R}\boldsymbol{T}_h]$. This shows that

$$\left[V_n^{-2}\sum_{i=1}^{n}\prod_{j=i+1}^{n}(\boldsymbol{I}+j^{-1}\boldsymbol{J}_g^*)\boldsymbol{T}_g^*\boldsymbol{R}\boldsymbol{T}_h\prod_{j=i+1}^{n}(\boldsymbol{I}+j^{-1}\boldsymbol{J}_h)\right]$$

has a limit given in Eq. (2.17).

Next, we consider the case $\tau=\frac{1}{2}$. We shall use

$$\int_{1/n}^1 x^{-1}\log^b(1/x)\,\mathrm{d}x\leqslant\sum_{i=1}^{n}i^{-1}\log^b(n/i)\leqslant\log^b n+\int_{1/n}^1 x^{-1}\log^b(1/x)\,\mathrm{d}x$$

or equivalently,

$$\frac{\log^{b+1} n}{b+1} \leqslant \sum_{i=1}^{n} i^{-1} \log^b(n/i) \leqslant \log^b n + \frac{\log^{b+1} n}{b+1}.$$

This implies that

$$\frac{b+1}{\log^{b+1} n} \sum_{i=1}^{n} i^{-1} \log^b(n/i) \to 1. \tag{2.18}$$

Furthermore, for need to the case where $\boldsymbol{H}$ has two or more Jordan blocks with the same order and same eigenvalue of real part 1/2, we need the following inequality.

If $c \neq 0$, then

$$\left| \sum_{j=1}^{n} j^{-1+ic} \log^b(n/j) \right| \leqslant C_4 \log^b n, \tag{2.19}$$

for some constant $C_4 > 0$. In fact, Eq. (2.19) follows from Abelian summation and the elementary inequality

$$\sup_k \left| \sum_{j=1}^{k} j^{-1+ic} \right| \leqslant C_4 < \infty.$$

In turn, the latter follows from the facts that

$$\left| \int_1^{n+1} x^{-1+ic}\,\mathrm{d}x \right| = |(ic)^{-1}((n+1)^{ic} - 1)| \leqslant 2/|c|$$

and

$$\begin{aligned} \left| \sum_{j=2}^{n} j^{-1+ic} - \int_2^{n+1} x^{-1+ic}\,\mathrm{d}x \right| &\leqslant \sum_{j=2}^{n} \left| \int_j^{j+1} (j^{-1+ic} - x^{-1+ic})\,\mathrm{d}x \right| \\ &\leqslant \sum_{j=2}^{n} j^{-1} \left| \int_0^1 (1 - (1 - x/j)^{-1+ic})\,\mathrm{d}x \right| \\ &\leqslant \sum_{j=2}^{n} j^{-2}(1 - 1/j)^{-|c|}. \end{aligned}$$

Corresponding to Eq. (2.17), we have

$$\sum_{s'=0}^{s-1} \sum_{t'=0}^{t-1} n^{-1} \log^{-2v+1} n \sum_{i=1}^{n} (n/i)^{\lambda_g + \bar{\lambda}_h} \frac{\log^{s'+t'}(n/i)}{s'!t'!} [\boldsymbol{T}_g^* \boldsymbol{R} \boldsymbol{T}_h]_{(t-s',t-t')}$$

$$\to \frac{1}{(v-1)!(v-1)!(2v-1)} [\boldsymbol{T}_g^* \boldsymbol{R} \boldsymbol{T}_h]_{(1,1)}, \tag{2.20}$$

if $s = t = v$, $v_g = v_h = v$ and $\lambda_g = \lambda_h$ with $\mathrm{Re}(\lambda_g) = \frac{1}{2}$. For all other cases, its limit is zero.

This shows that in both cases, we have $V_n^{-2}\boldsymbol{T}^*\mathrm{Var}(\boldsymbol{Y}_n)\boldsymbol{T}$ tends to a limit which implies that $V_n^{-2}\mathrm{Var}(\boldsymbol{Y}_n)$ tends to a limit denoted by Σ. The reader should note that although the limit of $V_n^{-2}\boldsymbol{T}^*\mathrm{Var}(\boldsymbol{Y}_n)\boldsymbol{T}$ may be complex, the limit Σ must be real $K\times K$ matrix since $V_n^{-2}\mathrm{Var}(\boldsymbol{Y}_n)$ is real.

To complete the proof of the theorem, we employ Corollary 3.2 of Hall and Heyde (1980) through verifying the Lindeberg-type condition. In fact, when $\tau<1/2$,

$$\boldsymbol{U}_1=\sum_{i=1}^n\widetilde{\boldsymbol{Q}}_i\prod_{j=i+1}^n(\boldsymbol{I}+j^{-1}\boldsymbol{J})$$

$$=\sum_{i=1}^n\widetilde{\boldsymbol{Q}}_i\,\mathrm{diag}\left(\prod_{j=i+1}^n(1+\lambda/j),\prod_{j=i+1}^n(\boldsymbol{I}+j^{-1}\boldsymbol{J}_1),\dots,\prod_{j=i+1}^n(\boldsymbol{I}+j^{-1}\boldsymbol{J}_s)\right)$$

$$=\mathrm{diag}\left[0,\sum_{i=1}^n\boldsymbol{Q}_i\boldsymbol{T}_1\prod_{j=i+1}^n(\boldsymbol{I}+j^{-1}\boldsymbol{J}_1),\dots,\sum_{i=1}^n\boldsymbol{Q}_i\boldsymbol{T}_s\prod_{j=i+1}^n(\boldsymbol{I}+j^{-1}\boldsymbol{J}_s)\right].$$

By Eq. (2.6), the gth element of the tth block $\sum_{i=1}^n\boldsymbol{Q}_i\boldsymbol{T}_t\prod_{j=i+1}^n(\boldsymbol{I}+j^{-1}\boldsymbol{J}_t)$ has the approximation

$$\sum_{h=0}^g\sum_{i=1}^n\frac{1}{h!}(n/i)^{\lambda_t}\log^h(n/i)\boldsymbol{Q}_i\boldsymbol{t}'_{t,g-h}.$$

Then, for any $\varepsilon>0$ and given $h=0,1,\dots,g$, we have

$$V_n^{-2}\sum_{i=1}^n(n/i)^{2\tau}\log^{2h}(n/i)\bar{\boldsymbol{t}}_{t,g-h}E(\boldsymbol{Q}_i'\boldsymbol{Q}_i|\mathscr{F}_{i-1})\boldsymbol{t}'_{t,g-h}I_{\{A_{ni}\}}$$

$$\leqslant V_n^{-2}\sum_{i=1}^n(n/i)^{2\tau}\log^{2h}(n/i)|\bar{\boldsymbol{t}}_{t,g-h}(E(\boldsymbol{Q}_i'\boldsymbol{Q}_i|\mathscr{F}_{i-1})-\boldsymbol{R})\boldsymbol{t}'_{t,g-h}|$$

$$+V_n^{-2}\sum_{i=1}^n(n/i)^{2\tau}\log^{2h}(n/i)\bar{\boldsymbol{t}}_{t,g-h}\boldsymbol{R}\boldsymbol{t}'_{t,g-h}I_{\{A_{ni}\}}\to 0,$$

where $A_i=\{\|(n/i)^{\lambda_t}\log^h(n/i)\boldsymbol{Q}_i\boldsymbol{t}_{t,g-h}\|\geqslant\varepsilon V_n\}$. Here, that the first term on the right tends to 0 is a consequence of Assumption 2.4 and that for the second term is due to $P(A_{ni})\to 0$ for any fixed i.

Similarly, we can show the case for $\tau=1/2$. By Corollary 3.2 of Hall and Heyde (1980), $V_n^{-1}(\boldsymbol{Y}_n-E\boldsymbol{Y}_n)$ is asymptotic normal with mean vector 0 and variance–covariance matrix Σ. The proof of the theorem is complete. □

3. Applications

We now apply the results in Section 2 to three real situations.

3.1. Clinical trials with a time trend in adaptive allocation rules

In Example 1 given in Section 1, the success probability p_{ik} of treatment response T_i varies between different subjects. If the ith subject is assigned to treatment k, i.e., a type k particle is drawn in the ith stage, define the draw outcome $\boldsymbol{X}_i$ as the vector whose kth component is 1 or all others are 0. To do statistical inference upon the parameters p_{ik}, we need to know the properties of $\boldsymbol{X}_i$ and $\boldsymbol{Y}_i$.

Here, we consider the GFU discussed in Wei (1979): at the ith stage, a success on treatment k generates one type k particle, and a failure on treatment k generates $1/(K-1)$ particles for all other types. Then, the generating matrix is

$$\boldsymbol{H}_i = \begin{pmatrix} p_{i1} & q_{i1}/(K-1) & \cdots & q_{i1}/(K-1) \\ q_{i2}/(K-1) & p_{i2} & \cdots & q_{i2}/(K-1) \\ \cdots & \cdots & \cdots & \cdots \\ q_{iK}/(K-1) & q_{iK}/(K-1) & \cdots & p_{iK} \end{pmatrix}.$$

Also we define

$$\boldsymbol{H} = \begin{pmatrix} p_{1} & q_{1}/(K-1) & \cdots & q_{1}/(K-1) \\ q_{2}/(K-1) & p_{2} & \cdots & q_{2}/(K-1) \\ \cdots & \cdots & \cdots & \cdots \\ q_{K}/(K-1) & q_{K}/(K-1) & \cdots & p_{K} \end{pmatrix}.$$

From Theorems 2.2 and 2.3 of Section 2 (the Assumptions 2.1–2.4 are trivially satisfied), we have

Corollary 1. *For a sequence of positive constants α_i such that $\sum \alpha_i/i<\infty$, if $|p_{ik}-p_k|\leqslant \alpha_i$ for all $k=1,\dots,K$, then*

(i) $E_{i-1}(\boldsymbol{X}_i|\mathscr{F}_{i-1})=\boldsymbol{Y}_{i-1}/\boldsymbol{Y}_{i-1}\boldsymbol{1}'\to \boldsymbol{v}$ *in probability as* $i\to\infty$;

(ii) $V_n^{-1}(\boldsymbol{Y}_n-(n+\beta)\boldsymbol{v})$ *is asymptotic normal with mean vector* 0 *and variance–covariance matrix* Σ, *where* $V_n^2=n$ *when* $\tau<1/2$, *and* $V_n^2=n\log^{2v-1} n$ *if* $\tau=1/2$. *The vector* $\boldsymbol{v}=(1/q_1,\dots,1/q_K)/\sum 1/q_i$, *the constants* τ, v *are defined in Assumption* 2.1 *and the matrix* Σ *is defined in Theorem* 2.3.

3.2. The urn model associated with covariates

Consider Example 2 in Section 1. Assume that $\xi_1,\dots,\xi_n$ are i.i.d. Consequently, the adding rule matrices $\boldsymbol{D}(\xi_i)$'s $(K\times K)$ are i.i.d. Define $\boldsymbol{H}_i=E\boldsymbol{H}(\xi_1)=\boldsymbol{H}$. Applying Athreya and Karlin (1968), we get

$$\boldsymbol{Y}_{i-1}/\boldsymbol{Y}_{i-1}\boldsymbol{1}'\to\boldsymbol{v} \text{ almost surely as } n\to\infty.$$

From the results in Section 2, we can get the asymptotic normality of $\boldsymbol{Y}_n$. The asymptotic covariance matrix depends on the covariance of the $\boldsymbol{D}_i$. We shall discuss the case $K=2$ as an illustration.

Consider the generalised play-the-winner rule, we assume the success probability of the kth treatment at the ith stage p_{ik} has the form $p_k(\xi_i)$, $k=1,2$ and $i=1,\dots,n$.

We assume that the covariates ξ_i's are i.i.d. Then, $E(p_k(\xi_i)) = p_k$. The adding rule matrices are denoted by

$$\boldsymbol{D}_i = \begin{pmatrix} d_1(\xi_i) & 1-d_1(\xi_i) \\ 1-d_2(\xi_i) & d_2(\xi_i) \end{pmatrix} \quad \text{and generating matrix} \quad \begin{pmatrix} p_1 & q_1 \\ q_2 & p_2 \end{pmatrix},$$

where $0 \leqslant d_k(\xi_i) \leqslant 1$ and $q_k = 1 - p_k$ for $k = 1,2$.

Also, assume that $\mathrm{Var}\,(d_k(\xi_1)) = a_k$ and $\mathrm{Cov}\,(d_1(\xi_1), d_2(\xi_1)) = b$. Then, $\lambda = 1$, $\lambda_1 = p_1 + p_2 - 1$ and $\tau = \max(0, \lambda_1)$. We can easily show the following:

$$\boldsymbol{R} = \frac{(a_1 q_2 + a_2 q_1)(q_1+q_2) + q_1 q_2 (p_1 - q_2)^2}{(q_1+q_2)^2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix},$$

$$\boldsymbol{T} = \begin{pmatrix} 1 & q_1 \\ 1 & -q_2 \end{pmatrix} \quad \text{and} \quad \boldsymbol{T}^{-1} = \frac{1}{q_1+q_2} \begin{pmatrix} q_2 & q_1 \\ 1 & -1 \end{pmatrix}.$$

When $\tau < 1/2$, $V_n = n$ and the limit corresponding to Eq. (2.17) is

$$c = \frac{2(a_1 q_2 + a_2 q_1)(q_1+q_2) + 2 q_1 q_2 (p_1 - q_2)^2}{1 - 2(p_1 + p_2 - 1)}.$$

So

$$n^{-1/2}(\boldsymbol{Y}_n - \boldsymbol{v}) \to \mathrm{N}(0, \Sigma)$$

in distribution, where

$$\Sigma = (\boldsymbol{T}^*)^{-1} \begin{pmatrix} 0 & 0 \\ 0 & c \end{pmatrix} \boldsymbol{T}^{-1} = \frac{c}{(q_1+q_2)^2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

For the case $\tau = \frac{1}{2}$, $V_n = n \log n$ and the limit (corresponding to Eq. (2.20)) is $c_1 = (a_1 q_2 + a_2 q_1)(q_1 + q_2) + q_1 q_2 (p_1 - q_2)^2$. Thus we have

$$(n \log n)^{-1/2}(\boldsymbol{Y}_n - \boldsymbol{v}) \to \mathrm{N}(0, \Sigma_1)$$

in distribution, where

$$\Sigma_1 = (\boldsymbol{T}^*)^{-1} \begin{pmatrix} 0 & 0 \\ 0 & c_1 \end{pmatrix} \boldsymbol{T}^{-1} = \frac{c_1}{(q_1+q_2)^2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

Wei and Durham (1978) considered the randomized play-the-winner rule which is a special case of the generalized play-the-winner rule where $d_k(\xi_i) = 1$ if the kth treatment successes on the ith patient, or 0 otherwise. In this case, $a_k = p_k q_k$, $k = 1,2$. The matrices Σ and Σ_1 depend only on a_1 and a_2. The asymptotic covariance matrix Σ was studied by Smythe and Rosenberger (1995).

3.3. GFU with homogeneous generating matrix *H*

The results of Smythe (1996) are based on the following assumptions: (i) $\lambda > 2\nu$, (ii) all complex eigenvalues are simple, and (iii) the eigenvectors are linearly independent. But in many cases, the generating matrix does not satisfy the above conditions. Thus, the theorems of Smythe (1996) will no longer apply. On the other hand, Theorems 2.1–2.3 in Section 2 still apply to these homogeneous generating matrices.

For example, let $K=3$ and the homogeneous generating matrix be one of

$$\boldsymbol{H}_1=\begin{pmatrix} 1/2 & 1/6 & 1/3 \\ 1/6 & 1/2 & 1/3 \\ 1/4 & 1/12 & 2/3 \end{pmatrix} \quad \text{and} \quad \boldsymbol{H}_2=\begin{pmatrix} 1 & 0 & 0 \\ 1/2 & 1/2 & 0 \\ 1/4 & 1/4 & 1/2 \end{pmatrix}.$$

Easily we can see that

$$\boldsymbol{H}_1=\boldsymbol{T}_1\boldsymbol{J}_1\boldsymbol{T}_1^{-1} \quad \text{and} \quad \boldsymbol{H}_2=\boldsymbol{T}_2\boldsymbol{J}_2\boldsymbol{T}_2^{-1},$$

where

$$\boldsymbol{J}_1=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/3 & 1 \\ 0 & 0 & 1/3 \end{pmatrix}, \qquad \boldsymbol{T}_1=\begin{pmatrix} 1 & -1 & 24 \\ 1 & -1 & 0 \\ 1 & 1 & -15 \end{pmatrix},$$

$$\boldsymbol{J}_2=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 1 \\ 0 & 0 & 1/2 \end{pmatrix}, \qquad \boldsymbol{T}_2=\begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 4 \\ 1 & 1 & 1 \end{pmatrix}.$$

4. Further reading

Bartlett et al., 1985 Flournoy and Rosenberger (1995), Rosenberger and Sriram (1997), Serfling (1980), Wei et al., 1990, Tamura et al. 1994

Acknowledgements

The authors are grateful to Professor W.F. Rosenberger for valuable discussions with him which led to the problems of this paper. We also thank Professor Chen Gui-Jing for his helpful comments.

References

Altman, D.G., Royston, J.P., 1988. The hidden effect of time. Statist. Med. 7, 629–637.

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

Athreya, K.B., Ney, P.E., 1972. Branching Processes. Springer, Berlin.

Bartlett, R.H., Roloff, D.W., Cornell, R.G., Andrews, A.F., Dillon, P.W., Zwischenberger, J.B., 1985. Extracorporeal circulation in neonatal respiratory failure: a prospective randomized study. Pediatrics 76, 479–487.

Coad, D.S., 1991. Sequential tests for an unstable response variable. Biometrika 78, 113–121.

Coad, D.S., 1992. A comparative study of some data-dependent allocation rules for Bernoulli data. J. Statist. Comput. Simulation 40, 219–231.

Flournoy, N., Rosenberger, W.F. (Eds.), 1995. Adaptive Designs. Institute of Mathematical Statistics, Hayward.

Hall, P., Heyde, C.C., 1980. Martingale Limit Theory and its Application. Academic Press, London.

Rosenberger, W.F., 1996. New directions in adaptive designs. Statist. Sci. 11, 137–149.

Rosenberger, W.F., Flournoy, N., Durham, S.D., 1997. Asymptotic normality of maximum likelihood estimators from multiparameter response driven designs. J. Statist. Plann. Inference 60, 69–76.

Rosenberger, W.F., Grill, S.E., 1997. A sequential design for psychophysical experiments: an application to estimating timing of sensory events. Statist. Med. 16, 2245–2260.

Rosenberger, W.F., Sriram, T.N., 1997. Estimation for an adaptive allocation design. J. Statist. Plann. Inference 59, 309–319.

Serfling, R.J., 1980. Approximation Theorems of Mathematical Statistics. Wiley, New York.

Smythe, R.T., 1996. Central limit theorems for urn models. Stochastic Process. Appl. 65, 115–137.

Smythe, R.T., Rosenberger, W.F., 1995. Play-the winner designs, generalized Pólya's urns, and Markov branching processes. In: Flournoy, N., Rosenberger, W.F. (Eds.), Adaptive Designs. Institute of Mathematical Statistics, Hayward, CA, pp. 13–22.

Tamura, R.N., Faries, D.E., Andersen, J.S., Heiligenstein, J.H., 1994. A case study of an adaptive clinical trial in the treatment of out-patients with depressive disorder. J. Amer. Statist. Assoc. 89, 768–776.

Wei, L.J., 1979. The generalized Pólya's urn design for sequential medical trials. Ann. Statist. 7, 291–296.

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

Wei, L.J., Smythe, R.T., Lin, D.Y., Park, T.S., 1990. Statistical inference with data-dependent treatment allocation rules. J. Amer. Statist. Assoc. 85, 156–162.