The Annals of Statistics
1979, Vol. 7, No. 2, 291–296

THE GENERALIZED PÓLYA'S URN DESIGN FOR SEQUENTIAL MEDICAL TRIALS

BY L. J. WEI

University of South Carolina-Columbia

In comparing $K(\geqslant 2)$ treatments in a medical trial, suppose that the response of the patient to treatment is dichotomous. We propose and analyze a class of simple and nondeterministic treatment assignment schemes which tend to put more patients on better treatments. These schemes are also applicable when we have delayed responses from patients.

1. Introduction. In comparing $K(\geqslant 2)$ treatments in a medical trial, suppose that eligible patients occur singly and must be treated when they arrive. The major goal of this trial is to gather sound data from current patients to derive information about the effectiveness of these K treatments for the benefit of future patients. Another goal of this trial is to treat each current patient in the best way which we can. This is due to the ethical problem of studies on human beings. These two goals are contradictory to some extent. Our purpose is to provide a treatment assignment scheme which tends to put more current patients on better treatments, but is also able to give us reliable information about treatment effectiveness after the trial is over.

To meet the ethical requirement, Zelen (1969) introduced the play the winner rule (PW) for comparing two treatments A and B into medical trials. This rule can be described as follows: 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. The PW rule may be implemented by placing in an urn balls marked with an "A" whenever a success is obtained with treatment A or a failure with treatment B. Similarly balls marked with a "B" are placed in the urn whenever a success is obtained with treatment B or a failure with treatment A. When a new patient enters the trial, the treatment assignment is determined by drawing a ball randomly from the urn *without* replacement; if the urn is empty, then the assignment is determined by the tossing of a fair coin. In practice, the time to observe the response of a patient to treatment is much longer than the time between patient entry. Therefore, most assignments are determined by the tossing of a fair coin. This results in an approximately equal number of patients on each treatment and the PW rule is of little value.

In the case when the response of the nth patient to treatment is known before the $(n + 1)$th patient enters the trial, i.e., the response is instantaneous, the PW rule can be modified so that after each success we continue to use the same treatment

Received October 1977; revised December 1977.
AMS 1970 *subject classifications.* Primary 62L05; Secondary 60K30.
Key words and phrases. The generalized Pólya's urn design, the multi-type continuous time branching process, experimental bias, cyclic play the winner rule.

and after each failure we switch to the other treatment. Zelen (1969) called this rule the "modified play the winner rule" (MPW) which has been studied extensively from selection and ranking theory approach. A good summary is given by Hoel, Sobel and Weiss (1975). The MPW rule tends to put more patients on the better treatment, but it is too deterministic and may introduce bias to the trial. Hoel and Sobel (1971) have extended the idea of the MPW rule to the case when more than two treatments are compared. They introduced the cyclic play the winner rule (PWC) which is defined in the next section. Again, this PWC rule is too deterministic and is not applicable when we have delayed responses from patients to treatments.

Efron (1971), Pocock and Simon (1975), and Wei (1977, 1978) have considered some nondeterministic treatment assignment rules in medical trials, but they did not take the ethical problem into consideration.

In this article, a class of new treatment assignment rules is proposed and analyzed in Sections 3 and 4. These rules tend to put more current patients on better treatments, but they are not deterministic and allow the delayed response of the patient to treatment. Another important feature of these new rules is that it is fairly easy to implement them in a real trial.

2. The cyclic play the winner rule (PWC). At the outset of the trial, we order the K given treatments at random and use this ordering in a cyclic manner. Suppose that the response of the patient to treatment is instantaneous and dichotomous. Let p_i denote the probability of the single trial success for treatment i and let $q_i = 1 - p_i$, where $i = 1, 2, \ldots, K$ and $0 < p_i < 1$. After each success, we treat the next patient with the same treatment; after each failure, we switch to the next treatment in the ordering scheme; after completing the cycle, we go back to the first treatment. This assignment scheme is called the cyclic play the winner rule (PWC) (Hoel and Sobel, 1971).

The PWC rule can be regarded as a Markov chain with K states corresponding to K treatments after the ordering of treatments is fixed. Define the random variable T_n to be $T_n = i$, if the nth patient receives treatment i. Then, the transition probabilities are:

$$P(T_n = i | T_{n-1} = i) = p_i, \qquad P(T_n = i + 1 | T_{n-1} = i) = q_i,$$
$$P(T_n = K | T_n = K) = p_K, \qquad \text{and} \quad P(T_n = 1 | T_n = K) = q_K,$$

where $i = 1, 2, \ldots, K - 1$. It is easy to see that this chain $\{T_n\}$ is irreducible, positive recurrent, aperiodic and has a stationary distribution $\xi = (\xi_1, \ldots, \xi_K)'$, where

$$\xi_i = (1/q_i)/\Sigma_{j=1}^{K}(1/q_j), \qquad i = 1, \ldots, K.$$

It follows that ξ is also the limiting distribution of the Markov chain $\{T_n\}$. Let $N_i(n)$ be the number of patients treated by treatment i after n assignments, then by

Theorem 4.2.1 of Kemeny and Snell (1960),

$$N_i(n)/n \to_p \xi_i, \qquad i = 1, 2, \ldots, K, \quad \text{as } n \to \infty.$$

If the ordered p values are denoted by $p_{[1]} \leqslant p_{[2]} \leqslant \ldots \leqslant p_{[K]}$, then the corresponding ξ values keep the same ordering, i.e., the PWC tends to put more patients on better treatments. But the PWC is completely deterministic after the first assignment and is not applicable when we have delayed responses from patients.

3. The generalized Pólya's urn design GPUD (W, α, β). We still randomly label these K treatments at the outset of the trial. The requirement of having instantaneous responses in the last section is not needed for this new assignment rule. This rule can best be explained by a generalized Pólya's urn scheme in the following manner. An urn has balls of K different colors. We start with w_i balls of color i, $i = 1, \ldots, K$. When an eligible patient arrives at the experimental site, a ball is selected at random from the urn. We observe its color i and return the ball to the urn. Treatment i is then assigned to this patient. When the response of a previous patient to treatment i is available, we perform the following operations: (i) if the response is a success, we add $\alpha(> 0)$ balls of color i; (ii) if the response is a failure, we add $\beta(> 0)$ balls of each color j, where $j = 1, 2, \ldots, K$ and $j \neq i$. This treatment assignment rule is called a generalized Pólya's urn design and is denoted by GPUD (**W**, α, β), where $\mathbf{W} = (w_1, \ldots, w_K)'$. Obviously, this design is applicable when we have delayed responses from patients to treatments.

If there is no information about the relative effectiveness of these K treatments at the outset of the trial, we let $w_i = w$, $i = 1, \ldots, K$. We will concentrate on the scheme GPUD ($w\mathbf{1}$, $K - 1$, 1), where $\mathbf{1} = (1, \ldots, 1)'$, because it most closely parallels the PWC scheme.

In the rest of this article, we analyze the GPUD (**W**, α, β) based on the assumption that the response of the patient is instantaneous. Let $\mathbf{X}_n = (X_{n1}, \ldots, X_{nK})'$ denote the composition of the urn after n successive assignments. Also, let $\mathbf{Z}(t) = (Z_1(t), \ldots, Z_K(t))'$ be a K-type continuous time Markov branching process (Mode, 1971; Athreya and Ney, 1972). Assume that (i) $\mathbf{Z}(0) = \mathbf{W}$; (ii) the life time of particles of all types are unit exponentials; and (iii) an ith type particle creates, on death, new particles of all types according to the probability generating function $f_i(\mathbf{s})$, where

$$f_i(\mathbf{s}) = s_1^{\beta} \ldots s_{i-1}^{\beta} s_i s_{i+1}^{\beta} \ldots s_K^{\beta} q_i + s_i^{\alpha+1} p_i.$$

Then, it is easy to see that the processes $\{\mathbf{X}_n;\ n = 0, 1, 2, \ldots\}$ and $\{\mathbf{Z}(\tau_n);\ n = 0, 1, 2, \ldots\}$ are equivalent, where $\{\tau_n;\ n = 0, 1, 2, \ldots\}$ denotes the split times of the process $\{\mathbf{Z}(t);\ t \geqslant 0\}$.

Since $f_i(\mathbf{0}) = 0$ for all i, it follows that $\tau_n \to_{\text{a.s.}} \infty$, as $n \to \infty$ (page 1807, Athreya and Karlin, 1968). Therefore, any limit relation which holds almost surely as $t \to \infty$ for the process $\{\mathbf{Z}(t);\ t \geqslant 0\}$ is also valid for the process $\{\mathbf{Z}(\tau_n);\ n = 0, 1, 2, \ldots\}$. Let $\mu_i(s) = f_i(s) - s_i$, $i = 1, 2, \ldots, K$. Then it can be shown that

$(\partial^2\mu_i(\mathbf{s}))/(\partial s_j \partial s_l)|_{\mathbf{s}=\mathbf{1}} < \infty$ for all j, l and $i = 1, 2, \ldots, K$. Also, since $m_{ij} = (\partial\mu_i(\mathbf{s}))/\partial s_j|_{\mathbf{s}=\mathbf{1}} > 0$, there exists a $t_0 > 0$ such that each element of the matrix $\exp(t_0 M)$ is positive, where

$$m_{ij} = \alpha p_i, \quad j = i,$$
$$= \beta q_i, \quad j \neq i,$$

and $M = \|m_{ij}\|_{K\times K}$. It follows that the process $\{\mathbf{Z}(t);\ t \geqslant 0\}$ is nonsingular, supercritical, and positive regular with extinction probability zero.

If we let $N_i(n)$ be the number of patients treated by treatment i after n assignment when GPUD $(\mathbf{W}, \alpha, \beta)$ is utilized, $i = 1, 2, \ldots, K$, then $N_i(n)$ essentially is the number of splits, among the first n, which are of type i, of the process $\{\mathbf{Z}(t);\ t \geqslant 0\}$. From Theorem 5 of Athreya and Karlin's paper (1967), we have

$$N_i(n)/n \to_{\text{a.s.}} \phi_i = v_i/\Sigma_{j=1}^K v_j, \tag{3.1}$$

where $\mathbf{v} = (v_1, \ldots, v_K)'$ is a left eigenvector (with positive components) of M corresponding to the maximal positive eigenvalue λ. Also, it is easy to show that for each i

$$Z_i(t)/\Sigma_{j=1}^K Z_j(t) \to_{\text{a.s.}} \phi_i \qquad \text{as} \quad t \to \infty.$$

It follows that

$$X_{ni}/\Sigma_{j=1}^K X_{nj} \to_{\text{a.s.}} \phi_i \qquad \text{as} \quad n \to \infty. \tag{3.2}$$

Since the matrix M is positive, there always exist a maximal positive eigenvalue λ and a corresponding left eigenvector with positive components. For $K = 2$, it is easy to get λ and $\mathbf{v}$ of M. The ratio $N_1(n)/N_2(n)$ then tends to $(\phi(p_1 - p_2) + (\phi^2(p_1 - p_2)^2 + 4q_1q_2)^{\frac{1}{2}})/2q_1$, with probability one, as $n \to \infty$, where $\phi = \alpha/\beta$. We note that this asymptotic ratio is increasing in ϕ when $p_1 > p_2$. Therefore, the larger ϕ is, the more patients the GPUD assigns to the better treatment. But, if ϕ is too large, then the GPUD becomes pretty deterministic and is vulnerable to experimental bias (c.f. Section 4).

When $K > 2$, λ and $\mathbf{v}$ are difficult to get analytically. However, it is not the case for GPUD $(\mathbf{W}, K - 1, 1)$. For this situation, obviously, $K - 1$ is an eigenvalue of the matrix M. But the maximal eigenvalue λ of M is bounded by $\max_j(\Sigma_{l=1}^K m_{jl}) = K - 1$. (c.f. page 120, Cox and Miller, 1965). It follows that the maximal positive eigenvalue $\lambda = K - 1$ and the ith component of the corresponding left eigenvector is $v_i = 1/q_i$. If the ordered p values are denoted by $p_{[1]} \leqslant p_{[2]} \leqslant \ldots \leqslant p_{[K]}$, then the corresponding ϕ values keep the same ordering: $\phi_{[1]} \leqslant \phi_{[2]} \leqslant \ldots \leqslant \phi_{[K]}$. We note that, for the design $G(\mathbf{W}, K - 1, 1)$, $\phi_i = \xi_i$, where ξ_i is defined in Section 2, $i = 1, 2, \ldots, K$. It follows that the PWC rule and the GPUD $(\mathbf{W}, K - 1, 1)$ tend to place the same proportions of patients on better treatments as the size of the trial increases.

We also report some results from small-sized trials. Table I shows the comparison between the PWC rule and the GPUD $(\mathbf{1}, K - 1, 1)$, where $K = 3$. In most cases of this table, the GPUD $(\mathbf{1}, 2, 1)$ tends to put more patients on better treatments than the PWC rule does.

Table I.
Comparisons between GPUD (1, 2, 1) *and* PWC *rule*

				GPUD (1, 2, 1)			PWC rule		
				expected no. of patients treated by treatment			expected no. of patients treated by treatment		
p_1	p_2	p_3	n	1	2	3	1	2	3
.4	.2	.1	6	2.2581	1.9399	1.8021	2.2292	1.9249	1.8458
			12	4.6710	3.8330	3.4960	4.4677	3.8581	3.6742
			18	7.1104	5.7155	5.1741	6.7093	5.7892	5.5016
			27	10.7933	8.5286	7.6781	10.0713	8.6857	8.2430
.6	.3	.2	6	2.3916	1.8751	1.7333	2.4108	1.8321	1.7571
			12	5.0679	3.6407	3.2914	4.8604	3.6591	3.4805
			18	7.8129	5.3747	4.8124	7.3087	5.4867	5.2046
			27	11.9885	7.9470	7.0645	10.9811	8.2281	7.7908
.8	.4	.2	6	2.5982	1.8419	1.5599	2.7568	1.6730	1.5702
			12	5.7051	3.4897	2.8052	5.6510	3.3047	3.0443
			18	8.9670	5.0595	3.9735	8.5457	4.9363	4.5180
			27	14.0068	7.3372	5.6560	12.8878	7.3837	6.7285
.9	.5	.3	6	2.6278	1.8338	1.5384	2.9476	1.5604	1.4920
			12	5.8504	3.4415	2.7045	6.1829	3.0068	2.8103
			18	9.2827	4.9459	3.7714	9.4169	4.4536	4.1295
			27	14.6445	7.0910	5.2645	14.2680	6.6238	6.1082

The scheme GPUD $(\mathbf{W}, K-1, 1)$ appears a little drastic in its early stages especially when K is large. Some alternatives may be considered. For example, we can take $\alpha = l$ for the lth success, $l = 1, 2, \ldots$. But for this scheme, the corresponding continuous time process $\{\mathbf{Z}(t)\}$ is no longer Markovian. In general, it is difficult to analyze a non-Markovian process.

4. Nondeterminicity of the GPUD $(\mathbf{W}, \alpha, \beta)$. If the physician knows or guesses which treatment a patient will receive before he selects the patient, then he may, consciously or subconsciously, bias the trail by his choice as to who is or is not a suitable patient. A natural measure of this kind of bias is the expected number of correct guesses of treatment assignments the physician can make if he guesses optimally. The best strategy against the GPUD $(\mathbf{W}, \alpha, \beta)$ is to guess the treatment on the basis of which is most likely to be used in the next assignment. It follows that the probability of having a correct guess at stage $(n+1)$ is

$$E\left(\max_{1\leqslant i\leqslant K}\left(X_{ni}/\Sigma_{j=1}^{K}X_{nj}\right)\right). \tag{4.1}$$

From (3.2) and the dominated convergence theorem, (4.1) converges to $\max_{1\leqslant i\leqslant K}\phi_i$, where ϕ_i is defined in (3.1). Since $M = \|m_j\|$ is positive, $\mathbf{e}_i$ cannot be a left eigenvector of M, where $\mathbf{e}_i$ is a $1 \times K$ vector whose components are all zero except for the ith component, $i = 1, 2, \ldots, K$. It follows that $\max_{1\leqslant i\leqslant K}\phi_i < 1$. Therefore, the GPUD $(\mathbf{W}, \alpha, \beta)$ is not deterministic and is not vulnerable to experimental bias.

Acknowledgments. I would like to thank the referee and the Associate Editor for valuable comments on the original version of this article and thank Dr. A. Cantor for computing Table I for me.

REFERENCES

[1] ATHREYA, K. B. and KARLIN, S. (1967). Limit theorems for the split times of branching processes. *J. Math. Mech.* **17** 257–277.

[2] ATHREYA, K. B. and KARLIN, S. (1968). Embedding of urn schemes into continuous time Markov branching process and related limit theorems. *Ann. Math. Statist.* **39** 1801–1817.

[3] ATHREYA, K. B. and NEY, P. E. (1972). *Branching Processes*. Springer-Verlag, New York.

[4] COX, D. R. and MILLER, H. D. (1965). *The Theory of Stochastic Processes*. Methuen & Co. Ltd., London.

[5] EFRON, B. (1971). Forcing a sequential experiment to be balanced. *Biometrika* **58** 403–417.

[6] HOEL, D. G. and SOBEL, M. (1972). Comparison of sequential procedures for selecting the best binomial population. *Proc. Sixth Berkeley Symp. Math. Statist. and Probabilities* **IV** 53–69.

[7] HOEL, D. G., SOBEL, M. and WEISS, G. H. (1975). A survey of adaptive sampling for clinical trials. In *Advances in Biometry*, (Ed. R. M. Elashoff). Academic Press, New York.

[8] KEMENY, J. G. and SNELL, J. L. (1960). *Finite Markov Chains*. D. Van Nostrand Co., New York.

[9] MODE, C. J. (1971). *Multitype Branching Processes*. American Elsevier, New York.

[10] POCOCK, S. J. and SIMON, R. (1975). Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. *Biometrics* **31** 103–115.

[11] WEI, L. J. (1977). A class of designs for sequential clinical trials. *J. Amer. Statist. Assoc.* **72** 382–386.

[12] WEI, L. J. (1978). Adaptive biased coin design. *Ann. Statist.* **6** 92–100.

[13] ZELEN, M. (1969). Play the winner rule and the controlled clinical trial. *J. Amer. Statist. Assoc.* **64** 131–146.

DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE
UNIVERSITY OF SOUTH CAROLINA
COLUMBIA, SOUTH CAROLINA 29208