Sequential Sampling for Selection: The Undiscounted Case

Stephen E. Chick[1]
Peter I. Frazier[2]

[1]Technology & Operations Management, INSEAD
[2]Operations Research & Information Engineering, Cornell University

Monday November 8, 2010
INFORMS Annual Meeting
Austin

Ranking and Selection for Discrete Event Simulation

- We have a discrete event simulator that can simulate the consequences of alternative real-world decisions, e.g.,
  - Designs of a queuing network.
  - Inventory policies for a supply chain.
  - Pricing strategies for a revenue management problem.
- Goal: find an alternative that works well, according to the simulator.
- Our simulator needs significant time to accurately characterize an alternative, and we do not have enough time to do so for each one.
- Which alternatives should we simulate and for how long?

Ranking and Selection for Discrete Event Simulation

- We have a discrete event simulator that can simulate the consequences of alternative real-world decisions, e.g.,
  - Designs of a queuing network.
  - Inventory policies for a supply chain.
  - Pricing strategies for a revenue management problem.
- Goal: find an alternative that works well, according to the simulator.
- Our simulator needs significant time to accurately characterize an alternative, and we do not have enough time to do so for each one.
- Which alternatives should we simulate and for how long?
- We study this problem using **Bayesian decision theory**, using **economic costs** of simulation and alternative selection.

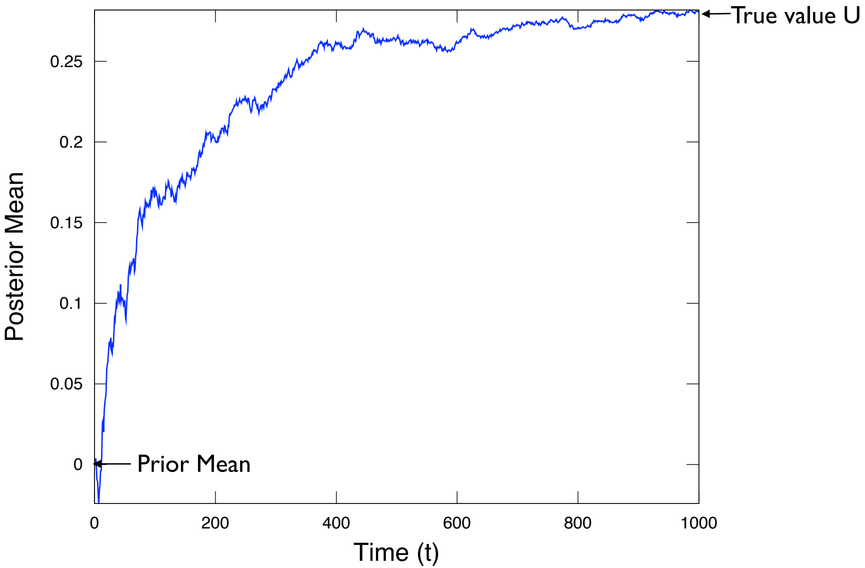
Ranking & Selection (R&S)

- We have k alternatives. Alternative $x \in \{1, \ldots, k\}$ has true value, U_x.
- There is also a standard option with known value U_0, e.g., $U_0 = 0$ is the value of "doing nothing."
- Sampling alternative x gives a noisy observation of U_x,

$$y \sim \mathrm{Normal}(U_x, \sigma_x^2),$$

  where we suppose the measurement variance σ_x^2 is known.
- To describe our belief about $U_1, \ldots, U_k$, we assume an independent normal prior distribution. Let $\vec{\Theta}_0$ be a vector containing the means and variances of the prior.
- After observing a sequence of samples, we will have a posterior distribution that is also normal. Let $\vec{\Theta}_t$ be this posterior.

Bayesian Posterior Probability Distribution



Cost and Benefit of Sampling

- A fully sequential policy π is a rule for adaptively choosing which alternative to sample at each point in time, and when to stop.
- Notation:
  - T is the total number of samples.
  - $I(T)$ is the alternative selected as the best.
  - c is the cost of one sample (can allow dependence on the alternative).
- Sampling incurs a direct cost, but improves our eventual choice $I(T)$.
- The value of a policy π given the information in the posterior Θ is

$$V^{\pi}(\Theta) = \mathrm{E}_{\pi}\left[-cT \; + \; U_{I(T)} \; | \vec{\Theta}_0 = \vec{\Theta}\right].$$

- Goal: find the policy with optimal value, $V^{*}(\vec{\Theta}) = \sup_{\pi} V^{\pi}(\vec{\Theta})$.

Previous Literature

This work builds on two related sections of the literature.

- Economics of simulation: [Chick and Gans, 2009] considers a discounted version of our problem. We extend this work by considering the undiscounted case, and by developing a **new and improved policy**.
- Knowledge-gradient: [Gupta and Miescke, 1996, Frazier et al., 2008] derive an allocation rule based on a single-step expected value of information calculation. [Frazier and Powell, 2008] extends this idea to stopping rules. We extend this work by considering **multi-step valuations of information**.

Special Case: $k = 1$

- Consider the special case of comparing a **single** alternative against a known standard:

$$V^*(\vec{\Theta}) \; = \; \sup_{\pi} \mathrm{E}_{\pi} \left[-cT \; + \; \max\{U_0, \mu_T\} | \vec{\Theta}_0 = \vec{\Theta} \right]$$

  where $\mu_T = \mathrm{E}[U_1 \mid \vec{\Theta}_T]$ is the posterior mean of the alternative.
- This is an optimal stopping problem.
- We relax this problem by allowing T to take real values, instead of just integers.
- Then the posterior mean μ_T becomes a diffusion, and the optimal stopping problem becomes a **free-boundary problem**.

Ease of Use

- If we solve for the optimal stopping boundary for standard values $c = 1$, $\sigma = 1$, $U_0 = 0$, a simple algebraic transformation provides the optimal stopping boundary for any values c, σ, U_0.
- Let $\pm b(t)$ be the optimal boundary for the standard problem.
- We use this approximation to $b(t)$, with little loss in performance.

$$b(t) \approx \begin{cases} .233s^2 & \text{if } s \leq 1 \\ .00537s^4 - .06906s^3 + .3167s^2 - .02326s & \text{if } 1 < s \leq 3 \\ .705s^{1/2}\ln(s) & \text{if } 3 < s \leq 40 \\ .642(s(2\ln(s))^{1.4} - \ln(32\pi))^{1/2} & \text{if } 40 < s, \end{cases}$$

where $s = 1/t$.

- **This approximation is easy to compute, and does not require solving the free-boundary problem**.

Multiple Alternatives

We now consider multiple alternatives, and derive or re-derive stopping and allocation rules using the idea of **value of information**.

- In general, the optimal stopping rule is

$$T = \inf\left\{t \geq 0 : V^*(\vec{\Theta}_t) - \max_{x=0,\ldots,k} \mu_{Tx} = 0\right\}.$$

  - $\max_x \mu_{Tx}$ is the value obtained by taking no more samples.
  - $V^*(\vec{\Theta}_t)$ is the maximal value that can be extracted given $\vec{\Theta}_t$.
  - $V^*(\vec{\Theta}_t) - \max_x \mu_{Tx} \geq 0$ is the **net value of continuing to sample** in an optimal way.
- $V^*(\vec{\Theta}_t)$ is hard to calculate for $k > 1$. We approximate it.

PDE Stopping Rule

- The optimal stopping rule is

$$T = \inf\left\{t \geq 0 : V^*(\vec{\Theta}_t) - \max_{x=0,\ldots,k} \mu_{Tx} = 0\right\}.$$

- $V^*(\vec{\Theta}_t)$ is hard to compute, so we approximate it as the maximum of the value functions $V^*_x(\vec{\Theta}_t)$ for "single-alternative problems".

$$V^*(\vec{\Theta}_t) \approx \max_{x=0,\ldots,k} V^*_x(\vec{\Theta}_t)$$

- In the single-alternative problem for x, we may only sample x, and upon stopping we can select either x or the best of the rest. $V_x(\vec{\Theta}_t)$ can be computed using the approximation for the $k = 1$ problem.
- We call this the **PDE stopping rule**, and it is easy to compute.

Stopping Rules in Numerical Study

We compared PDE against several other stopping rules derived using approximations to the value of information.

- PDE: **single** alternative, **adaptive** sample size.
  (This talk, optimal for $k = 1$)
- KG_1: **single** alternative, **single** sample.
  [Frazier et al., 2008]
- KG_*: **single** alternative, **deterministic** sample size.
  [Frazier and Powell, 2010]
- $EOC_{c,k}$: **multiple** alternatives, **deterministic** sample size.
  [Chick and Inoue, 2001]

Allocation Rules in Numerical Study

These approximations to the value of information also imply allocation rules.

- PDE: Sample the alternative whose posterior mean is furthest from the $k = 1$ stopping boundary.
- KG_1: Sample the alternative with the largest expected value of information (EVI). [Frazier et al., 2008]
- KG_*: Sample the alternative with the largest average EVI per sample (over deterministic rules). [Frazier and Powell, 2010]
- Sequential LL (based on EOC): Sample the alternative to which the most samples are allocated by the allocation with the best net EVI. [Chick and Inoue, 2001]

Numerical Results ($k > 1$)

- Table shows expected loss $E[cT + OC]$ for pairs of stopping and allocation rules. **Lower is better**.
- PDE stopping with KG_* allocation is the best policy.
- It is better than LL, $EOC_{c,1}$, which was best in the large empirical study [Branke et al., 2007].

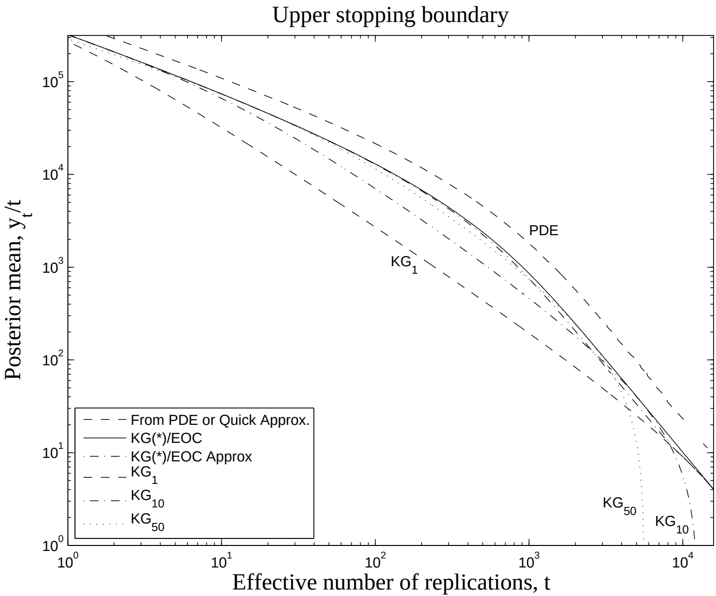
Alloc,Stop	k=3	10	20	50	100
KG_1,KG_1	3508 ± 12	7140 ± 18	8767 ± 19	10862 ± 67	12500 ± 72
KG_*,KG_*	674 ± 4	1445 ± 6	1761 ± 6	2245 ± 23	2666 ± 25
Equal,$EOC_{c,k}$	433 ± 2	1040 ± 3	1815 ± 3	4220 ± 16	8425 ± 29
LL,$EOC_{c,k}$	429 ± 2	821 ± 4	1095 ± 4	1577 ± 17	2168 ± 22
KG_*,$EOC_{c,k}$	424 ± 2	799 ± 3	1057 ± 4	1489 ± 16	2027 ± 19
KG_1,$EOC_{c,k}$	419 ± 2	728 ± 3	916 ± 3	1223 ± 11	1577 ± 11
KG_1, PDE	348 ± 2	694 ± 3	875 ± 3	1158 ± 10	1516 ± 10
PDE, PDE	344 ± 2	700 ± 3	856 ± 3	1075 ± 11	1308 ± 12
KG_*, PDE	327 ± 2	600 ± 2	722 ± 3	905 ± 8	1111 ± 9

Numerical Results ($k > 1$): Stopping Rule

- Consider the effect of the stopping rule.
- PDE is the best stopping rule, followed in order by $EOC_{c,k}$, KG_*, and KG_1. KG_1 performed badly because it underestimates the value of information.

	k=3	10	20	50	100
KG_1,KG_1	3508 ± 12	7140 ± 18	8767 ± 19	10862 ± 67	12500 ± 72
Equal,$EOC_{c,k}$	433 ± 2	1040 ± 3	1815 ± 3	4220 ± 16	8425 ± 29
KG_*,KG_*	674 ± 4	1445 ± 6	1761 ± 6	2245 ± 23	2666 ± 25
LL,$EOC_{c,k}$	429 ± 2	821 ± 4	1095 ± 4	1577 ± 17	2168 ± 22
KG_*,$EOC_{c,k}$	424 ± 2	799 ± 3	1057 ± 4	1489 ± 16	2027 ± 19
KG_1,$EOC_{c,k}$	419 ± 2	728 ± 3	916 ± 3	1223 ± 11	1577 ± 11
KG_1,PDE	348 ± 2	694 ± 3	875 ± 3	1158 ± 10	1516 ± 10
PDE,PDE	344 ± 2	700 ± 3	856 ± 3	1075 ± 11	1308 ± 12
KG_*,PDE	327 ± 2	600 ± 2	722 ± 3	905 ± 8	1111 ± 9

Stopping Boundary



Numerical Results ($k > 1$): Allocation Rule

- Consider the effect of the allocation rule.
- KG_* and KG_1 are the best allocation rules, despite poor performance as stopping rules. They consistently underestimate the value of information, but this bias cancels when making allocation decisions.
- The PDE allocation rule also performs well.

	k=3	10	20	50	100
KG_1, KG_1	3508 ± 12	7140 ± 18	8767 ± 19	10862 ± 67	12500 ± 72
KG_*,KG_*	674 ± 4	1445 ± 6	1761 ± 6	2245 ± 23	2666 ± 25
Equal,$EOC_{c,k}$	433 ± 2	1040 ± 3	1815 ± 3	4220 ± 16	8425 ± 29
LL,$EOC_{c,k}$	429 ± 2	821 ± 4	1095 ± 4	1577 ± 17	2168 ± 22
KG_*,$EOC_{c,k}$	424 ± 2	799 ± 3	1057 ± 4	1489 ± 16	2027 ± 19
KG_1,$EOC_{c,k}$	419 ± 2	728 ± 3	916 ± 3	1223 ± 11	1577 ± 11
KG_1,PDE	348 ± 2	694 ± 3	875 ± 3	1158 ± 10	1516 ± 10
PDE,PDE	344 ± 2	700 ± 3	856 ± 3	1075 ± 11	1308 ± 12
KG_*,PDE	327 ± 2	600 ± 2	722 ± 3	905 ± 8	1111 ± 9

Conclusion

- This approach balances the cost of sampling with the rewards of having information – financial criteria may be more appropriate than statistical criteria in most business decisions.
- We can solve the $k = 1$ case "exactly" with a PDE, and that solution supports understanding of the $k > 1$ problem.
- The resulting PDE stopping rule is empirically better than those from prior experiments.
- This approach may have application in more complex simulation optimization problems (e.g. unknown variances, CRN, correlated beliefs, metamodels), not just independent variance-known ranking and selection.

Thank You; Any Questions?

- If you are interested in these topics, please consider submitting a paper to an upcoming **special issue of IIE Transactions** devoted to simulation optimization and its applications.
- Due Date for Submission: June 2011
- Special Issue Editors: Loo Hay Lee; Ek Peng Chew; Samuel Qing-Shan Jia; Peter Frazier; Chun-Hung Chen

Branke, J., Chick, S., and Schmidt, C. (2007).
Selecting a selection procedure.
Management Sci., 53(12):1916–1932.

Chick, S. and Gans, N. (2009).
Economic analysis of simulation selection problems.
Management Sci., 55(3):421–437.

Chick, S. and Inoue, K. (2001).
New two-stage and sequential procedures for selecting the best simulated system.
Operations Research, 49(5):732–743.

Frazier, P. and Powell, W. (2008).
The knowledge-gradient stopping rule for ranking and selection.
Winter Simul. Conf. Proc., 2008.

Frazier, P. and Powell, W. (2010).
Paradoxes in learning and the marginal value of information.
Decision Analysis, page deca.1100.0190.

Frazier, P., Powell, W. B., and Dayanik, S. (2008).

Numerical Results ($k = 1$)

Expected loss of stopping rules for $k = 1$, $c = 1$, $\mu_0 = 0$, $t_0 = 100$, and $\sigma = 10^5$ calculated using Monte Carlo simulation with 10^6 samples. OC = Opportunity Cost = $\max_i U_i - U_{I(T)}$

Stopping Rule	E[cT]	E[OC]	E[cT+OC]	% sub-optimality
PDE	321.85 ± 0.25	263 ± 1	585 ± 1	—
$EOC_{c,k}$	142.53 ± 0.11	612 ± 2	755 ± 2	4.99%
KG_*	136.51 ± 0.11	634 ± 2	770 ± 2	5.45%
KG_1	10.53 ± 0.01	2505 ± 5	2515 ± 5	56.69%

Better approximations to the value of information give better performance.