Structural Properties of the Value of Information in Single-Stage Ranking and Selection

Peter Frazier[1]
Warren Powell[2]

[1]Operations Research & Information Engineering, Cornell University
[2]Operations Research & Financial Engineering, Princeton University

Sunday October 11, 2009
INFORMS Annual Meeting, San Diego

Ranking and Selection for Discrete Event Simulation

- We have a large discrete event simulator with which we can estimate the consequences of a future real-world decision. For example:
  - Designs of a queuing network.
  - Inventory policies for a supply chain.
  - Pricing strategies for a revenue management problem.
  - ...
- We would to find a real-world decision that will work well, according to the simulator.
- Our simulator requires significant time to accurately characterize a decision, and we do not have enough time to do so for each one.
- Question: which real-world decisions should we simulate, and how accurately?

Ranking & Selection (R&S)

- We have M alternatives.
- Each alternative has a true value, θ_x, which is unknown.
- We may sample alternative x and get a noisy observation of θ_x,

$$y \sim \text{Normal}(\theta_x, \lambda_x),$$

  where we suppose the measurement variance λ_x is known.
- We wish to decide which alternatives to sample in order to most efficiently find $\arg\max_x \theta_x$.
- Sampling is expensive. The central question is how to sample most efficiently.

Bayesian Ranking & Selection

- In the Bayesian formulation, we have a prior or posterior belief on θ.
  - In simulation, this belief is often the posterior resulting from a first stage of measurements.
- Let n_i be the number of measurements to take from alternative i, and let $\mathbf{n} = (n_1, \dots, n_M)$.
- Let $\mathbf{Y}$ be the set of observations resulting from these measurements.
- The value of this set of measurements $\mathbf{n}$ is

$$v(\mathbf{n}) = \mathbb{E}\left[\max_i \mathbb{E}[\theta_i \mid \mathbf{Y}, \mathbf{n}] \mid \mathbf{n}\right] - \max_i \mu_i.$$

The Value of Information

- One would generally like to find

$$\max v(\mathbf{n})$$
$$\text{subject to } n_x \geq 0, \quad \sum_x n_x \leq N,$$

  where N is a budget constraint.

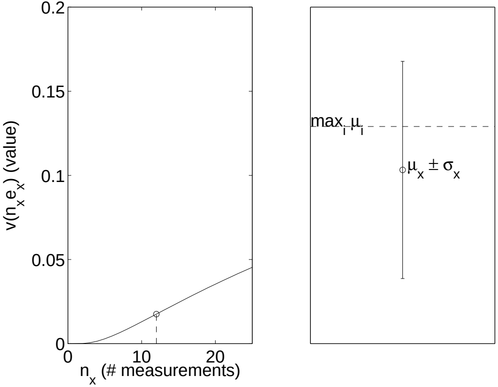
- We may also have an integrality constraint on n_x.
- Two approximation approaches are given in:
  - LL(B) [Chick & Inoue 2001].
  - OCBA for linear loss [He Chick Chen 2007].

Non-concavity of the value of information

- v is usually **not** concave.
- Non-concavity causes non-intuitive behavior by the optimal allocation.
- Non-concavity has consequenes for allocation heuristics.
- It is known that there are many cases outside of ranking & selection for which the value of information is not concave in the amount of information collected (see, e.g., [Howard 1966, 1988]).

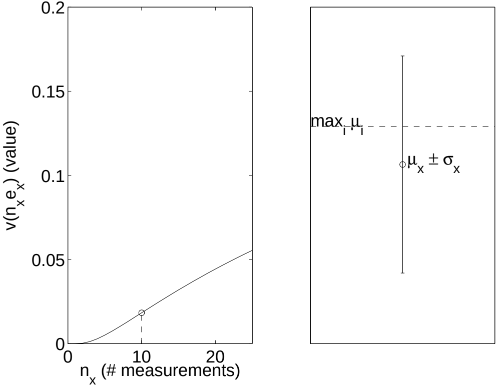
Measuring One Alternative

Plot shows the value $v(n_x e_x)$ of measuring only alternative x as a function of how many times it is measured. The measurement variance is $\lambda_x = 100$.



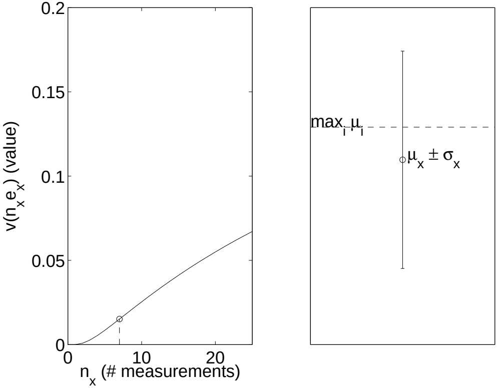
Measuring One Alternative

Plot shows the value $v(n_x e_x)$ of measuring only alternative x as a function of how many times it is measured. The measurement variance is $\lambda_x = 100$.



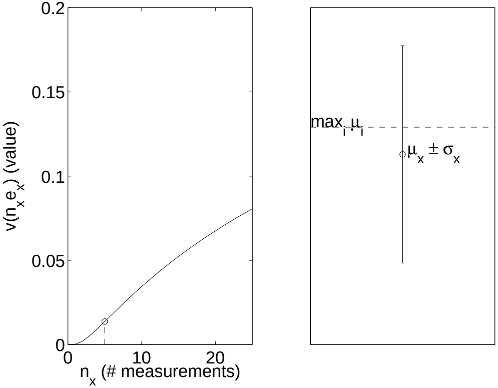
Measuring One Alternative

Plot shows the value $v(n_x e_x)$ of measuring only alternative x as a function of how many times it is measured. The measurement variance is $\lambda_x = 100$.



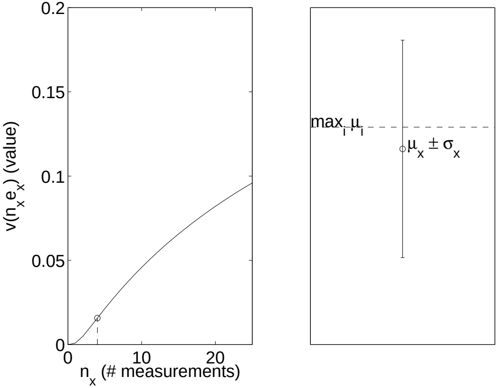
Measuring One Alternative

Plot shows the value $v(n_x e_x)$ of measuring only alternative x as a function of how many times it is measured. The measurement variance is $\lambda_x = 100$.



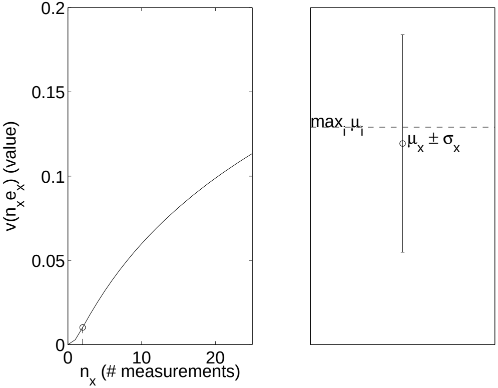
Measuring One Alternative

Plot shows the value $v(n_x e_x)$ of measuring only alternative x as a function of how many times it is measured. The measurement variance is $\lambda_x = 100$.



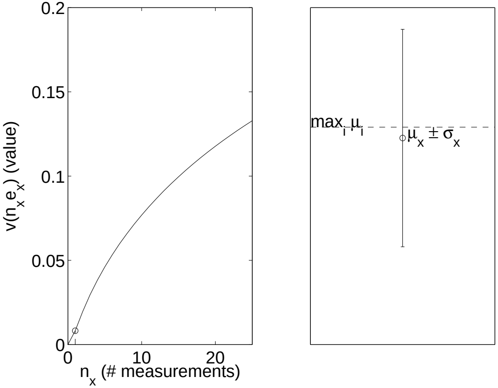
Measuring One Alternative

Plot shows the value $v(n_x e_x)$ of measuring only alternative x as a function of how many times it is measured. The measurement variance is $\lambda_x = 100$.



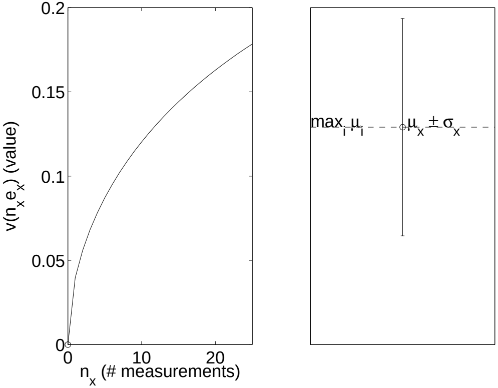
Measuring One Alternative

Plot shows the value $v(n_x e_x)$ of measuring only alternative x as a function of how many times it is measured. The measurement variance is $\lambda_x = 100$.



Measuring One Alternative

Plot shows the value $v(n_x e_x)$ of measuring only alternative x as a function of how many times it is measured. The measurement variance is $\lambda_x = 100$.



Measuring One Alternative

Let $\Delta_x = |\mu_x - \max_{x' \neq x} \mu_{x'}|$.

Theorem

$n_x \mapsto v(n_x e_x)$ is convex on $(0, n_x^*]$ and concave on (n_x^*, ∞), where

$$n_x^* = \frac{\lambda_x}{8\sigma_x^4}\left[\Delta_x^2 - \sigma_x^2 + \sqrt{\Delta_x^4 + 14\sigma_x^2\Delta_x^2 + \sigma_x^4}\right].$$

Corollary

If $\Delta_x = 0$, then $n_x \mapsto v(n_x e_x)$ is concave on $\mathbb{R}_{++}$.

The amount of non-concavity is increasing in Δ_x and λ_x.

Homogeneous Case

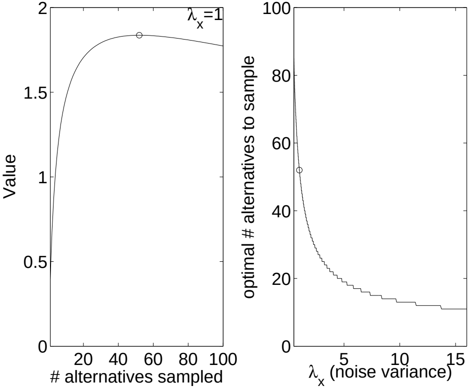
- Consider the homogeneous case, in which
  - We have M alternatives
  - $\mu_1 = \ldots = \mu_M$.
  - $\sigma_1^2 = \ldots = \sigma_M^2$.
- We might expect that the optimal allocation measures each alternative the same number of times... but this is not true.

Homogeneous Case

- Consider the homogeneous case, in which
  - We have M alternatives
  - $\mu_1 = \ldots = \mu_M$.
  - $\sigma_1^2 = \ldots = \sigma_M^2$.
- We might expect that the optimal allocation measures each alternative the same number of times... but this is not true.
- It may be better to randomly ignore some alternatives, and spread our budget over those that remain.

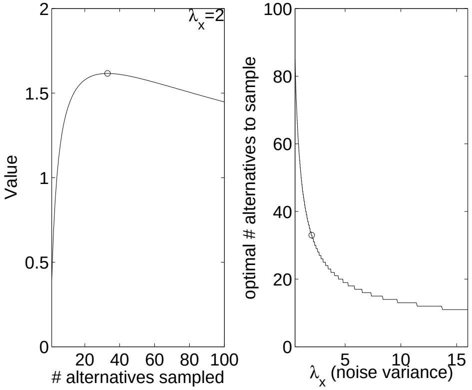
Homogeneous Case

$N = M = 100.$



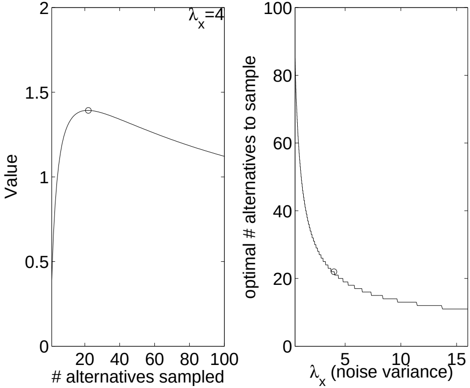
Homogeneous Case

$N = M = 100.$



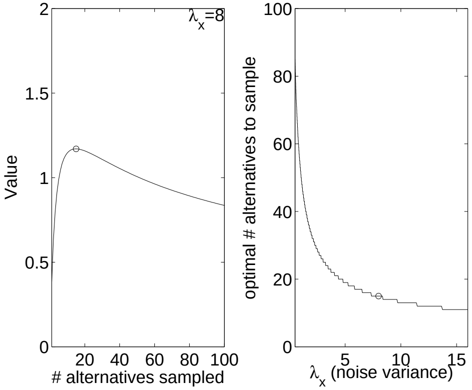
Homogeneous Case

$N = M = 100.$



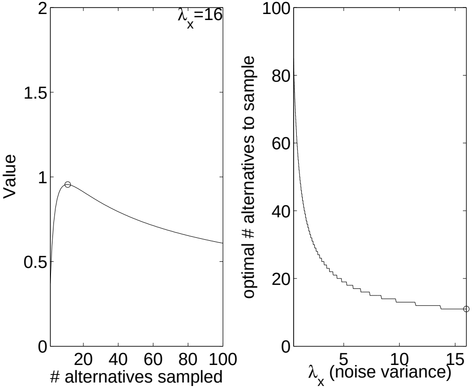
Homogeneous Case

$N = M = 100.$



Homogeneous Case

$N = M = 100.$



Implications for myopic policies for sequential allocation

- Now consider the **fully sequential** problem, where we choose each measurement based on all previous measurements.
- Let $\mu_x^n = \mathbb{E}_n \theta_x$ be the posterior mean of our belief about θ_x.
- x_{n+1} is the alternative to measure at time $n+1$, and depends upon all observations up to time n.
- The non-concavity of information has consquences for fully sequential policies as well.

Knowledge-Gradient Policy

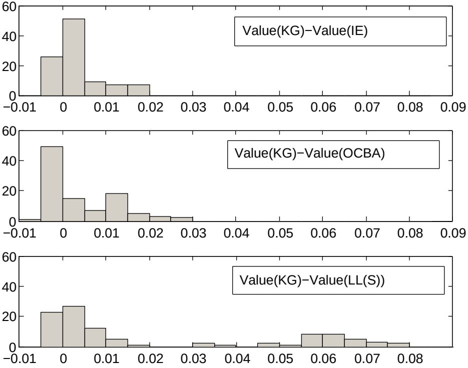
The **knowledge-gradient policy** (introduced as the (R1,...,R1) policy in [Gupta & Miescke 1996]) is defined to be the policy that chooses the x_{n+1} with the largest value of information.

$$\begin{aligned} x_{n+1} &\in \arg\max_x v(e_x) \\ &= \arg\max_x \mathbb{E}_n \left[\max_i \mu_i^{n+1} \mid x_{n+1} = x \right] - \max_i \mu_i^n. \end{aligned}$$

- This choice of x_{n+1} is optimal if $N = n+1$.
- μ_i^n is the expected value of i given our information at time n.
- $\max_i \mu_i^n$ is the best we can do given what we know at n.
- $\max_i \mu_i^{n+1}$ is the best we will be able to do given what we know at n and what we learn from our measurement x_{n+1}.

Knowledge-Gradient Policy

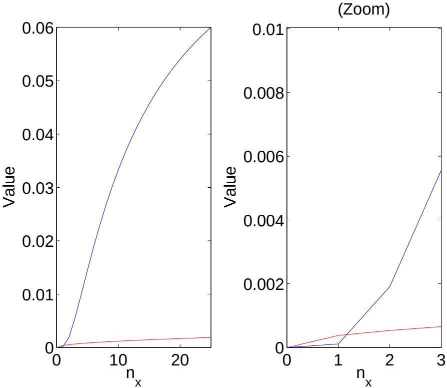
The KG policy is a reasonably good policy.



Histograms show difference in value between KG and other competing policies across 100 randomly generated priors.
Bars to the right of 0 are priors for which KG performed better.

Knowledge-Gradient Policy

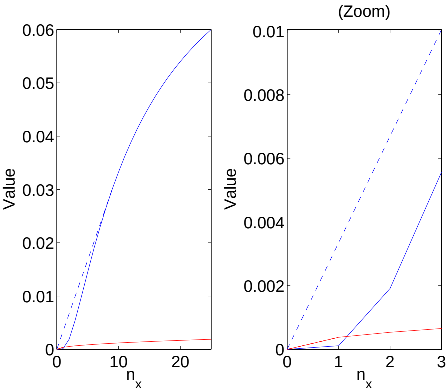
We will take more than one samples, but $v(e_x)$ and the KG policy ignore this. This can cause a problem.



- The blue alternative has $\Delta_x = 1$, $\sigma_x = 1.1$, $\lambda_x = 10$.
- The red alternative has $\Delta_x = 0$, $\sigma_x = 0.003$, $\lambda_x = 10$.
- The blue alternative is clearly better to measure, as long as $N > 1$.

Knowledge-Gradient Policy

We will take more than one samples, but $v(e_x)$ and the KG policy ignore this. This can cause a problem.



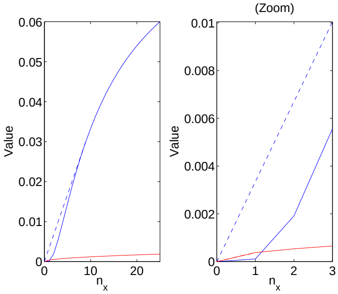
- The blue alternative has $\Delta_x = 1$, $\sigma_x = 1.1$, $\lambda_x = 10$.
- The red alternative has $\Delta_x = 0$, $\sigma_x = 0.003$, $\lambda_x = 10$.
- The blue alternative is clearly better to measure, as long as $N > 1$.

KG(*) Policy

We propose the KG(*) policy, which chooses

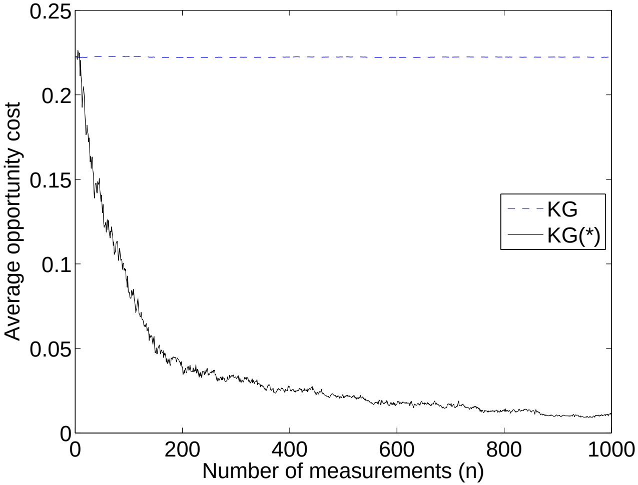
$$x_{n+1} \in \underset{m \in [1, N-n]}{\arg\max} \; v(me_x)/m.$$

- $v(me_x)/m$ is the **average value per measurement** when performing m measurements of x.
- $\max_{m \in [1, N-n]} v(me_x)/m$ is the **best possible** average value per measurement when measuring x.



Example

$\mu = [0, 0, -1]$, $\sigma^2 = [0, 10^{-4}, 2]$.
KG measures alternative 2, while KG(*) measures alternative 3.



KG(*) Policy

Proposition

$m \mapsto v(me_x)/m$ is strictly unimodal over $\mathbb{R}_{++}$ and its unique maximum m^* satisfies $\underline{m} \leq m^* \leq \overline{m}$, where

$$\underline{m} = \frac{\lambda_x}{4\sigma_x^2}\left(-1 + r + \sqrt{1 + 6r + r^2}\right),$$
$$\overline{m} = \frac{\lambda_x}{4\sigma_x^2}\left(1 + r + \sqrt{1 + 10r + r^2}\right),$$

and where $r = \Delta_x^2/\sigma_x^2$.

- The KG(*) policy requires calculating $\arg\max_{m \in [1, N-n]} v(me_x)/m$.
- We approximate this by $(\underline{m} + \overline{m})/2$, truncated above by $N - n$ and below by 1.

Thank You

Any questions?