ISE789
Sampling

19 October 2010

Paper

- Monte Carlo (Importance) Sampling Within Benders Decomposition Algorithm for Stochastic Linear Programs
- By Gerd Infanger
- Annals of Operations Research 39 (1992)

Outline

- Introduction
- Crude Monte Carlo
- Importance Sampling
- Importance Sampling in Benders
- Bounds (Clay and Omar)
  - Upper and Lower Bounds
  - Stopping
  - Confidence Bounds

Assumptions

- 2-Stage stochastic program
- Fixed recourse
- 2nd stage cost coefficients are deterministic
- Components of ξ or V are independent.

Introduction

- Main idea of importance sampling
  - Reduce time to achieve desired accuracy
- Need to sample if distribution of ξ is continuous or large and discrete
- Individual components of ξ don't have to have large discrete support if dimension is even modestly sized

Example

- ξ is 15-dimensional and each component is independent and uniformly distributed over the integers 1,2,…,10.
- How many scenarios are required to obtain universe solution?
- Why might this be a problem?
  - (say it takes 1 second to solve each LP…)

Crude Monte Carlo

$$z = \sum_{\omega \in \Omega} C(v^{\omega}) p(v^{\omega})$$

3.1. CRUDE MONTE CARLO

Suppose v^{ω}, $\omega = 1, \ldots, n$ are scenarios, sampled independently from their joint probability mass function, then $C^{\omega} = C(v^{\omega})$ are independent random variates with expectation z.

$$\bar{z} = (1/n) \sum_{\omega=1}^{n} C^{\omega} \tag{6}$$

is an unbiased estimator of z and its variance is

$$\sigma_{\bar{z}}^2 = \sigma^2/n,$$

$$\sigma^2 = \mathrm{var}(C(V)).$$

Thus, the standard error is decreasing with sample size n by $n^{-0.5}$. The convergence rate of $\bar{z}$ to z is independent of the dimension h of the random vector V.

Importance Sampling

- Reduce the variance of $\bar{z}$
  - Could just increase sample size
- Introduce a new probability mass function that gives greater weight to some scenarios

$$z = \sum_{\omega \in \Omega} C(v^{\omega})p(v^{\omega}) = \sum_{\omega \in \Omega} \frac{C(v^{\omega})p(v^{\omega})q(v^{\omega})}{q(v^{\omega})}$$

$$\bar{z} = \frac{1}{n} \sum_{\omega=1}^{n} \frac{C(w^{\omega})p(w^{\omega})}{q(w^{\omega})},$$

Intuitively…

- We want to sample scenarios which contribute most to the expectation
- Scenarios with high value
- Scenarios with high mass (pmf)
- Ideally, both!

How to choose q?

$$\mathrm{var}(\bar{z}) = \frac{1}{n} \sum_{\omega \in \Omega} \left(\frac{C(w^{\omega})p(w^{\omega})}{q(w^{\omega})} - z \right)^2 q(w^{\omega}).$$

- The perfect q requires what we are measuring

$$q^*(w^{\omega}) = \frac{C(w^{\omega})p(w^{\omega})}{\sum_{\omega \in \Omega} C(w^{\omega})p(w^{\omega})}$$

- Try to approximate the perfect q
  - Penalty for bad approximation is just higher sample size

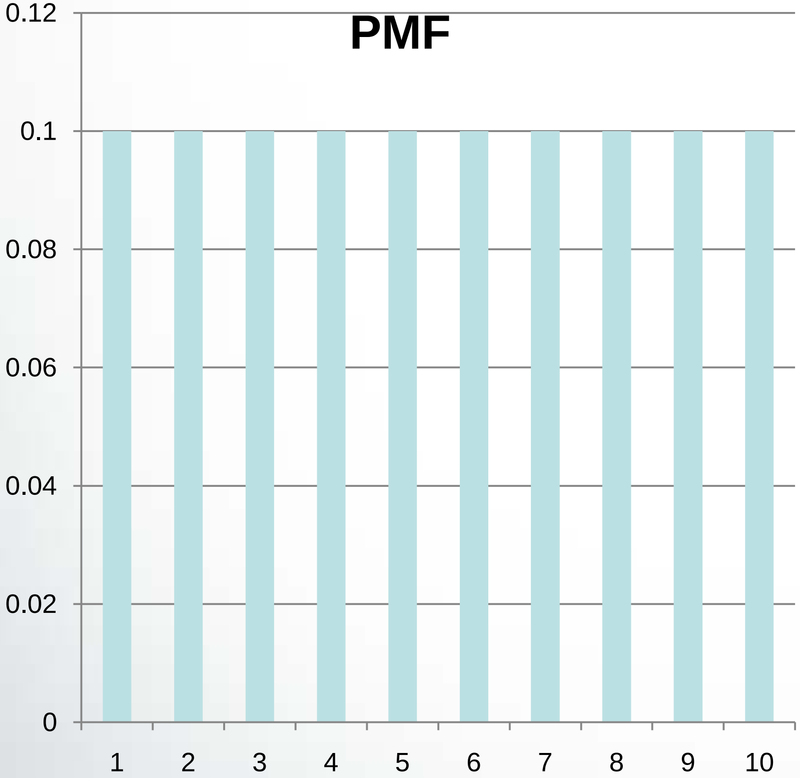
How to choose q?

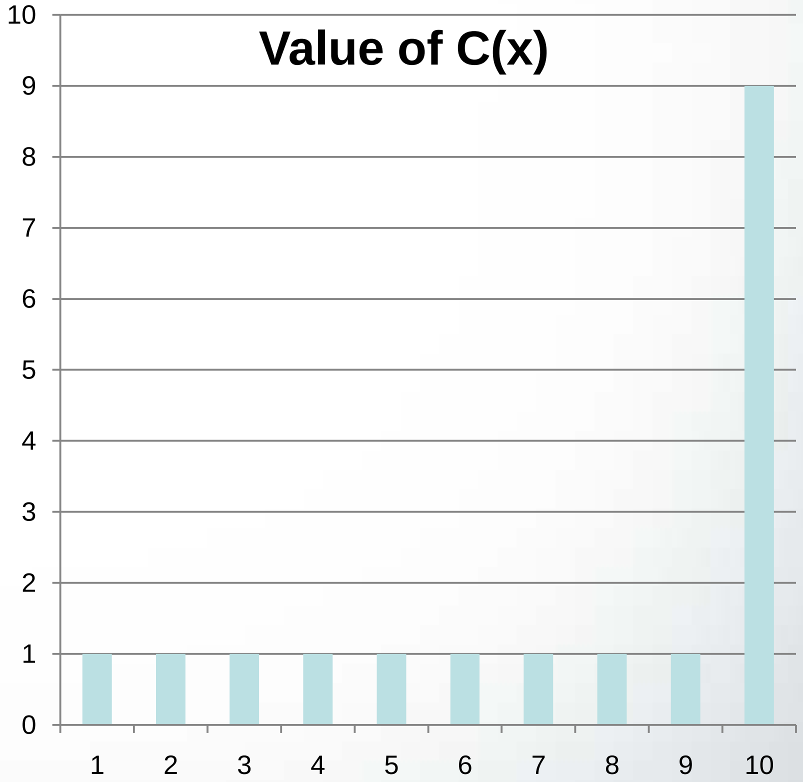
- Additivity reduces number of scenarios by isolating each component of ξ
  - No combinatorial explosion

$$C(V) \approx \sum_{i=1}^{h} C_i(V_i)$$

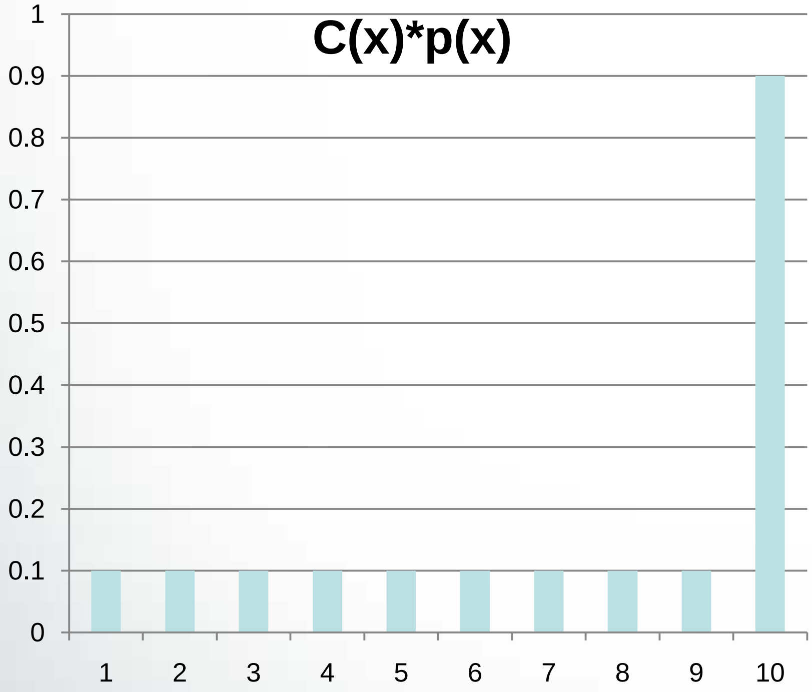
$$q(w^{\omega}) \approx \frac{C(w^{\omega})p(w^{\omega})}{\sum_{i=1}^{h}\sum_{\omega\in\Omega} C_i(w^{\omega})}.$$

Example:
Expectation of a random variable



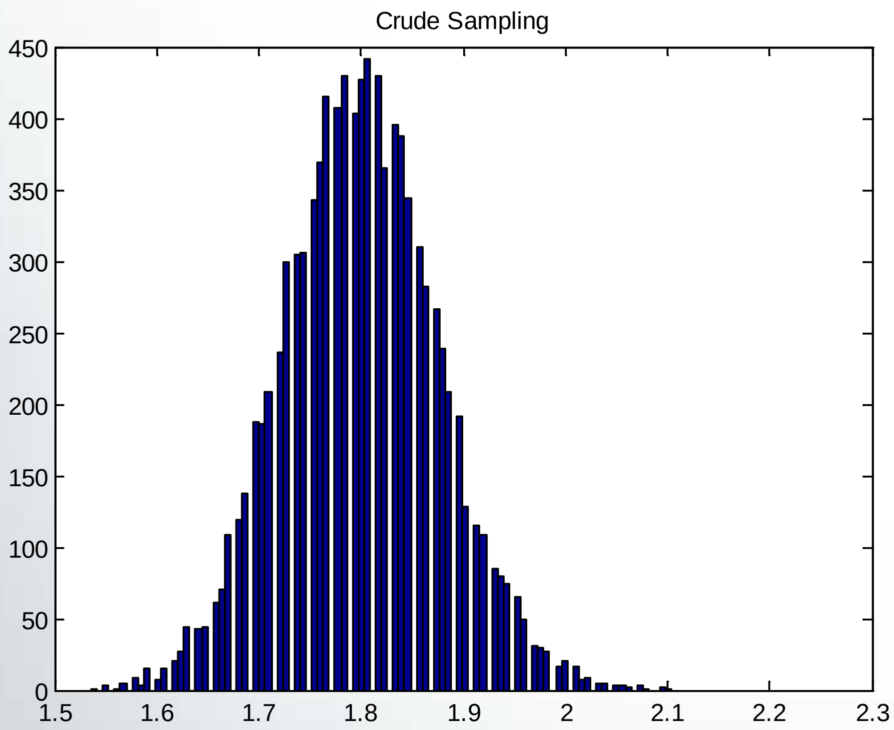


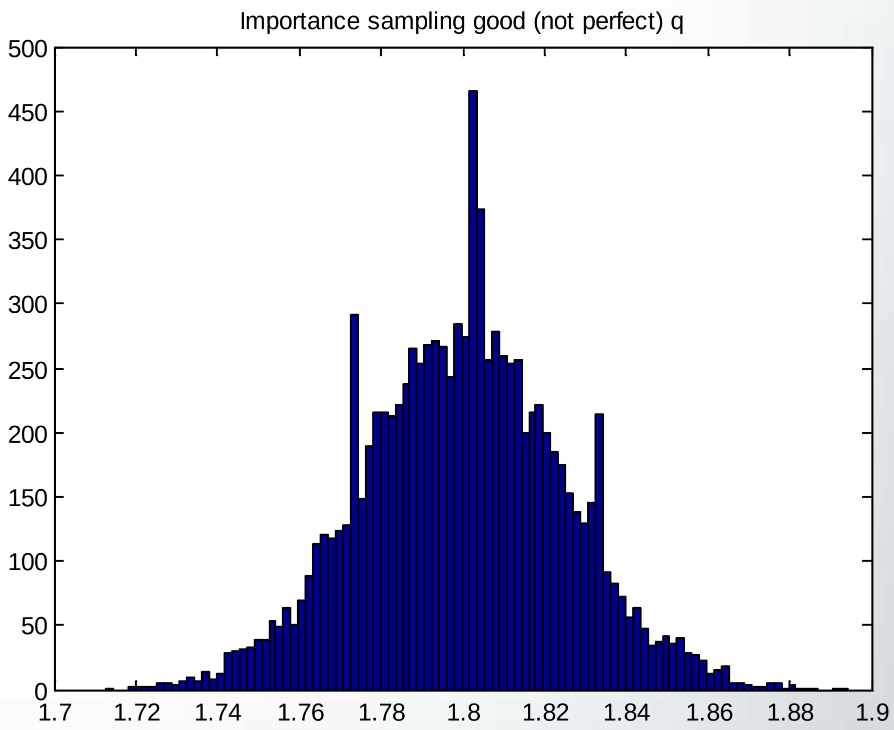
Importance of each scenario



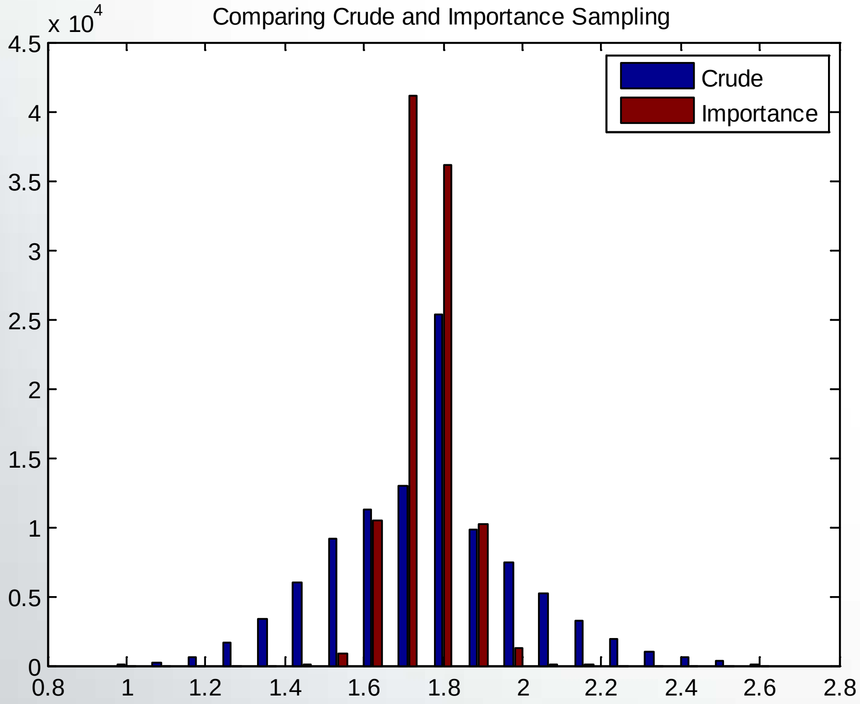
- Can you calculate q*?
  - Try it

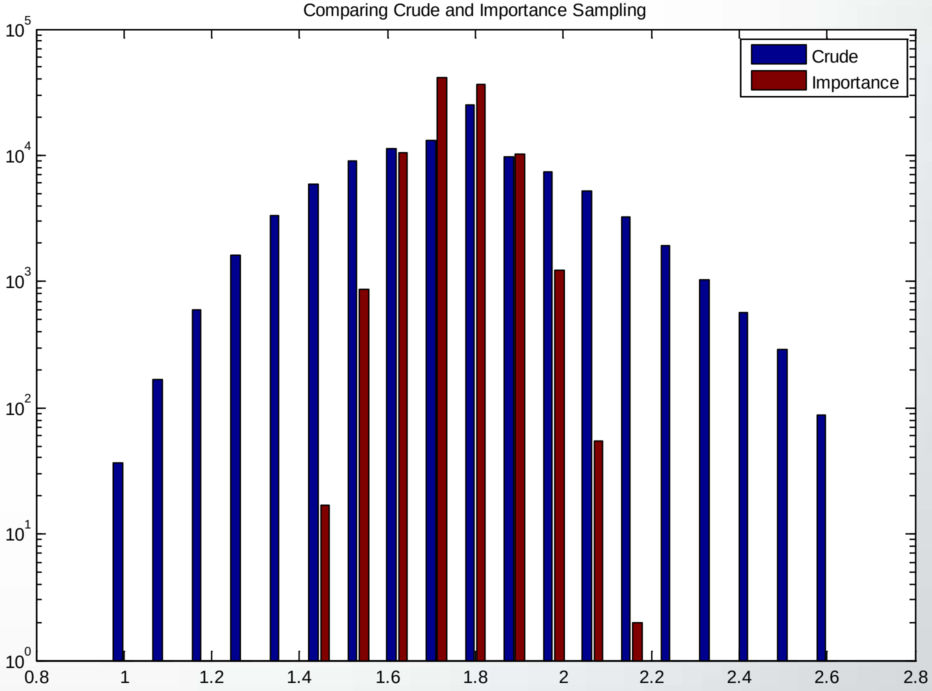
Matlab Example (1,000 points and 10,000 batches)



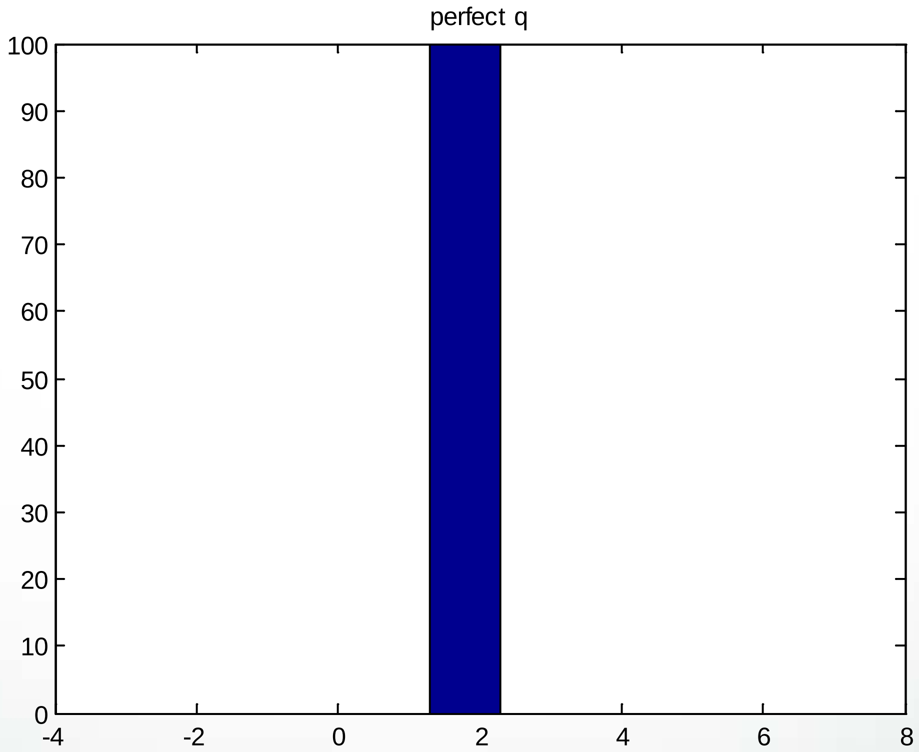


Matlab Example (1,00 points and 10,0000 batches)

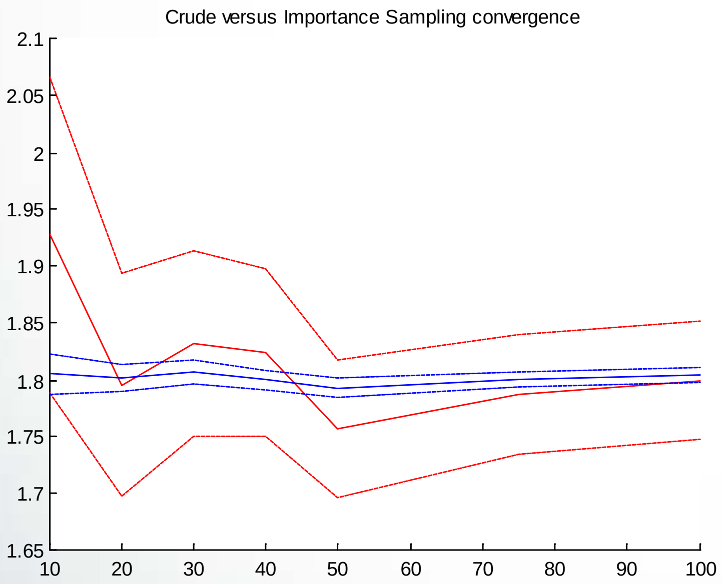


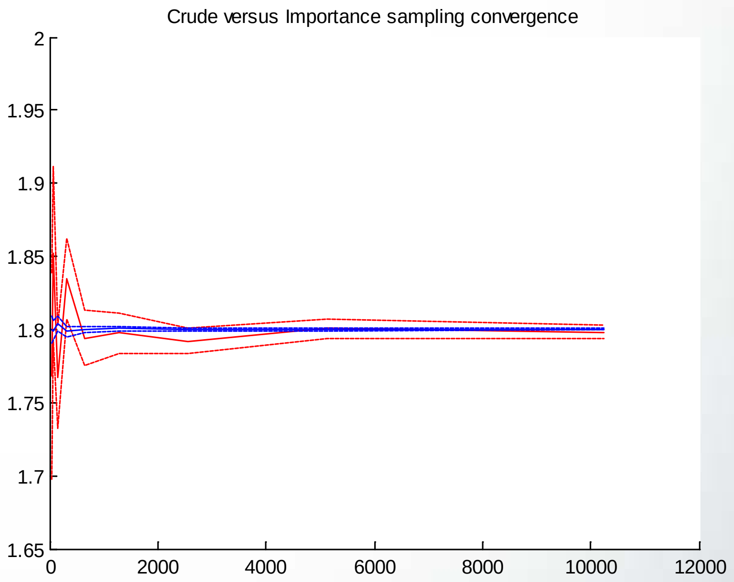


Matlab Example
(100 points and 100 batches)

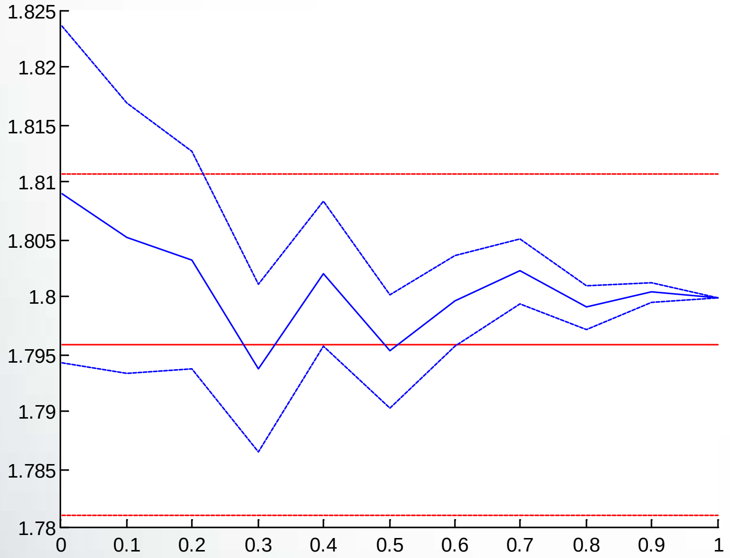


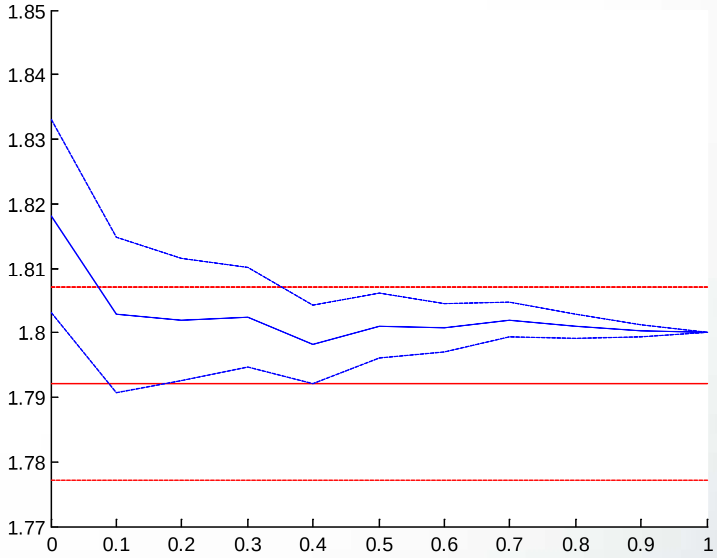
Convergence comparison





Importance of “good” q





How to choose q in practice?

- Entriken and Nakayama approach
  - Get huge reduction in variance for one experiment over Crude Monte Carlo
  - Via additive approximation and independence, we construct a base case
  - Then from the base case we can construct marginal cost functions
  - Then from those we can find weights (q values) which we sample from

$$C(V, \hat{X}) \approx \Gamma(V, \hat{X}) = C(\tau, \hat{X}) + \sum_{i=1}^{h} M_i(V_i, \hat{X}),$$
$$M_i(V_i, \hat{X}) = C(\tau_1, \ldots, \tau_{i-1}, V_i, \tau_{i+1}, \ldots, \tau_h, \hat{X}) - C(\tau, \hat{X}). \tag{7}$$

$$\bar{M}_i(\hat{X}) = E\, M_i(V_i, \hat{X}) = \sum_{\omega \in \Omega_i} M_i(v_i^{\omega}, \hat{X}) p(v_i^{\omega}) \tag{8}$$

$$F(v^{\omega}, \hat{X}) = \frac{C(v^{\omega}, \hat{X}) - C(\tau, \hat{X})}{\sum_{i=1}^{h} M_i(v_i^{\omega}, \hat{X})} \tag{9}$$

$$z(\hat{X}) = C(\tau, \hat{X}) + \sum_{i=1}^{h} \bar{M}_i(\hat{X}) \sum_{\omega \in \Omega} F(v^{\omega}, \hat{X}) \frac{M_i(v_i^{\omega}, \hat{X})}{\bar{M}_i(\hat{X})} \prod_{j=1}^{h} p_j(v_j^{\omega}). \tag{10}$$

B and L example, pgs 338-9

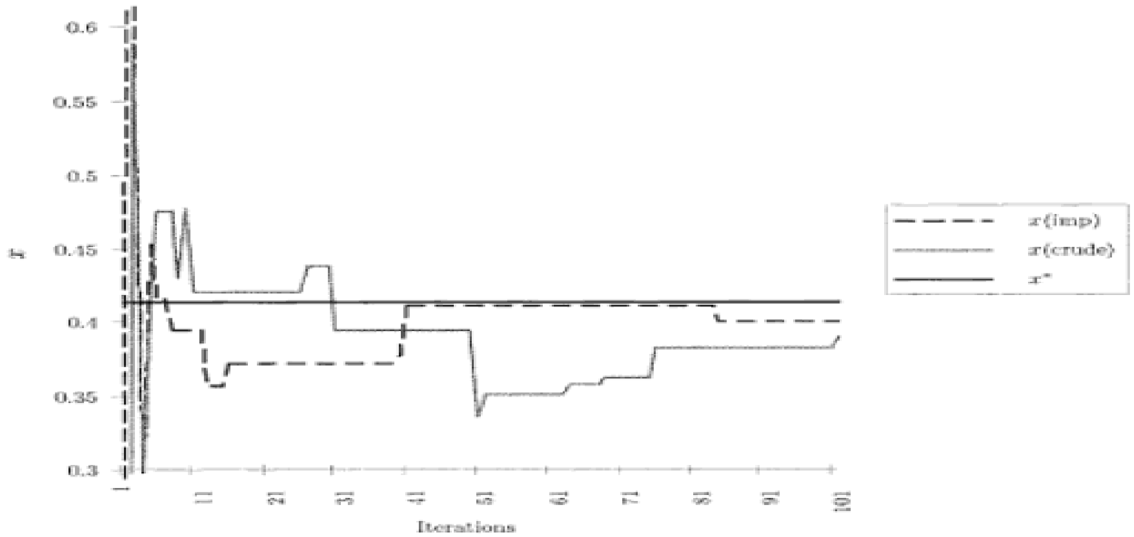


FIGURE 1. Solutions for crude Monte Carlo and importance sampling.

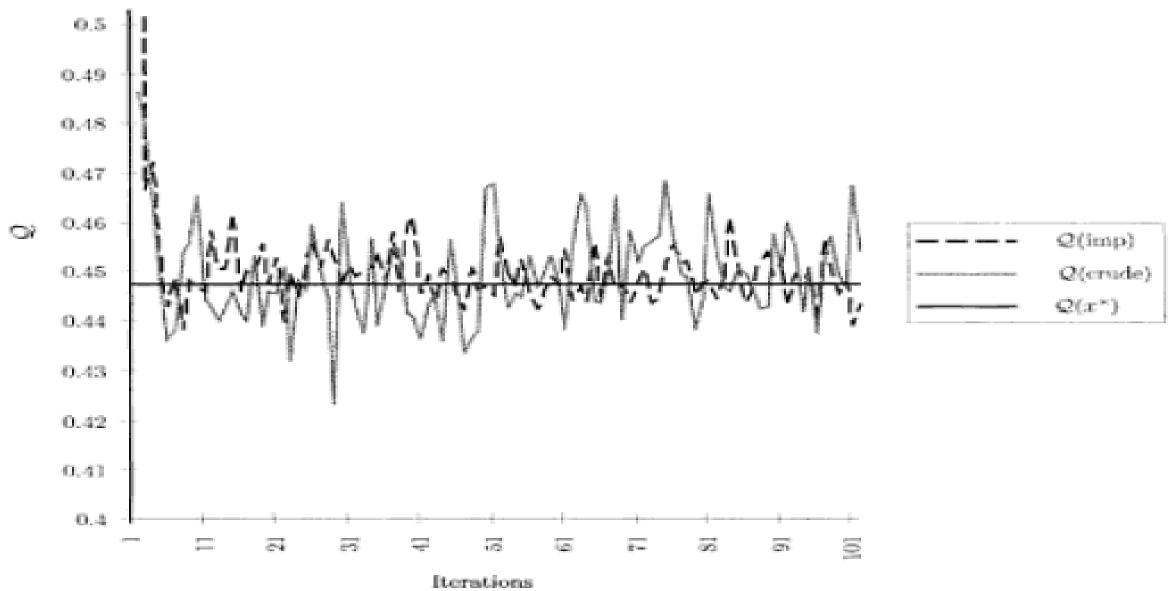


FIGURE 2. Objective values for crude Monte Carlo and importance sampling.

Other Good References on IS specifically

- Notes by Eric C Anderson
  - http://ib.berkeley.edu/labs/slatkin/eriq/classes/guest_lect/mc_lecture_notes.pdf
- P H Borcherds
  - http://iopscience.iop.org/0143-0807/21/5/305/pdf/ej0505.pdf

Other Good References on IS in SP

- Our Book
- Shapiro paper
  - Mathematical Programming 81 (1998)