第35卷第6期 中国科学技术大学学报 Vol. 35, No. 6
2005年12月 JOURNAL OF UNIVERSITY OF SCIENCE AND TECHNOLOGY OF CHINA Dec. 2005

Article ID: 0253-2778(2005)06-0789-07

Asymptotic Inference for Response-Adaptive Clinical Trial Design with General Outcomes*

ZHU Chun-hua[1,2], CHEN Gui-jing[2]

(1. *Department of Statistics and Finance, University of Science and Technology of China, Hefei* 230026, *China*;
2. *Department of Mathematics, Anhui University, Hefei* 230039, *China*)

Abstract: We consider the clinical trial with a continuous treatment response. First, applying the properties of the linear rank statistic, we obtain the asymptotic normality of the test statistic which is based on rank statistic proposed by Rosenberger and used to test the hypothesis of no treatment effect for a trial with continuous response. Then, we propose another assignment scheme for the trial with continuous response and obtain the asymptotic normality of the corresponding test statistic.

Key words: randomized play-the-winner rule; permutation test; linear rank statistic; asymptotic normality

CLC number: O212.7 **Document code**: A

AMS Subject Classification(2000): Primary 62G; Secondary 62L.

0 Introduction

In a clinical trial, response-adaptive treatment allocation rules[1,2] are to use accumulating information to assign patients to treatments with the goal of placing more patients on the more effective of two treatments. Ethical considerations make adaptive treatment assignment attractive, at least in principle. For a clinical trial with dichotomous response, the randomized play-the-winner (RPW)[3] rule has been proposed as a form of response-adaptive treatment allocation. And the Generalized Polya's Urn design[5~9] proposed later belongs to the forms of response-adaptive treatment allocation, too. For RPW rules, Rosenberger[4] derived the permutation tests under the null hypothesis of no treatment effect, and then, applying the martingale central limit theorem, proved that the

* **Received date**: 2003-12-17; **Revised date**: 2005-03-10

Foundation item: Supported by National Natural Science Foundation of China (10271001).

Biography: ZHU Chun-hua, female, born in 1976, PhD candidate. Research field: sequential analysis.
E-mail: zhuchunhua@ustc.edu

permutation test statistic is asymptotically normal under certain conditions on the sequence of responses. But for a clinical trial with continuous response, Rosenberger proposed a response-adaptive design with a large sample test statistic based on scores, while the asymptotic normality of the test statistic was just obtained by simulation results, which will be expressed in detail in Section 1.

In this paper, we study the asymptotic properties of the permutation test statistic based on rank statistic proposed by Rosenberger[4] for continuous response, and then propose another allocation rule and study the asymptotic properties of the corresponding test statistics.

1 A permutation test statistic based on scores calculated from a general variable

Rosenberger proposed a response-adaptive design along with a large sample test statistic based on scores as follows:

Let $Y_1, Y_2, \cdots, Y_n$ be a sequence of dichotomous treatment assignments, where $Y_j = 1$ or 0 according to whether patient j is assigned to treatment A or B, respectively. Write $n_A = \sum_{j=1}^{n} Y_j$, $n_B = n - n_A$. Then, among n responses, there are n_A responses of treatment A, denoted as $X_{n_1}, X_{n_2}, \cdots, X_{n_{n_A}}$, which distributed as $F(x)$; and n_B responses of treatment B, denoted as $X_{m_1}, X_{m_2}, \cdots, X_{m_{n_B}}$, which distributed as $G(x)$, where $1 \leqslant n_1 < n_2 < \cdots < n_{n_A} \leqslant n$, $1 \leqslant m_1 < m_2 < \cdots < m_{n_B} \leqslant n$.

To test the hypothesis of no treatment effect, we propose the null hypothesis as follows:

$$F(x) \equiv G(x), \qquad x \in R. \tag{1}$$

Under the null hypothesis, $X_{n_1}, X_{n_2}, \cdots, X_{n_{n_A}}, X_{m_1}, X_{m_2}, \cdots, X_{m_{n_B}}$ are identical independent distributed random variables. Then we can denote the n responses as $X_1, X_2, \cdots, X_n$, and $X_{(1)} < X_{(2)} < \cdots < X_{(n)}$ are their order statistics.

Let

$$R_{ij} = r, \quad \text{when} \quad X_i = X_{(r)}, i = 1, 2, \cdots, j, \tag{2}$$

where a larger rank indicates a better response to treatment. Define scores a_{ij} to be some function of the R_{ij}, $1 \leqslant i \leqslant j \leqslant n$, where $\sum_{i=1}^{j} a_{ij} = 0, j = 1, \cdots, n$. Define $a_{ij}^{+} = a_{ij} I(a_{ij} > 0)$, where I is the indicator function, and, let $\mathcal{F}_j = \sigma(Y_1, \cdots, Y_j; X_1, \cdots, X_j)$.

Let $\hat{p}_1 = \hat{p}_2 = \frac{1}{2}$, and

$$\hat{p}_i = E(Y_i \mid \mathcal{F}_{i-1}) = \frac{1}{2}\left\{1 + \frac{\sum_{j=1}^{i-1} a_{j,i-1}(Y_j - \frac{1}{2})}{\sum_{j=1}^{i-1} a_{j,i-1}^{+}}\right\}, i = 3, 4, \cdots \tag{3}$$

The better the responses of previous patients on treatment A, relative to those on B, the larger will be the probability that the next patient is assigned to A.

To test the hypothesis of no treatment effect, ref. [4] proposed the test statistic

$$S_n = \sum_{j=1}^{n} a_{jn}\left(Y_j - \frac{1}{2}\right). \tag{4}$$

Applying the martingale central limit theorem, Resenberger attempted to verify the asymptotic normality of S_n under some conditions. However, it is regretful that they only could give simulation results in some sense.

In the present paper, applying the properties of the linear rank statistic, we give the asymptotic normality of S_n.

Proposition 1.1 $n_A \to \infty, n_B \to \infty$, a. s. as $n \to \infty$.

Proof Let $a_{jn} = \phi\left(\frac{R_{jn}}{n+1}\right)$, and $\phi(u)\,(0<u<1)$ satisfy the following two conditions:

1) $\phi\left(\frac{1}{2}-u\right) = -\phi\left(\frac{1}{2}+u\right)$;

2) $\int_0^1 \phi^2(u)\,\mathrm{d}u < \infty$.

Write $A_M = \{n_A \leqslant M\}$, and $\widetilde{A} = \{\lim\limits_{n\to\infty} n_A < +\infty\} = \bigcup\limits_{M=1}^{\infty} A_M$. It is easily seen that there are less than M assignments on treatment A and more than $(n-M)$ on treatment B on A_M. Thus, among the n assignments the worst case is that M responses of treatment A are ranked before $(n-M)$ responses of treatment B.

For brevity, let $n = 2K$, K is a positive integer. Then $n \to \infty$ and $M < K$ gives

$$\hat{p}_{n+1} = \frac{1}{2}\left\{1 + \frac{\sum_{j=1}^{n} a_{jn}\left(Y_j - \frac{1}{2}\right)}{\sum_{j=1} 2a_{jn}^{+}}\right\} \geqslant$$

$$\frac{1}{2}\left\{1 + \frac{\frac{1}{2}\sum_{j=1}^{M}\phi\left(\frac{j}{n+1}\right) - \frac{1}{2}\sum_{j=M+1}^{K}\phi\left(\frac{j}{n+1}\right) - \frac{1}{2}\sum_{j=K+1}^{n}\phi\left(\frac{j}{n+1}\right)}{\sum_{j=1}^{n} a_{jn}^{+}}\right\} \to \frac{1}{2}.$$

So, there exists n_0, and when $n > n_0$, we have $\hat{p}_{n+1} \geqslant \frac{1}{3}$, that is, $1 - \hat{p}_{n+1} \leqslant \frac{2}{3}$, and

$$0 \leqslant P(A_M) = P(n_A \leqslant M) \leqslant \prod_{M=n_0+1}^{\infty}\left(\frac{2}{3}\right) = 0, \text{as } M \geqslant n_0.$$

Since A_M is increasing on M, then $P(A_M) = 0$, as $M \leqslant n_0$.

Hence
$$P(\widetilde{A}) = P\left(\bigcup_{M=1}^{\infty} A_M\right) \leqslant \sum_{M=1}^{\infty} P(A_M) = 0.$$

Then
$$n_A \to \infty, \quad \text{a. s.} \quad \text{as} \quad n \to \infty. \qquad \square$$

Write

$$\overline{Y} = \frac{1}{n}\sum_{j=1}^{n} Y_j = \frac{n_A}{n}, \bar{a} = \frac{1}{n}\sum_{j=1}^{n} a_{jn} = 0, Y_{nn}^2 = \sum_{j=1}^{n}(Y_j - \overline{Y})^2, a_n^2 = \sum_{j=1}^{n} a_{jn}^2. \tag{5}$$

Then, applying the properties of the linear rank statistic, we have

Theorem 1.2 S_n is asymptotical normality.

Proof There are infinite 0's and infinite 1's among the value of $Y_1, Y_2, \cdots$, since n_A, $n_B \to \infty$.

Then, condition on $Y_1 = y_1, Y_2 = y_2, \cdots, Y_n = y_n$,

$$\sum_{i=1}^{n}(Y_i - \overline{Y})^2 = \frac{n_A n_B}{n},$$

$$\max_{1\leqslant i\leqslant n}(Y_i - \overline{Y})^2 = \frac{(\max(n_A, n_B))^2}{n^2},$$

so
$$\frac{\sum_{i=1}^{n}(Y_i - \overline{Y})^2}{\max_{1\leqslant i\leqslant n}(Y_i - \overline{Y})^2} = \left[1 + \frac{\min(n_A, n_B)}{\max(n_A, n_B)}\right]\cdot \min(n_A, n_B) \to \infty, \text{ as } n \to \infty.$$

Hence, when $Y_1 = y_1, Y_2 = y_2, \cdots$, are given, the statistic $\frac{\sqrt{n-1}S_n}{\sqrt{Y_m^2 a_n^2}}$ converges in distribution to a standard normal variable. Then the theorem holds. □

2 A permutation test statistic for the urn model with general outcomes

In this section, for the trial with general outcomes, we propose another urn model as follows.

At the start of the trial, there are α balls of each other, say red and black, in the urn. When a patient is available for assignment to either of treatments A or B, a ball is drawn at random from the urn and replaced. A red ball generates an assignment to A, a black ball to B. When a response of a previously assigned patient becomes available, some additional red balls and black balls are added to the urn. Without loss of the generality, we assume the total number of two kinds of balls added to the urn is one at each stage.

For the continuous responses, we can discrete the responses to $(2L+1)$ levels, where L is a constant. Let $Y_1, Y_2, \cdots, Y_n$ be a sequence of dichotomous treatment assignments, where $Y_j = 1$ or 0 according to whether a patient is assigned to treatment A or B, respectively.

Let the responses of the treatment be $Z_1, Z_2, \cdots, Z_n$ and for $j = 1, 2, \cdots, n$, $\mathcal{F}_j = \sigma(Y_1, Y_2, \cdots, Y_j; Z_1, Z_2, \cdots, Z_n)$, the sigma algebra generated by the first j treatment assignments, and let $\mathcal{F}_0$ be the trivial sigma algebra.

Let $\tilde{p}_1 = \tilde{p}_2 = \frac{1}{2}$, and

$$E(Y_i \mid \mathcal{F}_{i-1}) = \tilde{p}_i = \frac{\alpha + \widetilde{S}_{i-1}}{2\alpha + i - 1}, i = 3, 4, \cdots, \tag{6}$$

where $\widetilde{S}_i = \sum_{j=1}^{i}\left\{Z_j\left(Y_j - \frac{1}{2}\right) + \frac{1}{2}\right\}$, which is the total number of red balls added to the urn after n assignments.

Now, we consider the null hypothesis of no treatment effect. Then, analogy to section

2, we assume: the responses are identical independent distributed sequences $Z_1, Z_2, \cdots$, which take on the value $\alpha_L, \alpha_{L-1}, \cdots, \alpha_1, \alpha_0, -\alpha_1, \cdots, -\alpha_L$, $(1=\alpha_L>\alpha_{L-1}>\cdots>\alpha_1>\alpha_0=0)$; and the corresponding probabilities are $p_L, p_{L-1}, \cdots, p_1, p_0, p_{-1}, \cdots, p_{-L}$ $\left(\sum_{j=-L}^{L} p_j = 1\right)$ respectively.

The test statistic of interest has the numerator

$$\sum_{j=1}^{n}\left\{Z_j\left(Y_j-\frac{1}{2}\right)\right\}, \tag{7}$$

which will take extreme values if there are significantly larger number of better responses on one treatment than on the other, leading to rejection of the hypothesis of equal treatment effects.

Let $\{\tilde{b}_{jn}, j=1,2,\cdots,n\}$ be a deterministic triangular array, chosen to make

$$\sum_{j=1}^{n}\left\{Z_j\left(Y_j-\frac{1}{2}\right)\right\}=\sum_{j=1}^{n}\tilde{b}_{jn}Z_j(Y_j-\tilde{p}_j), \tag{8}$$

for each n. This choice of $\tilde{b}_{jn}$ gives the equivalence of

$$\tilde{T}_n \equiv \frac{2\sum_{j=1}^{n}Z_j\left(Y_j-\frac{1}{2}\right)}{\gamma\left(\sum_{j=1}^{n}\tilde{b}_{jn}^2\right)^{\frac{1}{2}}}, \tag{9}$$

and $\widetilde{W}_{nn}$, where the array $\{\widetilde{W}_{nk}, k=1,2,\cdots,n,\}$ with

$$\widetilde{W}_{nk}=\frac{2\sum_{j=1}^{n}\tilde{b}_{jn}Z_j(Y_j-\tilde{p}_j)}{\left(\sum_{j=1}^{n}\tilde{b}_{jn}^2\right)^{\frac{1}{2}}}, \tag{10}$$

forms a martingale.

It is not difficult to verify that the desired sequence $\{\tilde{b}_{jn}\}$ is given by

$$\tilde{b}_{nn}=1, \tilde{b}_{jn}=\prod_{k=j+1}^{n}\left(1+\frac{1}{2\alpha+k-1}\cdot Z_k\right), j=1,\cdots,n-1. \tag{11}$$

Under the null hypothesis, we have the following results.

Denote $E(Z_k)=\beta, E(Z_k^2)=\gamma^2>0, k=1,2,\cdots,n$. Then we can get

$$\mathrm{Var}(\tilde{S}_n)=\frac{\gamma^2}{4}\sum_{j=1}^{n}\prod_{k=j+1}^{n}\left(1+\frac{2\beta}{2\alpha+k-1}\right). \tag{12}$$

Theorem 2.1　$\mathrm{Var}(\tilde{S}_n/n)\to 0$, as $n\to\infty$ for any $\beta<1$, where the expectation is taken over the $\{Z_j\}$ and $\{Y_j\}$.

Proof　From (12),

$$\mathrm{Var}(\tilde{S}_n)=\frac{\gamma^2}{4}\sum_{j=1}^{n}\prod_{k=j+1}^{n}\left(1+\frac{2\beta}{2\alpha+k-1}\right)\leqslant\frac{\gamma^2}{4}\sum_{j=1}^{n}\left(\frac{2\alpha+n}{2\alpha+i}\right)^{2\beta}. \tag{13}$$

Then, the rate of convergence of $\mathrm{Var}(\tilde{S}_n)$ is $o(n^2)$ for $\beta<\frac{1}{2}$. □

Corollary 2.2 $\frac{\widetilde{S}_n}{n} \to \frac{1}{2}$ in probability as $n \to \infty$.

Theorem 2.3 If $\beta < \frac{1}{2}$,

$$\frac{\max\limits_{1\leqslant j\leqslant n} \widetilde{b}_{jn}^2}{\sum\limits_{j=1}^{n} \widetilde{b}_{jn}^2} \to 0, \tag{14}$$

in probability as $n \to \infty$.

Proof We have that, for $j < n$, $\widetilde{b}_{jn} = \prod\limits_{k=j+1}^{n} \left(1 + \frac{Z_k}{2\alpha + k - 1}\right)$, so that

$$|\widetilde{b}_{jn}| \leqslant \exp\left\{\sum_{k=j+1}^{n} \frac{Z_k}{2\alpha + k - 1}\right\} = \exp(P_{jn})\exp\left(\beta \sum_{k=j+1}^{n} \frac{1}{2\alpha + k - 1}\right),$$

where $P_{jn} \sum\limits_{k=j+1}^{n} \frac{Z_k - \beta}{2\alpha + k - 1}$.

Write $A_n = \{\max\limits_{1\leqslant j<n} P_{jn} > \lambda_n\}$. By Kolmogorov' inequality, since

$$\mathrm{Var}\left(\sum_{j=a}^{n} \frac{Z_j - \beta}{2\alpha + k - 1}\right) = \sum_{j=a}^{n} \frac{1}{(2\alpha + k - 1)^2} < 2, \forall a = 1, 2, \cdots, n,$$

we have $P(A_n) = P(\max\limits_{1\leqslant j<n} P_{jn} > \lambda_n) \leqslant 2/\lambda_n^2$. On $A_n{}^c$, we have from that

$$|\widetilde{b}_{jn}| \leqslant \exp(\lambda_n + 1)\left(\frac{2\alpha + n - 1}{2\alpha + j - 1}\right)^{\beta},$$

so that

$$\max_{1\leqslant j\leqslant n} \widetilde{b}_{jn}^2 \leqslant \frac{\exp(2\lambda_n + 2)}{(2\alpha)^{2\beta}(2\alpha + n - 1)^{2\beta}} + o(n).$$

Now

$$E\left(\frac{\max\limits_{1\leqslant j\leqslant n} \widetilde{b}_{jn}^2}{\sum\limits_{j=1}^{n} \widetilde{b}_{jn}^2}\right) = \int_{A_n} \frac{\max\limits_{1\leqslant j\leqslant n} \widetilde{b}_{jn}^2}{\sum\limits_{j=1}^{n} \widetilde{b}_{jn}^2} \mathrm{d}P + \int_{A_n{}^c} \frac{\max\limits_{1\leqslant j\leqslant n} \widetilde{b}_{jn}^2}{\sum\limits_{j=1}^{n} \widetilde{b}_{jn}^2} \mathrm{d}P.$$

Let $\lambda_n = \ln(\ln n)$. Since $\frac{\max\limits_{1\leqslant j\leqslant n} \widetilde{b}_{jn}^2}{\sum\limits_{j=1}^{n} \widetilde{b}_{jn}^2} \leqslant 1$, the first integral does not exceed $P(A_n) \leqslant 2/\lambda_n^2 \to 0$.

For the second integral, it follows from an argument of ref. [10] that $\sum\limits_{j=1}^{n} \widetilde{b}_{jn}^2 \geqslant nc$ for a constant c.

Hence

$$\int_{A_n{}^c} \frac{\max\limits_{1\leqslant j\leqslant n} \widetilde{b}_{jn}^2}{\sum\limits_{j=1}^{n} \widetilde{b}_{jn}^2} \mathrm{d}P \leqslant \frac{1}{nc} E(\max_{1\leqslant j\leqslant n} \widetilde{b}_{jn}^2 I(A_n{}^c)) \leqslant \frac{\exp(2\lambda_n + 2)(2\alpha + n - 1)^{2\beta}}{nc(2\alpha)^{2\beta}} + \frac{o(n)}{n}.$$

This term tends to 0 as $n \to \infty$, provided $\beta < 1/2$. Hence the theorem holds. □

Remark In fact, when $\alpha_j = \frac{j}{L}$, $j = 1, 2, \cdots, L$, and $\sum\limits_{j=1}^{L} p_j - \sum\limits_{j=1}^{L} p_{-j} < \frac{L}{L+1}$, we can get $\beta < \frac{1}{2}$.

Theorem 2.4 $\widetilde{T}_n$ converges in distribution to a normal variable.

Proof By Corollary 3.1 of ref. [11], applying Corollary 2.2 and Theorem 2.3, we can easily obtain the normality of $\widetilde{T}_n$. □

3 Conclusion

For general responses, Rosenberger proposed the permutation test statistic based on the rank, and only gave the simulation results. Now, applying the properties of the linear rank statistic, we showed the asymptotic normality of the test statistic, and at the same time we proposed another urn design for general outcomes.

Reference

[1] Rosenberger W F, Sriram T N. Estimation for an adaptive allocation design [J]. J. Statist. Plann. Inf., 1997, 59: 309-319.

[2] Melfi V F, Page C. Estimation after adaptive allocation[J]. J. Statist. Plann. Inf., 2000, 87 (2): 353-363.

[3] Wei L J, Durham S D. The randomized play-the-winner rule in medical trials[J]. J. A. S. A., 1978, 73: 840-843.

[4] Rosenberger W F. Asymptotic inference with response-adaptive treatment allocation design [J]. Ann. Statis., 1993, 21(4): 2 098-2 107.

[5] Bai Z D, Chen G J, HU F F. Some theorems under urn models with time trends [J]. Chinese Annals of Mathematics, 2001, 22 (A): 89-96.

[6] Bai Z D, Chen G J, Hu F F. A kind of urn model for adaptive sequential design[J]. Acta Mathematica Scientica, 2001, 21 (B): 224-228.

[7] Bai Z D, Hu F F. Asymptotic theorems for Urn models with nonhomogeneous generating matrices[J]. Stochastic Processes and Their Applications, 1999, 80: 1, 87-101.

[8] Chen Gui-jing, Chen Yu-ming, Bai Zhi-dong, Hu Fei-fang. A kind of optimal design of urn models[J]. Acta Mathematica Scentica, 2003, 23A(2): 208-214.

[9] Wei L J. The generalized Polya's urn designs for sequential medical trials [J]. Ann. Statist., 1979, 7: 291-296.

[10] Wei L J, Smythe R T. Smith R L. K-treatment comparisons with restricted randomization rules in clinical trials [J]. Ann. Statist, 1986, 14: 265-274.

[11] Hall P, Heyde C C. Martingale Limit Theory and Its Application [M]. London: Academic Press, 1980.

具有一般反应的自适应临床设计的渐近推断

朱春华[1,2], 陈桂景[2]

(1.中国科学技术大学统计与金融系,安徽合肥 230026;2.安徽大学数学系,安徽合肥 230039)

摘要:研究了具有连续反应的临床试验.首先应用线性秩统计量的渐近理论,获得由Rosenberger 提出的无差异治疗效果假设检验统计量的渐近正态性;然后给出一种新的设计且获得相应统计量的渐近正态性.

关键词:随机优胜者优先原则;置换检验;线性秩统计量;渐近正态性