Proceedings of the 2016 Winter Simulation Conference
T. M. K. Roeder, P. I. Frazier, R. Szechtman, E. Zhou, T. Huschka, and S. E. Chick, eds.

PROPOSAL FOR FULLY SEQUENTIAL MULTIARM TRIALS WITH CORRELATED ARMS

Ozge Yapar, Noah Gans

OID Department – Wharton School
University of Pennsylvania
3730 Walnut Street, Suite 500
Philadelphia, PA 19104-6340, USA

Stephen E. Chick

Technology and Operations Management Area
INSEAD
Boulevard de Constance
77300 Fontainebleau, France

ABSTRACT

We focus on the design of multiarm multistage (MAMS) clinical trials, using ideas from simulation optimization, biostatistics, and health economics. From a trial design perspective, we build on the trend of comparing multiple treatments with a single control by allowing for more than two arms in a trial, and we allow for arbitrarily many stages of sampling by using a diffusion approximation that allows for adaptive stopping rules. From a simulation perspective, our techniques extend the correlated knowledge-gradient concept, which has been used in one-stage lookahead (knowledge gradient) procedures, to Bayesian fully sequential selection procedures.

1 INTRODUCTION

Economic and ethical considerations in medical research are driving increased attention to the design of clinical trials. Organizations continue look for ways to better balance the high costs of developing new therapies with budget constraints. Furthermore, ethical mandates motivate researchers to get safe and effective health technologies to market quickly and without excess expenditure and, conversely, to prevent technologies that are not safe or effective from going to market and harming patients.

Trends in trials design reflect these economic and ethical pressures. The use of multiple arms in a given trial, rather than only two arms, allows multiple technologies to be compared at the same time to an existing standard. (Arms might represent different doses or protocols or combinations of treatment decisions.) The adoption of multiple stages of sampling allows for treatments to be accepted (or rejected) earlier if the results of a trial emerge as particularly positive (or negative). Explicit accounting for cost effectiveness, not just clinical efficacy, is becoming more common.

The current work is inspired by those trends. We use a Bayesian expected value of information framework (Chick and Inoue 2001b, Frazier et al. 2008, Ryzhov et al. 2015) to account for both correlated means of rewards of alternatives and adaptive stopping times for sampling decisions. This approach extends existing correlated knowledge gradient results which account for correlated means but have used one-stage lookahead approximations (Chick and Inoue 2001a, Frazier et al. 2009, Frazier et al. 2016) and work on adaptive stopping times which has assumed independent means (Chick and Gans 2009, Chick and Frazier 2012) or only two alternatives (Chick et al. 2015).

2 MATHEMATICAL FORMULATION AND CONTRIBUTION

Suppose that we have M arms (alternatives) to choose from. Let $X_{i,t}$ be the reward of arm i when sampled at time $t (t = 0, 1, \ldots)$. We assume that each of the $X_{i,t}$s is a composite representation of health and cost data into a univariate outcome (e.g., via factors like \$75000/quality adjusted life year to convert health to money).

We assume samples are jointly independent and normally distributed, conditional on unknown mean θ_i and known variance λ_i. Let $\boldsymbol{\theta} = (\theta_1, .., \theta_M)^\mathsf{T}$ and let $\boldsymbol{\lambda}$ be the $M \times M$ diagonal matrix $\text{diag}(\lambda_1, .., \lambda_M)$. Our belief about $\boldsymbol{\theta}$ is distributed according to a multivariate normal prior, with $\boldsymbol{\theta} \sim N(\boldsymbol{\mu}^t, \boldsymbol{\Sigma}^t)$, given t observations so far. The decision maker specifies the initial prior distribution with $\boldsymbol{\mu}^0, \boldsymbol{\Sigma}^0$, and $\boldsymbol{\mu}^t, \boldsymbol{\Sigma}^t$ is computed with Bayes' rule, $\boldsymbol{\lambda}$, and observations. The means are correlated if $\boldsymbol{\Sigma}^0$ is not diagonal.

Let $c_i \geq 0$ be the cost of sampling arm i, and let $\Delta \in (0, 1]$ be a discount factor. We require $\Delta < 1$ or all $c_i > 0$, or both. Let a^t be the arm to sample next, given that t samples have been observed, for $t = 0, 1, \ldots, T-1$ for some stopping time, T, which is adapted to the sequence of observations.

A policy π gives a sequence of arms to test for patients in a trial, a stopping time T, and an arm a^T to adopt. We assume that P patients will benefit from the adoption decision and that a fixed cost I_i is incurred if arm i is selected for adoption. With this model, the expected reward of a non-anticipative policy π is

$$V^\pi(\boldsymbol{\mu}^0, \boldsymbol{\Sigma}^0) \;=\; \mathbb{E}_\pi \left[\sum_{t=0}^{T-1} -\Delta^t c_{a^t} \;+\; \Delta^T \{ P\mathbb{E}[X_{a^T, T+1} | \boldsymbol{\mu}^T, \boldsymbol{\Sigma}^T] - I_{a^T} \} \;\middle|\; \boldsymbol{\mu}^0, \boldsymbol{\Sigma}^0 \right]. \tag{1}$$

Our problem is to identify a policy π^* that maximizes the expected reward, $V^{\pi^*}(\boldsymbol{\mu}^0, \boldsymbol{\Sigma}^0) = \sup_\pi V^\pi(\boldsymbol{\mu}^0, \boldsymbol{\Sigma}^0)$.

3 SUMMARY OF RESULTS

Our initial results quantify the expected reward when sampling from a single alternative and when using an adaptive stopping rule for sampling before deciding to choose an alternative to select as best from the trial. We account for the fact that an observation from one alternative may change beliefs about the rewards of other alternatives, and we show that a sum (of terms involving the newsvendor loss function) similar to that in Frazier, Powell, and Dayanik (2009) is obtained for the case of correlated means and adaptive stopping rules. Our sum uses terms that are readily approximated by partial-differentiation approximations to the reward obtained by optimal sequential sampling. Those terms depend on whether rewards are discounted (Chick and Gans 2009) or not (Chick and Frazier 2012) and depend on the correlation structure of the unknown mean rewards.

The poster summarize our initial theoretical and numerical results. It represents an extension to the state-of-the-art related to simulation optimization using the correlated knowledge gradient approach.

REFERENCES

Chick, S. E., M. Forster, and P. Pertile. 2015. "A Bayesian Decision-Theoretic Model of Sequential Experimentation with Delayed Response". *INSEAD Technical Report*.

Chick, S. E., and P. I. Frazier. 2012. "Sequential Sampling for Selection with Economics of Selection Procedures". *Management Science* 58 (3): 550–569.

Chick, S. E., and N. Gans. 2009. "Economic Analysis of Simulation Selection Problems". *Management Science* 55 (3): 421–437.

Chick, S. E., and K. Inoue. 2001a. "New Procedures for Identifying the Best Simulated System Using Common Random Numbers". *Management Science* 47 (8): 1133–1149.

Chick, S. E., and K. Inoue. 2001b. "New Two-Stage and Sequential Procedures for Selecting the Best Simulated System". *Operations Research* 49 (5): 732–743.

Frazier, P., W. Powell, and S. Dayanik. 2009. "The knowledge-gradient policy for correlated normal beliefs". *INFORMS Journal on Computing* 21 (4): 599–613.

Frazier, P., W. B. Powell, and S. Dayanik. 2008. "A Knowledge-Gradient Policy for Sequential Information Collection". *SIAM Journal on Control and Optimization* 47:2410–2439.

Frazier, P., J. Xie, and S. Chick. 2016. "Bayesian Optimization via Simulation with Pairwise Sampling and Correlated Prior Beliefs". *Operations Research* 42 (2): 542–559.

Ryzhov, I. O., P. I. Frazier, and W. B. Powell. 2015. "A new optimal stepsize for approximate dynamic programming". *IEEE Transactions on Automatic Control* 60 (3): 743–758.