Randomly Reinforced Urn Designs whose Allocation Proportions Converge to Arbitrary Prespecified Values

Giacomo Aletti, Andrea Ghiglietti and Anna Maria Paganoni

Abstract We construct a response adaptive design, described in terms of two-color urn model targeting a fixed asymptotic allocation. We prove asymptotic results for the process of colors generated by the urn and for the process of its compositions. We also discuss the use of this urn model, for an estimation problem, in sequential clinical trials.

Key words: Reinforced processes; urn schemes; sequential clinical trials; stochastic processes

1 Introduction

Consider a clinical trial with two competitive treatments, say R and W. We construct a response adaptive design, described in terms of two-color urn model, targeting any optimal fixed asymptotic allocation. A large class of response-adaptive randomized designs is based on urn models, a classical tool to guarantee a randomized device [3, 7], to balance the allocations [2] or to construct designs which asymptotically assign all subjects to the best treatment [4, 6]. The *two-color, Randomly Reinforced Urn* (RRU) introduced in [5] and studied in [6], is a randomized device that is able to target the optimal treatment, see [6]. Here we modify the reinforcement scheme

Giacomo Aletti
Dipartimento di Matematica "F. Enriques", Università degli studi di Milano,
via Saldini 50, 20133 Milano, Italy, e-mail: giacomo.aletti@unimi.it

Andrea Ghiglietti
Dipartimento di Matematica "F. Brioschi", Politecnico di Milano
Piazza Leonardo da Vinci 32, 20123 Milano, Italy, e-mail: andrea.ghiglietti@mail.polimi.it

Anna Maria Paganoni
Dipartimento di Matematica "F. Brioschi", Politecnico di Milano
Piazza Leonardo da Vinci 32, 20123 Milano, Italy, e-mail: anna.paganoni@polimi.it

of the urn in order to target asymptotically an optimal allocation proportion. Let us consider two probability distributions μ_R and μ_W with support contained in $[\alpha,\beta]$, where $0 \leq \alpha \leq \beta < +\infty$ and a sequence $(U_n)_n$ of independent uniform random variables on $(0,1)$. We will interpret μ_R and μ_W as the laws of the responses to treatment R and W, respectively. We assume that both the means $m_R = \int_\alpha^\beta x\mu_R(dx)$ and $m_W = \int_\alpha^\beta x\mu_W(dx)$ are strictly positive. To get some intuition into the process X_n, visualize an urn initially containing r_0 balls of color R and w_0 balls of color W. Set

$$R_0 = r_0,\ W_0 = w_0,\ Z_0 = \frac{R_0}{D_0}.$$

At time $n = 1$, a ball is sampled from the urn; its color is $X_1 = \mathbf{1}_{[0,Z_0]}(U_1)$, a random variable with Bernoulli (Z_0) distribution. Let M_1 and N_1 be two independent random variables with distribution μ_R and μ_W, respectively; assume that X_1, M_1 and N_1 are independent. Next, if the sampled ball is R, it is replaced in the urn together with X_1M_1 balls of the same color if $Z_0 < \eta$, where $\eta \in (0,1)$ is a suitable parameter, otherwise the urn composition does not change; if the sampled ball is W, it is replaced in the urn together with $(1-X_1)N_1$ balls of the same color if $Z_0 > \delta$, where $\delta < \eta \in (0,1)$ is a suitable parameter, otherwise the urn composition does not change. So we can update the urn composition in the following way

$$R_1 = R_0 + X_1M_1\mathbf{1}_{[Z_0<\eta]},\ W_1 = W_0 + (1-X_1)N_1\mathbf{1}_{[Z_0>\delta]},\ Z_1 = \frac{R_1}{D_1}.$$

Now iterate this sampling scheme forever. Thus, at time $n+1$, given the sigma-field $\mathscr{F}_n$ generated by $X_1,...,X_n,M_1,...,M_n$ and $N_1,...,N_n$, let $X_{n+1} = \mathbf{1}_{[0,Z_n]}(U_{n+1})$ be a Bernoulli(Z_n) random variable and, independently of $\mathscr{F}_n$ and X_{n+1}, assume that M_{n+1} and N_{n+1} are two independent random variables with distribution μ_R and μ_W, respectively. Set

$$R_{n+1} = R_n + X_{n+1}M_{n+1}\mathbf{1}_{[Z_n<\eta]},\ W_{n+1} = W_n + (1-X_{n+1})N_{n+1}\mathbf{1}_{[Z_n>\delta]},\ Z_{n+1} = \frac{R_{n+1}}{D_{n+1}}.$$

We thus generate two infinite sequences $(X_n)_{n\in\mathbb{N}}$ and $(Z_n)_{n\in\mathbb{N}}$ of random variables, representing the color of the ball sampled from the urn and the proportion of balls of color R, respectively.

In [1] the following asymptotic convergence result is proved.

Theorem 1. *The sequence of proportions $(Z_n)_{n\in\mathbb{N}}$ of the urn process converges almost surely to the following limit*

$$lim_{n\to\infty} Z_n = \begin{cases} \eta \text{ if } \int_\alpha^\beta x\mu_R(dx) > \int_\alpha^\beta x\mu_W(dx), \\ \\ \delta \text{ if } \int_\alpha^\beta x\mu_R(dx) < \int_\alpha^\beta x\mu_W(dx). \end{cases}$$

The urn proportion process $(Z_n)_{n\in\mathbb{N}}$ converges to a value which depends on the unknown means of the reinforcement distributions. This aspect characterizes the

adaptive nature of the design based on the urn model. In particular this modified urn model generates a process $(Z_n)_{n\in\mathbb{N}}$ that converges to one of the values $\{\delta,\eta\}$, according the reinforcement with the greatest mean. When $m_R = m_W$ we don't have the explicit form of the asymptotic distribution of the urn proportion Z_n. Nevertheless, we know that $(Z_n)_{n\in\mathbb{N}}$ converges to a random variable Z_∞ whose distribution has no atoms and its support is $S_\infty = [\delta,\eta]$.

2 An application to estimation in clinical trials

Let us consider a treatment W, and suppose its mean effect on patients to be estimated during an adaptive clinical trial. Suppose that the distribution of random effect on patients of the competitor R can be modified arbitrarily. The aim of the experiment is to infer the mean effect of the treatment W by modifying suitably the mean effect of treatment R. Let us consider K urns with the same initial composition (R_0,W_0). Red balls are associated with treatment R, while white balls with treatment W. We will denote with $Z^j = (Z_n^j)_{n\in\mathrm{N}}$ the process of the urn proportion in the j^{th} urn, for $j \in \{1,2,..,K\}$.

At the beginning, we choose the distribution of treatment R, in order to set the mean of the random response to an initial value $m_{R,1}$. Let K urn processes start simultaneously. At each step, we draw a ball form every urn and we update the composition of each urn independently, following the model described in section 1. After reinforcements have been performed, we will have the compositions of K urns, and we can compute the empirical cumulative distribution function $\widehat{F}_n$ of the random variable Z_n. Thanks to the Theorem 1, for every $x \in [0,1]$, $\widehat{F}_n(x)$ must converge to

$$\begin{cases} F_\eta(x) = 1_{\{x\geq\eta\}} \text{ if } m_{R,1} > m_W, \\ \\ F_\delta(x) = 1_{\{x\geq\delta\}} \text{ if } m_{R,1} < m_W. \end{cases}$$

In the case of $m_{R,1} = m_W$, we compute off line the asymptotic cumulative distribution $\widehat{F}_e$ of $Z_\infty = Z_e$. In other words, we simulate other M urns, in order to get the empirical distribution of the limit Z_∞. For this purpose M, the number of urns and m the number of draws can be arbitrarily large.

$$\frac{1}{M}\sum_{i=1}^{M} 1_{\{Z_m^i<x\}} \simeq \widehat{F}_e(x), \qquad \text{for large m and M.}$$

At each step, once every urn has been reinforced, we use the Wasserstein distance (d_W) to compute the distances between Z_n and the three asymptotic possible distributions. Then, we take the minimum among these distances and if it is lower than a suitable threshold α we can assume the proportion Z_n has reached its limit.

Otherwise, the urn processes go on with further draws. We stop the algorithm at step $\tilde{n}$ if

$$\min\left\{d_W(Z_{\tilde{n}},\delta_\eta)\ ,d_W(Z_{\tilde{n}},Z_e)\ ,d_W(Z_{\tilde{n}},\delta_\delta)\right\}=$$

$$\min\left\{\int_0^1|F_{\tilde{n}}(x)-F_\eta(x)|dx\ ,\ \int_0^1|F_{\tilde{n}}(x)-\widehat{F}_e(x)|dx\ ,\ \int_0^1|F_{\tilde{n}}(x)-F_\delta(x)|dx\right\}<\alpha.$$

When the stopping rule is verified, different scenarios are possible. If the minimum distance is $d(Z_{\tilde{n}},\delta_\eta)$ we can assume $m_{R,1}$ was greater than the unknown mean m_W. For this reason, we change the composition of treatment R to decrease the mean effect to a new suitable value $m_{R,2}<m_{R,1}$. Alternatively, if the lowest distance was $d(Z_{\tilde{n}},\delta_\delta)$ we increase the mean effect of treatment R, in order to have a mean $m_{R,2}>m_{R,1}$. In any case, we have that the difference between the two means is decreasing, i.e., $|m_{R,2}-m_W|<|m_{R,1}-m_W|$. At this point, we start over with K urn processes, with the same initial composition (r_0,w_0). The model is the same as before, with the only difference that the mean of the reinforcement of red balls has been changed. Now, random responses to treatment R follow a distribution law with an updated mean value equal $m_{R,2}$, whereas random responses to treatment W have the previous mean value equal to m_W.

The experiment goes on until the stopping rule is verified and $d(Z_{\tilde{n}},Z_e)<\alpha$. If we call i_0 the number of times we have modified the mean of the random responses to treatment R, we can suppose that $m_{R,i_0}=m_W$ and this is a good estimate of the unknown mean m_W. Compared to a non-adaptive design this procedure, not only provides a good estimate of m_W, but also allocates a greater proportion of patients to the best treatment.

References

1. Aletti, G., Ghiglietti, A., Paganoni, A.M.: A modified randomly reinforced urn design. In: Mox Report n. 32-2011, Dipartimento di Matematica, Politecnico di Milano. http://mox.polimi.it/it/progetti/pubblicazioni/quaderni/32-2011.pdf
2. Baldi Antognini, A., Giannerini, S.: Generalized Polya urn designs with null balance. In: Journal of Applied Probability, 44, 661–669 (2007).
3. Cheung, S.H., Zhang, Li-X., Hu, F.: Asymptotic theorems of sequential estimation-adjusted urn models. In: The Annals of Applied Probability, 16, 340–369 (2006).
4. Flournoy, N., May, C.: Asymptotic theorems of sequential estimation-adjusted urn models. In: The Annals of Statistics, 37, 1058–1078 (2009).
5. May, C., Paganoni, A.M., Secchi, P.: On a two-color generalized Plya urn. In: Metron, LXIII n.1, 115-134 (2005).
6. Muliere, P., Paganoni, A.M., Secchi, P.: A randomly reinforced urn. Journal of Statistical Planning and Inference, 136, 1853–1874 (2006).
7. Rosenberger, W. F.: Randomized urn models and sequential design. Sequential Analysis, 21(1&2), 1–28 (2002).