# ISE

Industrial and Systems Engineering

**Reducing Bias in Stochastic Linear Programs
with Sampling Methods**

**Michael B. Freimer
Douglas J. Thomas
The Pennsylvania State University**

**Jeffrey T. Linderoth
Lehigh University**

**Report No. 05T-002**

LEHIGH
University

# Reducing Bias in Stochastic Linear Programs with Sampling Methods

Michael B. Freimer, Douglas J. Thomas

The Smeal College of Business, The Pennsylvania State University, {460, 463} Business Building, State College, Pennsylvania 16802-3603, USA, {mbf10@psu.edu, dthomas@psu.edu}

Jeffrey T. Linderoth

Industrial and Systems Engineering, Lehigh University, 200 W. Packer Ave., Bethlehem, PA 18015, USA, {jtl3@lehigh.edu}

Two-stage stochastic linear programs can be solved approximately by drawing a subset of all possible random scenarios and solving the problem based on this subset, an approach known as sample path optimization. Sample path optimization creates two kinds of objective function bias. First, the expected optimal objective function value for the sampled problem is lower (for minimization problems) than the optimal objective function value for the true problem. Second, if the stage-one decision from the solution to a sampled problem is implemented, the expected objective function value achieved is greater than the optimal objective value for the full problem. We investigate how two alternative sampling techniques, antithetic variates and Latin Hypercube sampling, affect these two biases relative to the alternative of drawing samples independently. We focus primarily on the first of these two types of bias, although we also characterize the bias in expected actual cost. For a simple example, we analytically express the reductions in bias obtained by these two sampling methods. We provide a general condition under which using antithetic variates reduces the bias of the expected optimal objective function value for the sampled problem. For seven test problems from the literature, we computationally investigate the bias impact of these sampling methods.

*Key words:* stochastic programming; sample path optimization; antithetic variates; Latin Hypercube sampling

*History:*

## 1. Introduction

Two-stage stochastic linear programming problems arise in a variety of production and inventory planning settings. At the first stage, values are chosen for a set of design variables;

for example, the design variables may represent a set of line capacities. The objective function of the first-stage problem requires us to evaluate the expected value of the solution to a second-stage linear program (LP), some of whose parameters (e.g., demand) are stochastic. Furthermore, the design variables from the first stage appear in the constraints of the second-stage LP. Early formulations of this problem were given by Dantzig (1955) and Beale (1955). The motivation for this paper is the efficient solution of such two-stage design problems. In the remainder of this introduction we discuss some alternative solution techniques and set the stage for the results of the subsequent sections. Throughout the paper we adopt a modified version of the notation for the two-stage stochastic programming problem presented by Kleywegt and Shapiro (2001):

$$\text{MP}: \quad z^*_{\text{MP}} \stackrel{\text{def}}{=} \min_x \mathbb{E}_\omega\left[Q(x,\omega)\right] + g(x), \text{ s.t. } Ax = b, \ x \geq 0,$$

where $g(x)$ is a deterministic function of $x$, and $Q(x,\omega)$ represents the optimal objective function value of the second-stage problem:

$$\text{P}: \quad Q(x,\omega) \stackrel{\text{def}}{=} \min_y q(\omega)^T y, \text{ s.t } T(\omega)x + W(\omega)y = h(\omega), \ y \geq 0.$$

Here $q(\omega) \in \mathbb{R}^n, T(\omega) \in \mathbb{R}^{\ell \times m}, W(\omega) \in \mathbb{R}^{\ell \times n}$, and $h(\omega) \in \mathbb{R}^\ell$ may be random (i.e., functions of the realization $\omega$). When $g(x) = c^T x$ and $W(\omega)$ is deterministic, we have a two-stage stochastic linear program with fixed recourse.

For most problems of interest, the objective function of MP,

$$\mathbb{E}_\omega\left[Q(x,\omega)\right] + g(x), \tag{1.1}$$

is a nonlinear function of the decision vector $x$, and standard gradient-based nonlinear programming techniques may be applied to MP. Given the stochastic nature of the objective function however, it may be impossible to evaluate $\mathbb{E}_\omega\left[Q(x,\omega)\right]$ or its subgradient exactly. For such problems, one might hope to find an unbiased estimator for this subgradient.

Sample path optimization is a common approach for solving MP that avoids this difficulty. The idea is to draw $N$ realizations (sample paths) of problem MP and to optimize over this representative sample. More specifically, let $MP_N(\omega_1, \ldots, \omega_N)$ denote a realization of the $N$-sample path problem. That is,

$$\text{MP}_N: \quad z^*_{\text{MP}_N(\omega_1,\ldots,\omega_N)} \stackrel{\text{def}}{=} \min_x N^{-1} \sum_{i=1}^{N} Q_i(x,\omega_i) + g(x), \text{ s.t. } Ax = b, \ x \geq 0,$$

2

where $Q_i(x, \omega_i)$ represents the optimal objective function value of the problem:

$$Q_i(x, \omega_i) \stackrel{\text{def}}{=} \min_y q(\omega_i)^T y, \text{ s.t. } T(\omega_i)x + W(\omega_i)y = h(\omega_i), \ y \geq 0.$$

The problem $MP_N$ can also be rewritten as the following problem:

$$\text{MP}'_N : \min_{x, y_1, \ldots y_N} N^{-1} \sum_{i=1}^N q(\omega_i)^T y_i + g(x), \text{ subject to:}$$

$$Ax = b, \ x \geq 0, \ T(\omega_i)x + W(\omega_i)y_i = h(\omega_i), \ y_i \geq 0, \ i = 1, 2, \ldots N.$$

Problem $\text{MP}_N$ or $\text{MP}'_N$ can be used as an approximation to the original problem MP, and both are likely to be easier to solve than MP. (If $g(\cdot)$ is linear, $MP'_n$ is a linear program.)

Under fairly general conditions, the solution to $\text{MP}_N$ approaches that of MP with probability 1 as the number of realizations $N$ increases (Dupačová and Wets, 1988). However, the solution to $MP_N$ is biased in the sense that the expectation of the optimal objective function value of $MP_N$ is less than that of $MP$. Mak, Morton, and Wood (1999) show that:

$$\mathbb{E}_{(\omega_1, \ldots, \omega_N)} \left[ z^*_{\text{MP}_N(\omega_1, \ldots, \omega_N)} \right] \leq \mathbb{E}_{(\omega_1, \ldots, \omega_{N+1})} \left[ z^*_{\text{MP}_{N+1}(\omega_1, \ldots, \omega_{N+1})} \right] \leq z^*_{\text{MP}} \ \forall N. \qquad (1.2)$$

A related issue is that the optimal solution $x^*_N(\omega_1, \ldots, \omega_N)$ of $\text{MP}_N$ may be suboptimal with respect to the objective function $\mathbb{E}_\omega [Q(x, \omega)] + g(x)$ of MP. We refer to:

$$\mathbb{E}_\omega \left[ Q(x^*_{\text{MP}_N(\omega_1, \ldots, \omega_N)}, \omega) \right] + g \left( x^*_{\text{MP}_N(\omega_1, \ldots, \omega_N)} \right)$$

as the *actual* cost of the sample path problem and $z^*_{\text{MP}_N(\omega_1, \ldots, \omega_N)}$ as the *perceived* cost.

In Section 2 we present an analytic example of both of these difficulties based on the newsvendor problem, and in Sections 2.1 and 2.2 we describe how the situation may be improved using antithetic variates (AV) and Latin Hypercube (LH) sampling. These are sampling techniques usually prescribed for reducing the variance of an unbiased estimator. Suppose $X(\omega)$ is a random variable (e.g., a component of the data $\{q, T, W, h\}$) having invertible cdf $F$. Under independent sampling (IS) we generate $N$ independent numbers $\{U_1, \ldots U_N\}$ uniformly distributed on [0,1], and

$$\hat{\psi}_{IN} \stackrel{\text{def}}{=} N^{-1} \sum_{i=1}^N \left[ \psi(F^{-1}(U_i)) \right] \qquad (1.3)$$

is an unbiased estimate of $\mathbb{E}[\psi(X(\omega))]$ for an arbitrary function $\psi$. Under AV, rather than drawing $N$ independent values $\{U_1, \ldots U_N\}$, we draw $N/2$ antithetic pairs $\{(U_i, 1 - U_i), i =$

3

$1, 2, \ldots N/2\}$ and combine these $N$ values to compute $\hat{\psi}_{AV}$ via (1.3). If $\psi$ is monotone, $Var[\hat{\psi}_{AV}] \leq Var[\hat{\psi}_{IN}]$ (Law and Kelton, 2000). Under LH, the interval [0,1] is divided into $N$ segments, $[(i-1)/N, i/N]$, $i = 1 \ldots, N$, and a sample is generated uniformly from each segment. These samples are shuffled to obtain $U_1, \ldots, U_N$.

Higle (1998) investigates the use of AV and other techniques to reduce the variance of

$$N^{-1} \sum_{i=1}^{N} Q_i(x, \omega_i) + g(x),$$

which is an unbiased estimate of of $\mathbb{E}_\omega[Q(x, \omega)] + g(x)$ for an arbitrarily chosen value of $x$. In Section 2.1 we estimate $\mathbb{E}_\omega[Q(x, \omega)] + g(x)$ at the unknown *optimal* value $x^*_{\mathrm{MP}}$ using the solution to the sample path problem $\mathrm{MP}_N$. The solution to this problem, $z^*_{\mathrm{MP}_N}$, is by (1.2) a biased estimate of $\mathbb{E}_\omega[Q(x^*_{\mathrm{MP}}, \omega)] + g(x^*_{\mathrm{MP}})$ $(= z^*_{\mathrm{MP}})$. However we show that using AV or LH to compute $z^*_{\mathrm{MP}_N}$ can reduce this bias (as well as mean squared error) for our analytic example. In a related paper, Linderoth et al. (2002) examine the impact of LH on the bias of $z^*_{\mathrm{MP}_N}$ and on an upper bound for $z^*_{\mathrm{MP}}$ with a set of empirical examples. This paper extends that work by providing analytical evidence of bias reduction and performing more extensive computational work, including results showing bias reduction effects with antithetic variates.

In Section 3 we demonstrate results similar to those of Section 2 using a series of computational examples.

## 2. Sample Path Optimization

In this section we discuss the difficulties associated with sample path optimization described in the introduction. We begin with an analytic example based on the newsvendor problem, which can be expressed as a two-stage stochastic program as follows. In the first stage we choose an order quantity $x$. After demand $D$ has been realized, we decide how many of the available papers $y$ to sell. Assume demand is uniformly distributed on the interval $[0, 1]$, and there is a shortage cost $\alpha \in (0, 1)$ and an overage cost $1 - \alpha$. The second stage problem is

$$\text{P}: \quad Q(x, D) \stackrel{\text{def}}{=} \min_y \{(1 - \alpha)(x - y) + \alpha(D - y) \mid y \leq x, y \leq D\}.$$

The solution to P is $\min(x, D)$. Let $TC(x)$ be the expected total cost associated with order quantity $x$:

$$TC(x) \stackrel{\text{def}}{=} \mathbb{E}[Q(x, D)]$$
$$= \mathbb{E}\left[\min_y \{(1 - \alpha)(x - y) + \alpha(D - y) \mid y \le x, y \le D\}\right], \qquad (2.1)$$

so MP is $\min_x TC(x)$. Furthermore:

$$\begin{aligned} TC(x) &= (1 - \alpha)\mathbb{E}(x - D)^+ + \alpha\mathbb{E}(D - x)^+ \\ &= (1 - \alpha)\int_0^x (x - z)dz + \alpha \int_x^1 (z - x)dz \\ &= (1 - \alpha)\frac{x^2}{2} + \alpha\frac{(1 - x)^2}{2}. \end{aligned}$$

The cost-minimizing solution is therefore $x^* = \alpha$, and the optimal expected total cost is:

$$TC^* \stackrel{\text{def}}{=} TC(\alpha) = (1 - \alpha)\frac{\alpha^2}{2} + \alpha\frac{(1 - \alpha)^2}{2} = \frac{\alpha(1 - \alpha)}{2}. \qquad (2.2)$$

The $N$-sample path version of this problem is:

$$z^*_{\text{MP}_N(D_1,\ldots,D_N)} \stackrel{\text{def}}{=} \min_x N^{-1} \sum_{i=1}^N \left[(1 - \alpha)(x - D_i)^+ + \alpha(D_i - x)^+\right]. \qquad (2.3)$$

The optimal solution $\hat{x}$ to (2.3) is the $\lceil \alpha N \rceil^{\text{th}}$ order statistic of the demands $\{D_1, \ldots, D_N\}$. Therefore $\hat{x}$ has a Beta distribution with parameters $\lceil \alpha N \rceil$ and $(N - \lceil \alpha N \rceil + 1)$, so:

$$\mathbb{E}[\hat{x}] = \frac{\lceil \alpha N \rceil}{N + 1} \quad \text{and} \quad \mathbb{E}[\hat{x}^2] = \frac{\lceil \alpha N \rceil(\lceil \alpha N \rceil + 1)}{(N + 1)(N + 2)}. \qquad (2.4)$$

We next examine the expected performance of $\hat{x}$ with respect to the original objective function $TC(\cdot)$. The expected *actual* total cost using the sample path optimization solution is:

$$\begin{aligned} \mathbb{E}_{\hat{x}}[TC(\hat{x})] &= \int_0^1 TC(z)f_{\hat{x}}(z)dz = \frac{1 - \alpha}{2}\mathbb{E}[\hat{x}^2] + \frac{\alpha}{2}\left(1 - 2\mathbb{E}[\hat{x}] + \mathbb{E}[\hat{x}^2]\right) \\ &= \frac{1}{2}\mathbb{E}[\hat{x}^2] - \alpha\mathbb{E}[\hat{x}] + \frac{\alpha}{2}. \end{aligned} \qquad (2.5)$$

The expected *perceived* cost of the sample path optimization solution (i.e., $\mathbb{E}[z^*_{\text{MP}_N(D_1,\ldots,D_N)}]$)

is:

$$\mathbb{E}_{D_1,\ldots D_N}\left[N^{-1}\sum_{i=1}^{N}(1-\alpha)(\hat{x}-D_i)^+ + \alpha(D_i-\hat{x})^+\right]$$

$$= \int_0^1\left[(1-\alpha)\left(\frac{\lceil\alpha N\rceil-1}{N}\right)\frac{1}{u}\int_0^u(u-z)dz + \alpha\left(\frac{N-\lceil\alpha N\rceil}{N}\right)\frac{1}{1-u}\int_u^1(z-u)dz\right]f_{\hat{x}}(u)du$$

$$= \frac{(1-\alpha)}{2}\left(\frac{\lceil\alpha N\rceil-1}{N}\right)\mathbb{E}\left[\hat{x}\right] + \frac{\alpha}{2}\left(\frac{N-\lceil\alpha N\rceil}{N}\right)(1-\mathbb{E}\left[\hat{x}\right]). \quad (2.6)$$

We derive the second line by conditioning on the value of $\hat{x}$, the $\lceil\alpha N\rceil^{\text{th}}$order statistic, in which case $\lceil\alpha n\rceil - 1$ of the demand values are uniformly distributed below $\hat{x}$, and the remaining $N - \lceil\alpha N\rceil$ are distributed above.

If $\alpha N$ is integer, the expected actual cost of the sample path solution (SPS) computed by substituting (2.4) into (2.5) is:

$$\frac{\alpha}{2}\left[1-\frac{\alpha N^2+4\alpha N-N}{(N+1)(N+2)}\right],$$

and the expected perceived cost of the sample path solution computed by substituting (2.4) into (2.6) is:

$$\frac{\alpha(1-\alpha)}{2}\left(\frac{N}{N+1}\right).$$

Therefore using (2.2):

$$\frac{\text{Expected Actual Cost SPS}}{\text{Optimal Cost}} = \left(\frac{1}{1-\alpha}\right)\left(1-\frac{\alpha N^2+4\alpha N-N}{(N+1)(N+2)}\right), \text{ and} \quad (2.7)$$

$$\frac{\text{Expected Perceived Cost SPS}}{\text{Optimal Cost}} = \frac{N}{N+1}. \quad (2.8)$$

Expressions (2.7) and (2.8) approach 1 (from above and below respectively) as $N$ increases.

## 2.1   The Use of Antithetic Variates to Reduce Bias

Some of the bias in expressions (2.7) and (2.8) can be reduced with the use of antithetic variates (AV), a technique usually prescribed for reducing the variance of an unbiased estimator. For our newsvendor problem with uniform [0,1] demand, we draw $N/2$ antithetic pairs $\{(D_i, 1-D_i), i=1,2,\ldots N/2\}$, rather than $N$ independent values $\{D_1,\ldots D_N\}$. These correlated values are used in the sample path problem (2.3).

In the subsequent analysis we suppose $\alpha > 0.5$, although similar computations can be performed for lower values. In this case $\hat{x}_{AV}$, the solution to the sample path problem

6

with AV, is the $\lceil \alpha N - N/2 \rceil^{\text{th}}$ order statistic of $N/2$ random variables uniformly distributed on $[0.5, 1]$. Hence $\hat{x}_{AV} = 1/2 + X/2$, where $X$ has a Beta distribution with parameters $\lceil \alpha N - N/2 \rceil$ and $N/2 - \lceil \alpha N - N/2 \rceil + 1$. If $\alpha N$ and $(\alpha N - N/2)$ are integers, then $\mathbb{E}[\hat{x}] < \mathbb{E}[\hat{x}_{AV}] < \alpha$, so the expectation of the AV solution is closer to the optimal solution of the original problem. Furthermore $\text{Var}[\hat{x}_{AV}] < \text{Var}[\hat{x}]$.

To derive the expected perceived cost under AV, we condition on the value of $\hat{x}_{AV}$. The three terms in the integrand correspond to the antithetic partner of $\hat{x}_{AV}$, those demand values lying below $\hat{x}_{AV}$, and those demand values lying above $\hat{x}_{AV}$ (but whose antithetic partners lie below):

$$\int_{0.5}^{1} [N^{-1}(1-\alpha)(u-(1-u))$$
$$+ N^{-1}(1-\alpha)(\lceil \alpha N \rceil - N/2 - 1)\left(\frac{1}{u-1/2}\right)\int_{0.5}^{u}((u-z)+(u-1+z))dz$$
$$+ N^{-1}(N - \lceil \alpha N \rceil)\frac{1}{1-u}\int_{u}^{1}(\alpha(z-u)+(1-\alpha)(u-1+z))dz ] f_{\hat{x}_{AV}}(u)du$$
$$= \frac{(-2N\alpha + N + \lceil \alpha N \rceil)\mathbb{E}[\hat{x}_{AV}] - \lceil \alpha N \rceil + \alpha N}{2N}.$$

The expected actual cost under AV is computed from (2.5) using the distribution of $\hat{x}_{AV}$. Figure 1 plots the expected actual and perceived costs as percentages of the optimal cost with both independent sample paths and antithetic pairs of sample paths, using a cost ratio of $\alpha = 0.8$. (To facilitate comparison across sampling methods, this figure and Figure 3 below include lines for Latin Hypercube sampling which will be discussed in the next section.) Note that the use of antithetic pairs reduces the gaps between the expected cost of the optimal solution and both the actual and perceived costs of the sample path solution.

In this example, with cost ratio $\alpha = 0.8$, the use of AV also reduces the variance of the optimal objective function estimator for the stochastic LP (i.e., the variance of the perceived cost $z^*_{\text{MP}_N}$), although we show below that this is not always the case. In a related paper, Higle (1998) investigates variance reduction in the estimation of $\mathbb{E}_\omega[Q(x,\omega)]$ for a fixed value of $x$. Here we are estimating $\mathbb{E}_\omega[Q(x,\omega)]$ at the optimal value of $x$, in other words we are using $z^*_{\text{MP}_N}$ to estimate $\mathbb{E}_\omega[Q(x^*_{\text{MP}},\omega)]$.

As with the bias, the change in variance for the newsvendor problem can be computed exactly; a derivation is included in the Appendix. Interestingly, while the use of AV decreases the bias for all values of $\alpha$ in the range $(0.5, 1)$, it increases the variance for some values of $\alpha$ and $N$. The combination of the two effects is reflected in the Mean Squared Error (MSE).
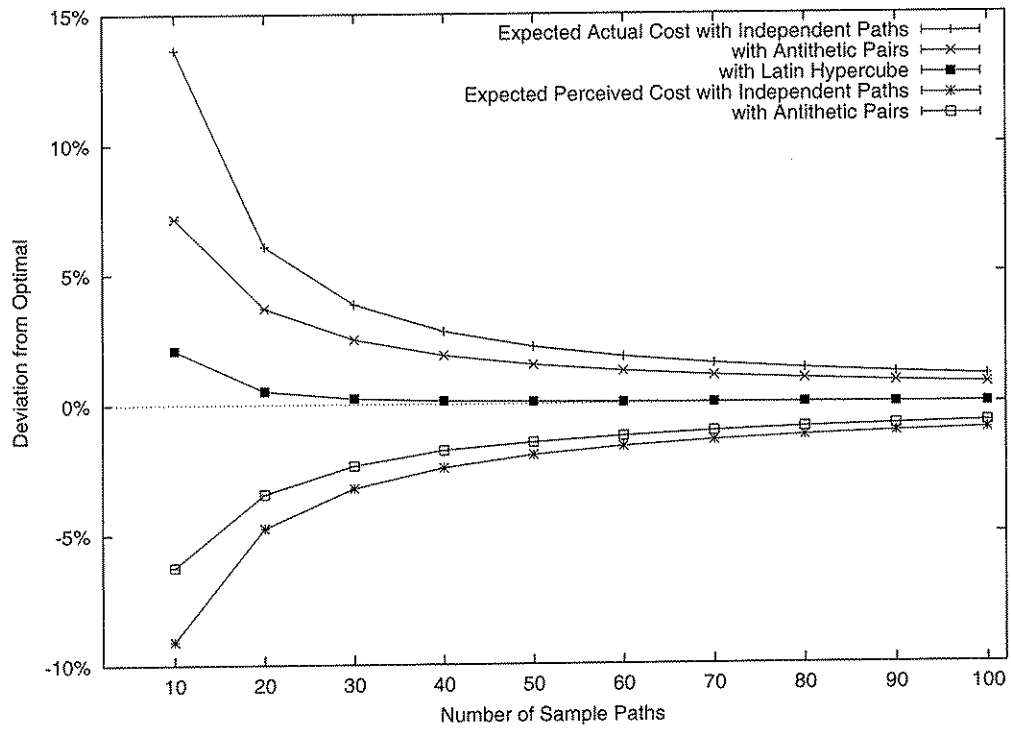
7

Figure 1: Expected performance of sample path solution under IS, AV, and LH for the newsvendor problem, as a function of the number of sample paths

Figure 2 shows the change in MSE obtained by using antithetic pairs. Note that for $\alpha = 0.6$ and $\alpha = 0.7$, use of antithetic pairs increases MSE (the "reduction" is negative).
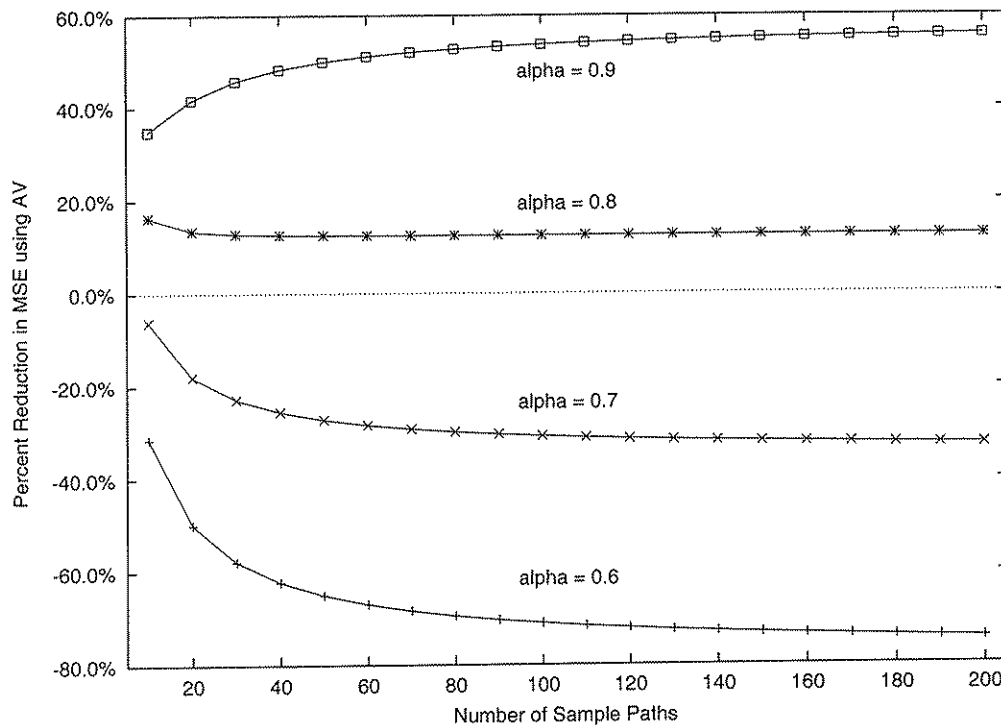


Figure 2: Change in MSE obtained by using antithetic variates instead of independent samples for the newsvendor problem, as a function of the number of sample paths

If $\alpha N$ is integer, the expected actual cost of the sample path solution under AV simplifies to:

$$\frac{\alpha(1-\alpha)N^2 + (8\alpha - 1)(1 - \alpha)N + 2}{2(N+2)(N+4)},$$

and the expected perceived cost under AV is:

$$\frac{1-\alpha}{2} \frac{\alpha N + 1}{N+2}.$$

Therefore using (2.2), under AV we have:

$$\frac{\text{Expected Actual Cost}}{\text{Optimal Cost}} = \frac{N^2}{(N+2)(N+4)} + \frac{(8\alpha - 1)(1 - \alpha)N + 2}{\alpha(1-\alpha)(N+2)(N+4)}; \qquad (2.9)$$

$$\frac{\text{Expected Perceived Cost}}{\text{Optimal Cost}} = \frac{\alpha N + 1}{\alpha(N+2)}. \qquad (2.10)$$

9

Expressions (2.9) and (2.10) approach 1 (from above and below respectively) as $N$ increases.

In addition to the expected performance of the sample path solution demonstrated in Figure 1, we may be interested in the distribution of its actual performance. Figure 3 plots the probability that the actual performance of the sample path solution will be within $x\%$ of the optimal solution, for various values of $x$, and with $N = 10$. Plots for both independent sampling and antithetic variates are shown, based on the distributional assumptions for $\hat{x}$ and $\hat{x}_{AV}$ given above and a cost ratio of 0.8.
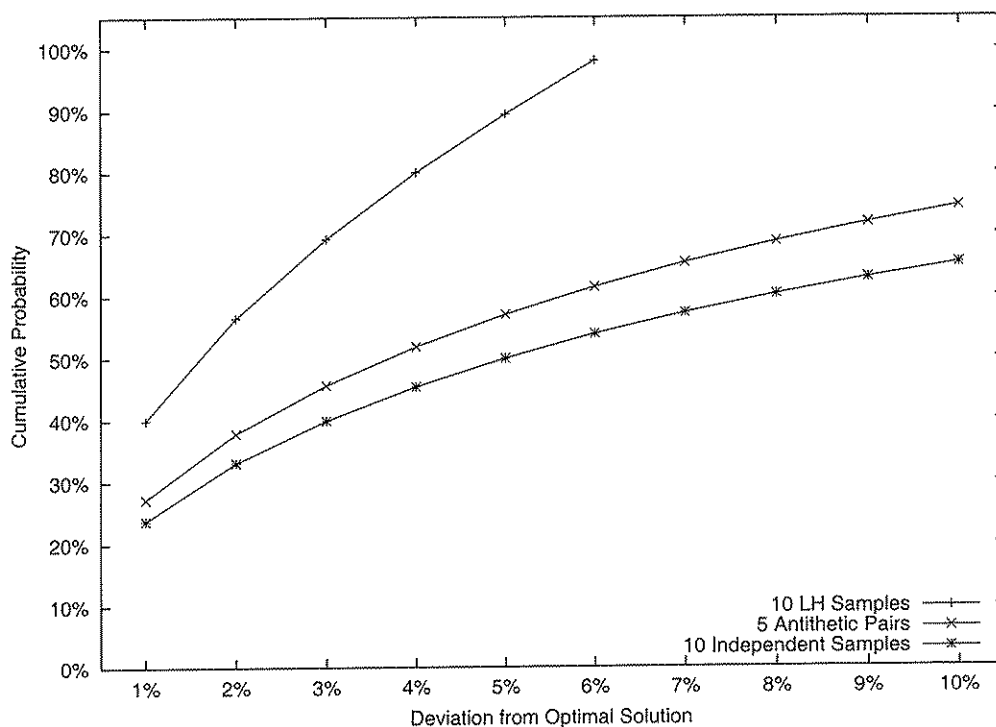


Figure 3: Probability that the actual performance of the sample path solution lies within $x\%$ of the optimal solution for newsvendor problem

## 2.2 The Use of Latin Hypercube Sampling to Reduce Bias

We can also attack the bias in expressions (2.7) and (2.8) using Latin Hypercube sampling (LH), another technique usually prescribed for variance reduction (McKay et al., 1979). In this one-dimensional problem, we divide the interval $[0, 1]$ into $N$ equal segments; the $i^{th}$

demand value $D_i$ is drawn uniformly from the $i^{th}$ segment. The solution to the sample path problem under Latin Hypercube sampling, $\hat{x}_{LH}$, is the demand value drawn from the $\lceil \alpha N \rceil^{\text{th}}$ segment, which is uniformly distributed on $[(\lceil \alpha N \rceil - 1)/N, \lceil \alpha N \rceil/N]$. Under AV we have $\mathbb{E}[\hat{x}] < \mathbb{E}[\hat{x}_{AV}] < \alpha$ when $\alpha N$ is integer. Similarly we now have $\mathbb{E}[\hat{x}_{LH}] < \alpha$; however the relationships between $\mathbb{E}[\hat{x}_{LH}]$ and the values $\mathbb{E}[\hat{x}_{AV}]$ and $\mathbb{E}[\hat{x}]$ depend on the choice of $\alpha$ and $N$.

The derivation of the expected perceived cost under LH is straightforward:

$$\mathbb{E}_{D_1,\ldots D_N} \left[ N^{-1} \sum_{i=1}^{N} (1-\alpha)(\hat{x}_{LH} - D_i)^+ + \alpha(D_i - \hat{x}_{LH})^+ \right]$$

$$= N^{-1} \mathbb{E}_{D_1,\ldots D_N} \left[ \sum_{i=1}^{\lceil \alpha N \rceil - 1} (1-\alpha)(D_{\lceil \alpha N \rceil} - D_i) + \sum_{i=\lceil \alpha N \rceil + 1}^{N} \alpha(D_i - D_{\lceil \alpha N \rceil}) \right]$$

$$= N^{-1} \left[ \sum_{i=1}^{\lceil \alpha N \rceil - 1} N^{-1}(1-\alpha)(\lceil \alpha N \rceil - i) + \sum_{i=\lceil \alpha N \rceil + 1}^{N} N^{-1}\alpha(i - \lceil \alpha N \rceil) \right]$$

$$= \frac{\lceil \alpha N \rceil (\lceil \alpha N \rceil - 2\alpha N - 1) + \alpha N(N+1)}{2N^2}.$$

When $\alpha N$ is integer this expression reduces to $\alpha(1-\alpha)/2$; comparing this to (2.2) we see the perceived cost estimate is unbiased. The expected actual cost under LH is computed from (2.5):

$$\mathbb{E}_{D_1,\ldots D_N} \left[ \hat{x}_{LH}^2 \right]/2 - \alpha \mathbb{E}_{D_1,\ldots D_N} \left[ \hat{x}_{LH} \right] + \alpha/2$$

$$= \frac{1}{2} \left[ \frac{1}{12N^2} + \left( \frac{\lceil \alpha N \rceil}{N} - \frac{1}{2N} \right)^2 \right] - \alpha \left( \frac{\lceil \alpha N \rceil}{N} - \frac{1}{2N} \right) + \frac{\alpha}{2}$$

$$= \frac{1}{8N^2} \left[ \frac{1}{3} + (2\lceil \alpha N \rceil - 1)^2 \right] - \frac{\alpha}{2N}(2\lceil \alpha N \rceil - 1) + \frac{\alpha}{2}.$$

When $\alpha N$ is integer this expression reduces to $\alpha(1-\alpha)/2 + 1/(6N^2)$. Figure 1 plots the expected actual cost under Latin Hypercube sampling as a percentage of the optimal cost. We see that LH reduces the gap between the expected actual cost of the sample path solution and the optimal cost more effectively than AV. Furthermore, as shown in Figure 3, the quality of the LH solution dominates both AV and IS solutions.

As with AV, the use of LH also can also reduce the variance of the optimal objective function estimator (i.e., the variance of the perceived cost) for the stochastic LP. (Refer to the Appendix for the derivation.) This is true for any value of $\alpha$ when $N$ is greater than 3. Again, the combination of the two effects is reflected in the MSE, although as we noted

above, the bias for this particular newsvendor case is equal to zero, so the MSE and the variance are equal. Figure 4 shows the dramatic reduction obtained in MSE by using Latin Hypercube sampling.
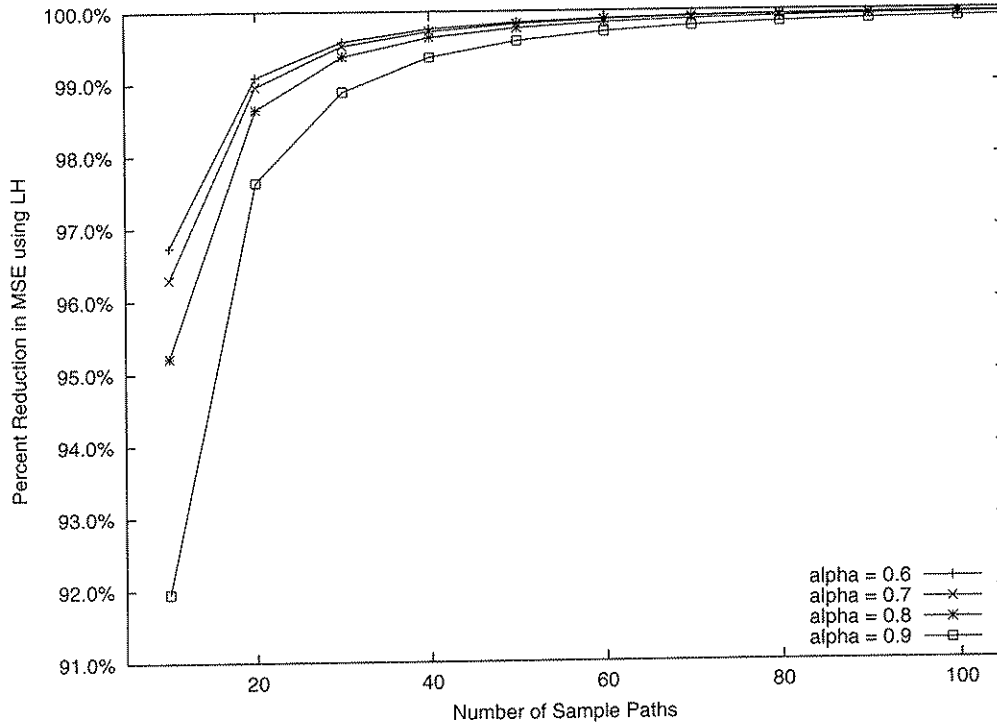


Figure 4: Change in MSE obtained by using Latin Hypercube sampling for the newsvendor problem, as a function of the number of sample paths

## 2.3   A Sufficient Condition for Reduced Bias under AV

In section 2.1 we showed analytically that the use of AV reduces the bias of the perceived performance of the sample path solution to the newsvendor problem. In Section 3 we also provide several computational examples. We now provide a sufficient condition for a reduction in bias when AV is applied to one of the variables in a stochastic LP. For the stochastic LP described in Section 1, generate an $N = 2$ realization of the sample path problem using uniform random numbers $u_1$ and $u_2$, and let $f(u_1, u_2)$ be the optimal objective function value for this realization. Since this is a minimization problem, $\mathbb{E}\left[f(u_1, u_2)\right]$ is bounded above by

12

the optimal objective function value to the original problem. This is true when $u_1$ and $u_2$ are generated independently or under AV; the results of Mak et al. (1999) hold in either case.

We would like a condition under which the expected perceived performance of the sample path solution under AV is greater than the performance under independent sampling (IS):

$$\mathbb{E}_{AV}\left[f(u_1, u_2)\right] > \mathbb{E}_{IS}\left[f(u_1, u_2)\right]$$
$$\int_0^1 f(u_1, 1 - u_1)du_1 > \int_0^1 \int_0^1 f(u_1, u_2)du_2 du_1. \tag{2.11}$$

To derive this condition, we first define the transformation $(v_1, v_2) = g(u_1, u_2)$:

$$v_1 = \frac{\sqrt{2}}{2}(u_1 + u_2), \quad v_2 = \frac{\sqrt{2}}{2}(u_1 - u_2).$$

This transformation is shown in Figure 5. Let $h(v)$ be the density of $v_1$. Since $(u_1, u_2)$ were
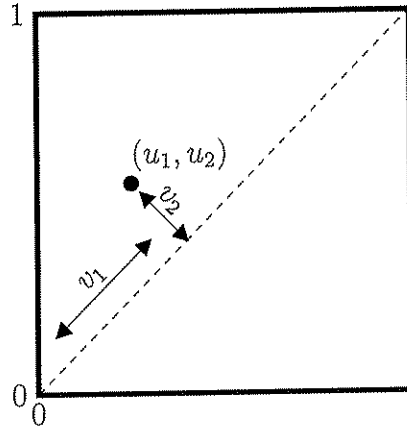


Figure 5: Change of variables from $(u_1, u_2)$ to $(v_1, v_2)$

independent uniforms, $h(v)$ is triangular with support $[0, \sqrt{2}]$. For fixed $v_1, v_2$ is uniformly distributed on $[-v_1 + 2(v_1 - \sqrt{2}/2)^+, v_1 - 2(v_1 - \sqrt{2}/2)^+]$. Furthermore:

$$\mathbb{E}_{IS}\left[f(u_1, u_2)\right] = \int_0^{\sqrt{2}} \int_{-v_1 + 2(v_1 - \sqrt{2}/2)^+}^{v_1 - 2(v_1 - \sqrt{2}/2)^+} \frac{f(g^{-1}(v_1, v_2))}{2(v_1 - 2(v_1 - \sqrt{2}/2)^+)} dv_2 h(v_1) dv_1, \text{ and}$$

$$\mathbb{E}_{AV}\left[f(u_1, u_2)\right] \int_{-\sqrt{2}/2}^{\sqrt{2}/2} \frac{f(g^{-1}(\sqrt{2}/2, v_2))}{\sqrt{2}} dv_2.$$

Therefore a sufficient condition for (2.11) is:

$$\int_{-v_1 + 2(v_1 - \sqrt{2}/2)^+}^{v_1 - 2(v_1 - \sqrt{2}/2)^+} \frac{f(g^{-1}(v_1, v_2))}{2(v_1 - 2(v_1 - \sqrt{2}/2)^+)} dv_2 \text{ is strictly concave in } v_1. \tag{C}$$

Condition (C) is not necessary for (2.11); it is possible to construct examples in which the condition fails but AV still reduces bias. However (C) holds for the newsvendor problem. In this case $f(u_1, u_2)$ is derived from (2.3):

$$\min_x (1/2) \left[ (1 - \alpha)(x - u_1)^+ + \alpha(u_1 - x)^+ + (1 - \alpha)(x - u_2)^+ + \alpha(u_2 - x)^+ \right].$$

The optimal solution is the $\lceil 2\alpha \rceil^{\text{th}}$ order statistic, which is $\min(u_1, u_2)$ if $\alpha < 0.5$. Therefore:

$$\begin{aligned}
f(u_1, u_2) =& 1/2[(1 - \alpha)(\min(u_1, u_2) - u_1)^+ + \alpha(u_1 - \min(u_1, u_2))^+ \\
& + (1 - \alpha)(\min(u_1, u_2) - u_2)^+ + \alpha(u_2 - \min(u_1, u_2))^+] \\
=& (\alpha/2)((u_1 - u_2)^+ + (u_2 - u_1)^+).
\end{aligned}$$

Translating to $(v_1, v_2)$ coordinates we have $f(g^{-1}(v_1, v_2)) = (\alpha/2)\sqrt{2}|v_2|$, so:

$$\begin{aligned}
\int_{-v_1 + 2(v_1 - \sqrt{2}/2)^+}^{v_1 - 2(v_1 - \sqrt{2}/2)^+} \frac{f(g^{-1}(v_1, v_2))}{2(v_1 - 2(v_1 - \sqrt{2}/2)^+)} dv_2 &= \int_{-v_1 + 2(v_1 - \sqrt{2}/2)^+}^{v_1 - 2(v_1 - \sqrt{2}/2)^+} \frac{\alpha\sqrt{2}|v_2|}{4\left(v_1 - 2(v_1\sqrt{2}/2)^+\right)} dv_2 \\
&= \frac{\alpha\sqrt{2}\left(v_1 - 2\left(v_1 - \sqrt{2}/2\right)^+\right)}{4}.
\end{aligned}$$

This is strictly concave in $v_1$. A similar result holds if $\alpha \geq 0.5$.

Unfortunately the same approach does not yield a condition under which the expected perceived performance of the sample path solution with LH is greater than the performance with IS.

# 3.   Computational Examples

In Sections 2.1 and 2.2 we analytically characterized the effects of sampling methods on the bias and variance of the solution to a simple sample path problem. Here, we present empirical results of applying these sampling methods to a set of more complicated test problems. The section contains a brief description of our test problems, a description of our approach for obtaining statistical estimates for perceived and actual cost, a description of our computational platform, and the results of the experiments.

## 3.1   Test Problems

The test problems are two-stage stochastic linear programs with recourse that were obtained from the literature. Table 1 contains details about each of the problems. The

14

| Name | Application | Source | Scenarios |
|---|---|---|---|
| 20term | Vehicle Positioning | Mak et al. (1999) | $1.1 \times 10^{12}$ |
| fleet | Fleet Planning | Powell and Topaloglu (2005) | $8.5 \times 10^{113}$ |
| gbd | Aircraft Allocation | Dantzig (1963) | $6.5 \times 10^5$ |
| LandS | Electrical Investment Planning | Louveaux and Smeers (1988) | $10^6$ |
| snip | Stochastic Network Interdiction | Janjarassuk and Linderoth (2005) | $3.7 \times 10^{19}$ |
| ssn | Telecommunication Network Design | Sen et al. (1994) | $10^{70}$ |
| storm | Flight Scheduling | Mulvey and Ruszczyński (1995) | $6 \times 10^{81}$ |

Table 1: Description of test instances

problem *fleet* is a fleet management problem available from the page `http://www.orie.cornell.edu/~huseyin/research/research.html#Fleet_20_3`. The problem *snip* is a (linear relaxation) of a stochastic network interdiction problem available at the page `http://coral.ie.lehigh.edu/sp-instances/`. The remaining problems are described in Linderoth et al. (2002) and available from the companion web site `http://www.cs.wisc.edu/~swright/stochastic/sampling/`.

## 3.2 Methodology

**Perceived Cost Estimates** As indicated by the inequalities in (1.2), and previously shown by Norkin et al. (1998) and Mak et al. (1999), the expected perceived cost:

$$\mathbb{E}_{(\omega_1,...,\omega_N)}\left[z^*_{\mathrm{MP}_{N(\omega_1,...,\omega_N)}}\right]$$

is a biased estimate of $z^*_{MP}$, the value of the optimal solution. First, we generate $M$ independent (and identically distributed) samples of size $N$: $(\omega_1^1,...,\omega_N^1),...,(\omega_1^M,...,\omega_N^M)$. We define $\ell_j, j = 1, 2, ... M$, to be the solution value of the $j$th sample path problem:

$$\ell_j \overset{\mathrm{def}}{=} z^*_{\mathrm{MP}_{N(\omega_1^j,...,\omega_N^j)}}$$

and compute the value:

$$\mathcal{L}_{N,M} \overset{\mathrm{def}}{=} \frac{1}{M}\sum_{j=1}^{M}\ell_j.$$

The statistic $\mathcal{L}_{N,M}$ provides an unbiased estimate of $\mathbb{E}_{(\omega_1,...,\omega_N)}\left[z^*_{\mathrm{MP}_{N(\omega_1,...,\omega_N)}}\right]$. Since the $M$ samples are i.i.d, we can construct an approximate $(1-\alpha)$ confidence interval for $\mathbb{E}_{(\omega_1,...,\omega_N)}\left[z^*_{\mathrm{MP}_{N(\omega_1,...,\omega_N)}}\right]$

$$\left[\mathcal{L}_{N,M} - \frac{z_{\alpha/2}s_{\mathcal{L}}(M)}{\sqrt{M}}, \mathcal{L}_{N,M} + \frac{z_{\alpha/2}s_{\mathcal{L}}(M)}{\sqrt{M}}\right], \tag{3.1}$$

where

$$s_{\mathcal{L}}(M) \stackrel{\text{def}}{=} \sqrt{\frac{1}{M-1} \sum_{j=1}^{M} (\ell^j - \mathcal{L}_{N,M})^2}. \qquad (3.2)$$

For small values of $M$, one can use $t_{\alpha/2,M-1}$ critical values instead of $z_{\alpha/2}$, which will produce slightly bigger confidence intervals.

**Actual Cost Estimates**   Since a solution to the sample path problem $\mathrm{MP}_N$ may be sub-optimal with respect to the true objective function (1.1), we estimate the expected actual cost of $x_N^*(\omega_1, \ldots, \omega_N)$, an optimal solution to $\mathrm{MP}_N$. We estimate the expected actual cost of a sample path problem of size $N$ in the following manner. First, we generate $M$ samples of size $N$: $(\omega_1^1, \ldots, \omega_N^1), \ldots (\omega_1^M, \ldots, \omega_N^M)$ and solve the sample-path problem $MP_N$ for each sample yielding:

$$x_j^* \in \arg \min_{Ax=b, x \geq 0} N^{-1} \sum_{i=1}^{N} Q_i(x, \omega_i^j) + g(x), \ j = 1, 2, \ldots M.$$

Note that this is the same calculation necessary to compute a lower bound on the optimal objective value, and the computational effort required is to solve $M$ sample path problems, each containing $N$ scenarios. Next, for each candidate solution $x_j^*$, we take a new, Latin Hypercube sample of size $N'$, $(\omega_1^j, \ldots, \omega_{N'}^j)$ and compute the quantity:

$$a_j = \sum_{i=1}^{N'} Q(x_j^*, \omega_i^j) + g(x_j^*). \qquad (3.3)$$

(Latin Hypercube sampling appears to be superior to the other two methods for variance reduction; thus, we use this technique to estimate expected actual cost no matter what sampling method was used to obtain $x_j^*$.) Since $x_j^*$ is fixed, this computation required the solution of $N'$ independent linear programs. The quantity:

$$\mathcal{A}_{N,M} \stackrel{\text{def}}{=} \frac{1}{M} \sum_{j=1}^{M} a_j$$

is an unbiased estimate of the expected actual cost:

$$\mathbb{E}_{(\omega_1, \ldots, \omega_N)} \left[ \mathbb{E}_\omega \left[ Q(x_{\mathrm{MP}_{N(\omega_1, \ldots, \omega_N)}}^*, \omega) \right] + g(x_{\mathrm{MP}_{N(\omega_1, \ldots, \omega_N)}}^*) \right].$$

Since the random quantities $a^j$ are i.i.d., we can construct an approximate $(1 - \alpha)$ confidence interval for:

$$\mathbb{E}_{(\omega_1, \ldots, \omega_N)} \left[ \mathbb{E}_\omega \left[ Q(x_{\mathrm{MP}_{N(\omega_1, \ldots, \omega_N)}}^*, \omega) \right] + g(x_{\mathrm{MP}_{N(\omega_1, \ldots, \omega_N)}}^*) \right]$$

16

as:

$$\left[ \mathcal{A}_{N,M} - \frac{z_{\alpha/2} s_{\mathcal{A}}}{\sqrt{M}}, \mathcal{A}_{N,M} + \frac{z_{\alpha/2} s_{\mathcal{A}}}{\sqrt{M}} \right],$$ (3.4)

where:

$$s_{\mathcal{A}}(M) \overset{\text{def}}{=} \sqrt{\frac{1}{M-1} \sum_{j=1}^{M} \left( a^j - \mathcal{A}_{N,M} \right)^2}.$$ (3.5)

## 3.3  Computational Platform

The computational experiments presented here were performed on a non-dedicated, distributed computing platform known as a *computational grid* (Foster and Kesselman, 1999). The computational platform was created with the aid of the Condor software toolkit (Livny et al., 1997), which can be configured to allow for the idle cycles of machines to be donated to a "Condor pool". Table 2 shows the characteristics of the computing environment used to solve our test instances.

| # of CPUs | Operating System | Processor Type | Clock Speed |
|---|---|---|---|
| 110 | Linux | Opteron | 1.8GHz |
| 48 | Linux | Xeon | 1.4GHz |
| 96 | Linux | Pentium III | 1.1GHz |

Table 2: CPU Resources used for experiments

In order to create the sampled problems, we use the SUTIL software toolkit (Czyzyk et al., 2005). Specifically for this work, SUTIL was equipped with the ability to sample two-stage stochastic programs using an antithetic variates sampling technique. An important feature of SUTIL, necessary when running in a distributed and heterogeneous computing environment, is its ability to obtain the *same* value for a random vector $\omega^j$ on different processors and at different points of the solution algorithm (say different iterations of the LShaped method). This is a nontrivial implementation issue and is accomplished in SUTIL by employing an architecture- and operating-system-independent random number stream, storing and passing appropriate random seed information to the participating processors, and performing some recalculation of random vectors in the case that the vectors in a sample are correlated.

In order to solve the sampled problems, we use the code **atr** of Linderoth and Wright (2003). The algorithm is a variation of the well-known LShaped algorithm (Van Slyke and

Wets, 1969) that has been enhanced with mechanisms for reducing the synchronization requirements of the algorithm (useful for the distributed computing environment), and also with a $\|\cdot\|_\infty$-norm trust region to help stabilization of the master problem. The initial iterate of the algorithm was taken to be the solution of a sampled instance of intermediate size.

## 3.4 Computational Results

Our computational experiments were designed to examine the impact of different sampling methods on the bias and variance of the perceived cost of 2-stage stochastic linear programs solved via sample path optimization. Recall that bias and variance reduction combine to improve the mean squared error of the solution to the sample path problem. In the results presented here, the optimal solution to the full problem, $z^*_{\mathrm{MP}}$, is unknown, so we cannot calculate the bias. Since we know that for minimization problems, the expected value of the sample path solution, $\mathbb{E}\left[z^*_{\mathrm{MP}_N}\right]$, is less than the true optimal solution, $z^*_{\mathrm{MP}}$, we can test whether or not one sampling method reduces bias as compared to another by testing whether or not the expected value of the sample path solution, $\mathbb{E}\left[z^*_{\mathrm{MP}_N}\right]$, is significantly larger and therefore closer to the true optimal solution, $z^*_{\mathrm{MP}}$.

Using samples drawn in an independent fashion, samples drawn using antithetic variates, and samples drawn using Latin Hypercube sampling, the following experiment was performed. For each of the instances described in Table 1, confidence intervals for both expected perceived cost and expected actual cost (as defined in 3.1 and 3.4) were computed for $M = 50$ for $N \in \{50, 100, 500, 1000\}$, and $M = 10$ for $N \in \{5000, 10000, 50000\}$. The value $N'$ used in the calculation of $a_j$ (3.3) was $N' = 20,000$ in each case. The complete experiment required the solution of 1,134,682,000 linear programs, so the ability to run in the powerful distributed setting of the computational grid was of paramount importance to this work.

Tables 5—11 in the Appendix summarize the confidence intervals for expected perceived cost and expected actual cost, and Tables 3 and 4 show the results of $t$-tests for bias reduction and $F$-tests for variance reduction for the expected perceived cost. (Since each trial was independently generated, and variances are significantly different in some cases, we use one-sided, unpaired student $t$-tests, assuming unequal variance.) We use the symbol $\succ$ to indicate when a test assumes one method is preferred to another. Note that for this set of problems, statistically significant bias reduction with AV occurs occasionally. Bias reduction is observed

with LH slightly more frequently, particularly for problem *ssn* (more on this in a moment). Interestingly, any time AV results in bias reduction, LH does as well. Both AV and LH sampling methods are effective in reducing variance, with LH reducing variance as compared to IS in almost all cases.

It is worth noting that statistically significant bias reduction may not be detected either because it does not exist or because there is too much variability in the estimate of $\mathbb{E}[z^*_{\mathrm{MP}_N}]$. Figures 6 and 7 show estimates of expected perceived and expected actual cost with confidence intervals for problem *fleet*. All confidence intervals shown in the figures and tables use $z_{0.975} \approx 1.96$ when $M = 50$ ($N \in \{50, 100, 500, 1000\}$) and $t_{0.025, M-1} \approx 2.685$ when $M = 10$ ($N \in \{5000, 10000, 50000\}$). A horizontal reference line is shown on the perceived and actual cost figures for each problem. In Figure 7, one can easily see that both AV and LH reduce variance for smaller values of $N$. While bias may be reduced also, it is difficult to tell given the large variance. Figures 8 and 9 tell a different story for problem *ssn*. It is quite clear from Figure 9 that LH reduces bias for small values of $N$, while variance reductions are less obvious. For *ssn* with $N$=50, the reduction in bias is approximately 4.5 (see values in Table 10), while the reduction in variance is not statistically significant. For *fleet* with $N$=50, the variance reduction is approximately 62,690, while the bias reduction is not significant. Similar perceived and actual cost figures for the other five problems are in the Appendix.

As was shown analytically in Section 2.1, AV can *increase* variance for the newsvendor problem with certain values of $\alpha$. In our computational experiments, we found no cases where either AV or LH significantly increased variance. We did, however, encounter a few cases where bias was increased by AV or LH. Table 4 summarizes the four (out of 441) cases where there was a statistically significant increase in bias. Two of the cases are depicted in Figure 7 with $N = 10000$ and Figure 9 with $N = 1000$. Five of the six cases have a small number of observations, $M = 10$. For these small $M$ cases, we perform Wilcoxon-Mann-Whitney rank sum tests, noting that two of the five cases are not significant at the 5% level according to the rank-sum test.

| Instance | $N$ | Bias Reduction (t-test) | | | Variance Reduction (F-test) | | |
|---|---|---|---|---|---|---|---|
| | | AV $\succ$ IS | LH $\succ$ IS | LH $\succ$ AV | AV $\succ$ IS | LH $\succ$ IS | LH $\succ$ AV |
| 20 | 50 | 0.0002 | 0.0001 | | 0.0000 | 0.0000 | |
| | 100 | | | | 0.0000 | 0.0000 | 0.0386 |
| | 500 | | | | 0.0000 | 0.0000 | |
| | 1000 | | | | 0.0000 | 0.0000 | |
| | 5000 | | | | 0.0123 | | |
| | 10000 | | | | 0.0009 | 0.0003 | |
| | 50000 | | | 0.0069 | 0.0011 | 0.0011 | |
| fleet | 50 | | | | 0.0078 | 0.0000 | 0.0002 |
| | 100 | | | | 0.0000 | 0.0000 | |
| | 500 | | | | 0.0003 | 0.0000 | 0.0263 |
| | 1000 | | | | 0.0135 | 0.0000 | 0.0001 |
| | 5000 | | | | 0.0177 | 0.0028 | |
| | 10000 | | | | | 0.0000 | 0.0052 |
| | 50000 | | | | | 0.0025 | 0.0077 |
| gbd | 50 | | | 0.0112 | 0.0037 | 0.0000 | 0.0000 |
| | 100 | | 0.0076 | | 0.0213 | 0.0000 | 0.0000 |
| | 500 | | | | | 0.0000 | 0.0000 |
| | 1000 | | | | | 0.0000 | 0.0000 |
| | 5000 | | | | | 0.0000 | 0.0000 |
| | 10000 | | | | | 0.0000 | 0.0000 |
| | 50000 | | | | | 0.0000 | 0.0000 |
| LandS | 50 | | | | 0.0000 | 0.0000 | |
| | 100 | | | | 0.0000 | 0.0000 | 0.0000 |
| | 500 | | | | 0.0000 | 0.0000 | 0.0000 |
| | 1000 | | | | 0.0000 | 0.0000 | 0.0095 |
| | 5000 | 0.0012 | 0.0007 | | 0.0000 | 0.0000 | |
| | 10000 | | | | 0.0000 | 0.0000 | 0.0300 |
| | 50000 | | | | 0.0000 | 0.0000 | 0.0207 |
| snip | 50 | | 0.0052 | 0.0432 | | 0.0000 | 0.0000 |
| | 100 | 0.0203 | 0.0003 | | | 0.0000 | 0.0000 |
| | 500 | | | | 0.0006 | 0.0000 | 0.0000 |
| | 1000 | | | | | 0.0000 | 0.0000 |
| | 5000 | | | | | 0.0032 | 0.0008 |
| | 10000 | | | | | 0.0001 | 0.0013 |
| | 50000 | | | | | 0.0000 | 0.0024 |
| ssn | 50 | | 0.0000 | 0.0000 | | | |
| | 100 | | 0.0000 | 0.0000 | | 0.0149 | 0.0001 |
| | 500 | | 0.0010 | 0.0000 | | 0.0010 | 0.0004 |
| | 1000 | | 0.0010 | 0.0000 | | 0.0302 | 0.0213 |
| | 5000 | | | 0.0052 | | 0.0065 | |
| | 10000 | | | | | | |
| | 50000 | | 0.0070 | | | | 0.0204 |
| storm | 50 | | | | 0.0000 | 0.0000 | 0.0000 |
| | 100 | | | | 0.0000 | 0.0000 | 0.0000 |
| | 500 | | | | 0.0000 | 0.0000 | 0.0000 |
| | 1000 | | | | 0.0000 | 0.0000 | 0.0000 |
| | 5000 | | | | 0.0000 | 0.0000 | |
| | 10000 | | | | 0.0000 | 0.0000 | 0.0076 |
| | 50000 | | | | 0.0000 | 0.0000 | 0.0321 |

Table 3: $p$-values (from $t$ and $F$ test) for cases where LH or AV sampling methods result in a statistically significant (at 5% or better) reduction of bias or variance

| Instance | $N$ | AV $\prec$ IS | | LH $\prec$ IS | |
|---|---|---|---|---|---|
| | | $t$-test | Rank-sum test | $t$-test | Rank-sum test |
| 20 | 50000 | 0.0179 | 0.1212 | | |
| fleet | 10000 | 0.0059 | 0.01726 | 0.0121 | 0.0258 |
| gbd | 10000 | 0.0369 | 0.06402 | 0.0029 | 0.02575 |
| ssn | 1000 | 0.0176 | | | |

Table 4: $p$-values (from $t$ and Wilcoxon-Mann-Whitney Rank Sum tests) for "opposite" cases where AV and LH sampling methods result in a statistically significant *increase* in bias

# 4. Conclusion

Sample path optimization is a convenient method for solving stochastic programs; however a gap is introduced between the optimal solution and both the expected actual and expected perceived cost of the sample path solution. We have investigated two variations of sample path optimization where samples are drawn in antithetic pairs or using Latin Hypercube sampling. For a version of the simple newsvendor problem, we show that both the antithetic samples approach and the Latin Hypercube approach, techniques commonly used for variance reduction, reduce the solution *bias* as compared to sample path optimization with independent samples. For the newsvendor problem, the Latin Hypercube approach reduces variance of the sample path solution, while antithetic variates may increase or decrease the variance, depending on the cost parameters. In addition, we provide a sufficient condition for when sampling by antithetic variates would reduce the bias.

Using a computational grid, we perform extensive computational experiments investigating these same sampling methods on large-scale, two-stage, stochastic programs from the literature. We find that both sampling techniques are effective at reducing variance. For many of our problems, bias reduction is difficult to detect, however, for one of our instances, *ssn*, Latin Hypercube sampling dramatically reduces the bias.
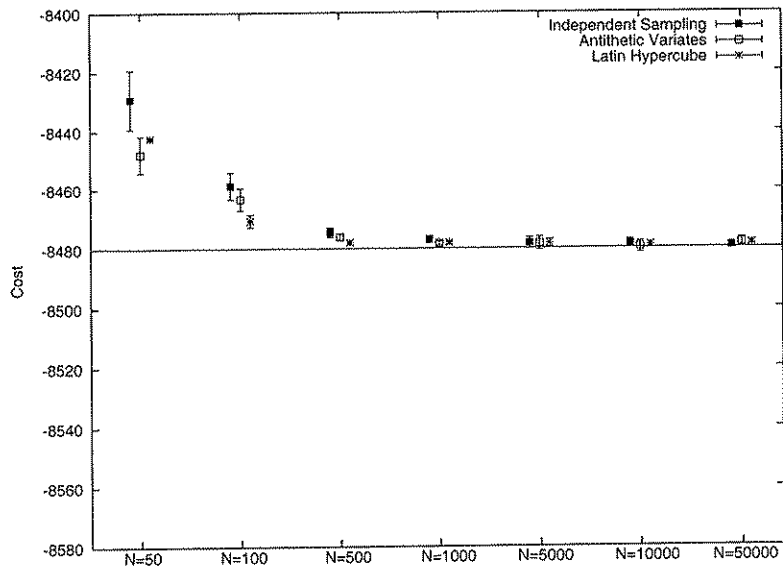
# Acknowledgments

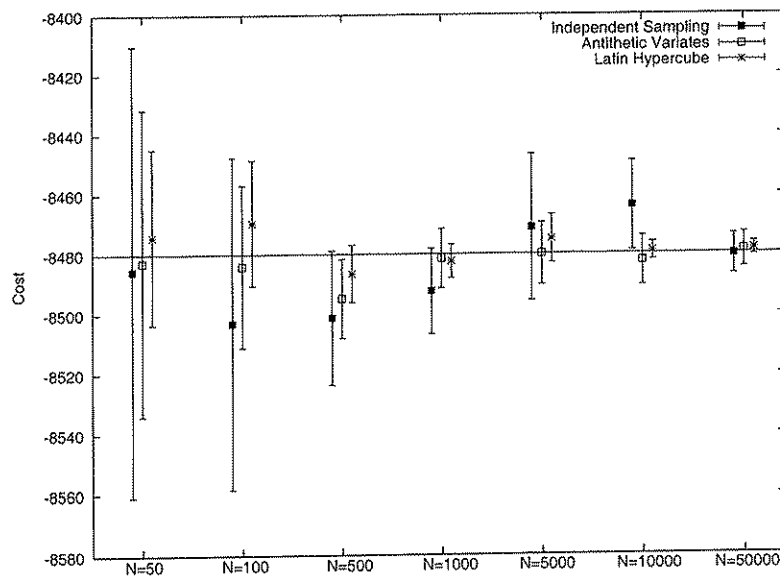Figure 6: Expected Actual Cost Estimates for **fleet**



Figure 7: Expected Perceived Cost Estimates for **fleet**
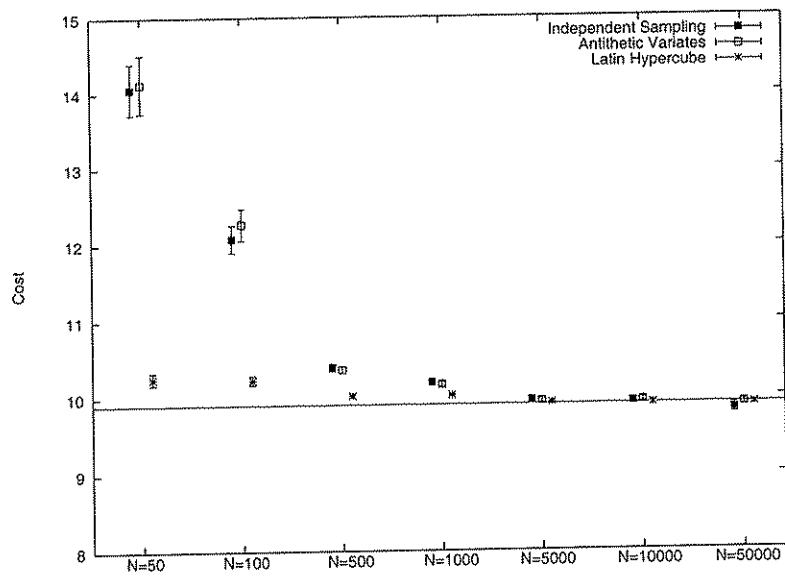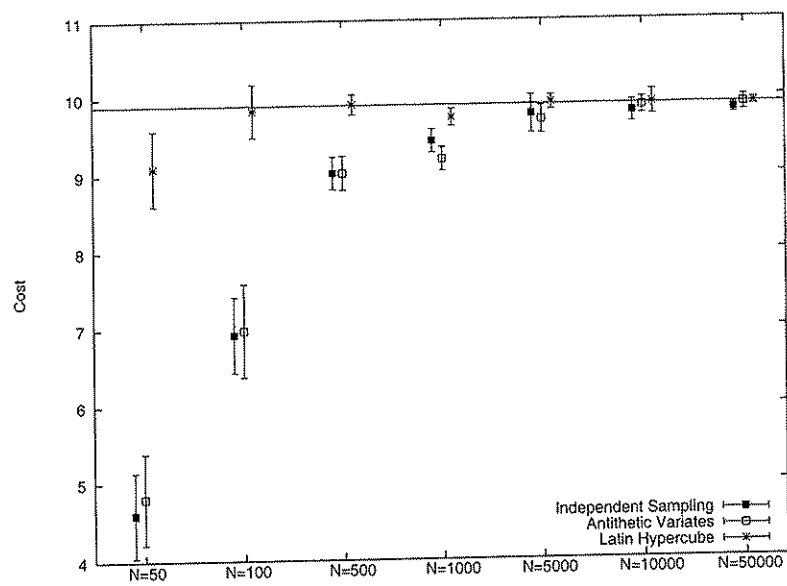
Figure 8: Expected Actual Cost Estimates Estimates for **ssn**



Figure 9: Expected Perceived Cost Estimates for **ssn**

23

# References

Beale, E. M. 1955. On minimizing a convex function subject to linear inequalities. *Journal of the Royal Statistical Society* **17B** 173–184.

Czyzyk, J., J. Linderoth, J. Shen. 2005. SUTIL: A utility library for handling stochastic programs. User's Manual.

Dantzig, G. B. 1955. Linear programming under uncertainty. *Management Science* **1** 197–206.

Dantzig, G. B. 1963. *Linear Programming and Extensions*. Princeton University Press, Princeton, NJ.

Dupačová, J., R. J.-B. Wets. 1988. Asymptotic behavior of statistical estimators and of optimal solutions of stochastic optimization problems. *The Annals of Statistics* **16** 1517–1549.

Foster, I., C. Kesselman. 1999. Computational grids. I. Foster, C. Kesselman, eds., *The Grid: Blueprint for a New Computing Infrastructure*. Morgan Kaufmann. Chapter 2.

Higle, J. L. 1998. Variance reduction and objective function evaluation in stochastic linear programs. *INFORMS Journal on Computing* **10** 236–247.

Janjarassuk, U., J. Linderoth. 2005. Reformulation and sampling to solve a stochastic network interdiction problem. In preparation.

Kleywegt, A. J., A. Shapiro. 2001. Stochastic optimization. *Handbook of Industrial Engineering*, 3rd ed. John Wiley & Sons, New York, 2625–2649.

Law, A., D. Kelton. 2000. *Simulation Modeling and Analysis*. 3rd ed. McGraw-Hill, Boston.

Linderoth, Jeffrey T., Alex Shapiro, Stephen J. Wright. 2002. The empirical behavior of sampling methods for stochastic programming. Tech. Rep. Optimization Technical Report 02-01, Computer Sciences Department, University of Wisconsin-Madison.

Linderoth, Jeffrey T., Stephen J. Wright. 2003. Implementing a decomposition algorithm for stochastic programming on a computational grid. *Computational Optimization and Applications* **24** 207–250.

Livny, M., J. Basney, R. Raman, T. Tannenbaum. 1997. Mechanisms for high throughput computing. *SPEEDUP* **11**.

Louveaux, F., Y. Smeers. 1988. Optimal investments for electricity generation: A stochastic model and a test problem. Y. Ermoliev, R. J.-B Wets, eds., *Numerical techniques for stochastic optimization problems*. Springer-Verlag, Berlin, 445–452.

Mak, W. K., D. P. Morton, R. K. Wood. 1999. Monte carlo bounding techniques for determining solution quality in stochastic programs. *Operations Research Letters* **24** 47–56.

McKay, M. D., R. J. Beckman, W. J. Conover. 1979. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* **21** 239–245.

Mulvey, J. M., A. Ruszczyński. 1995. A new scenario decomposition method for large scale stochastic optimization. *Operations Research* **43** 477–490.

Norkin, V., G. Pflug, A. Ruszczyński. 1998. A branch and bound method for stochastic global optimization. *Mathematical Programming* **83** 425–450.

Powell, W., H. Topaloglu. 2005. Fleet management. S. Wallance, W. Ziemba, eds., *Applications of Stochastic Programming*. MPS-SIAM Series on Optimization, Philadelphia.

Sen, S., R. D. Doverspike, S. Cosares. 1994. Network planning with random demand. *Telecommunications Systems* **3** 11–30.

Van Slyke, R., R.J-B. Wets. 1969. L-shaped linear programs with applications to control and stochastic programming. *SIAM Journal on Applied Mathematics* **17** 638–663.

# Appendix

In this appendix we compute $Var[z^*_{\text{MP}_N(D_1,...,D_N)}]$, the variance with respect to demands $D_1, \ldots, D_n$ of the perceived cost of the newsvendor sample path solution. The analysis is based on the following expression for variance involving random variables $X$ and $Y$:

$$Var(X) = Var_Y[E(X|Y)] + E_Y[Var(X|Y)].$$

(See, for example, (Law and Kelton, 2000).) We have:

$$
\begin{aligned}
Var[z^*_{\text{MP}_N(D_1,...,D_N)}] =& Var\left[\frac{1}{N}\sum_{i=1}^{N}(1-\alpha)(\hat{x}-D_i)^+ + \alpha(D_i-\hat{x})^+\right] \\
=& Var_{\hat{x}}\left[\mathbb{E}\left[\frac{1}{N}\sum_{i=1}^{N}(1-\alpha)(\hat{x}-D_i)^+ + \alpha(D_i-\hat{x})^+\,\middle|\,\hat{x}\right]\right] \\
& + \mathbb{E}_{\hat{x}}\left[Var\left[\frac{1}{N}\sum_{i=1}^{N}(1-\alpha)(\hat{x}-D_i)^+ + \alpha(D_i-\hat{x})^+\,\middle|\,\hat{x}\right]\right].
\end{aligned}
\tag{4.1}
$$

Recalling that $\hat{x}$ is the $\lceil\alpha N\rceil^{th}$ order statistic of the demand values $D_1, \ldots, D_n$, we analyze the two terms on the right side of (4.1) when demands are sampled under IS, AV, and LH.

## Independent Sampling (IS)

Under independent sampling, $\hat{x}$ has a Beta distribution with parameters $\lceil\alpha N\rceil$ and $(N - \lceil\alpha N\rceil + 1)$, so:

$$
\mathbb{E}[\hat{x}] = \frac{\lceil\alpha N\rceil}{N+1}
\tag{4.2}
$$

$$
\mathbb{E}[\hat{x}^2] = \frac{(\lceil\alpha N\rceil)(\lceil\alpha N\rceil+1)}{(N+1)(N+2)}
\tag{4.3}
$$

$$
Var[\hat{x}] = \frac{(\lceil\alpha N\rceil)(N-\lceil\alpha N\rceil+1)}{(N+1)^2(N+2)}.
\tag{4.4}
$$

We condition on the value of $\hat{x}$. Since $\hat{x}$ is the $\lceil\alpha N\rceil^{th}$ order statistic, $\lceil\alpha N\rceil - 1$ of the demand values are uniformly distributed below $\hat{x}$, and the remaining $N - \lceil\alpha N\rceil$ are uniformly distributed above. The first term on the right side of (4.1) becomes:

$$
\begin{aligned}
& Var_{\hat{x}}\left[\mathbb{E}\left[\frac{1}{N}\sum_{i=1}^{N}(1-\alpha)(\hat{x}-D_i)^+ + \alpha(D_i-\hat{x})^+\,\middle|\,\hat{x}\right]\right] \\
& = Var_{\hat{x}}\left[(1-\alpha)\left(\frac{\lceil\alpha N\rceil-1}{N}\right)\frac{1}{\hat{x}}\int_0^{\hat{x}}(\hat{x}-z)dz + \alpha\left(\frac{N-\lceil\alpha N\rceil}{N}\right)\frac{1}{1-\hat{x}}\int_{\hat{x}}^1(z-\hat{x})dx\right] \\
& = \left(\frac{\lceil\alpha N\rceil-1+\alpha-\alpha N}{2N}\right)^2 Var(\hat{x}).
\end{aligned}
\tag{4.5}
$$

The second term on the right side of (4.1) becomes:

$$\mathbb{E}_{\hat{x}}\left[Var\left[\frac{1}{N}\sum_{i=1}^{N}(1-\alpha)(\hat{x}-D_i)^+ + \alpha(D_i-\hat{x})^+ \Big| \hat{x}\right]\right]$$

$$= \mathbb{E}_{\hat{x}}\left[\frac{1}{N^2}\left[(1-\alpha)^2(\lceil\alpha N\rceil-1)\frac{\hat{x}^2}{12} + \alpha^2(N-\lceil\alpha N\rceil)\frac{(1-\hat{x})^2}{12}\right]\right]. \tag{4.6}$$

Combining (4.1) through (4.6) gives $Var[z^*_{\mathrm{MP}_N(D_1,...,D_N)}]$.

## Antithetic Variates (AV)

Under antithetic variates, $\hat{x}_{AV} = 1/2 + X/2$, where $X$ has a Beta distribution with parameters $\lceil\alpha N - N/2\rceil$ and $(N/2) - \lceil\alpha N - N/2\rceil + 1$, so

$$\mathbb{E}[\hat{x}_{AV}] = \frac{\lceil\alpha N\rceil + 1}{N+2}, \tag{4.7}$$

and

$$Var[\hat{x}_{AV}] = \frac{(\lceil\alpha N\rceil - N/2)(N - \lceil\alpha N\rceil + 1)}{4(N/2+1)^2(N/2+2)}. \tag{4.8}$$

We again condition on the value of $\hat{x}_{AV}$. The first term on the right side of (4.1) becomes (4.9), where the three terms inside the brackets on the right side of (4.9) correspond to the antithetic partner of $\hat{x}_{AV}$, those demand values lying below $\hat{x}_{AV}$, and those demand values lying above $\hat{x}_{AV}$ (but whose antithetic partners lie below):

$$Var_{\hat{x}_{AV}}\left[\mathbb{E}\left[\frac{1}{N}\sum_{i=1}^{N}(1-\alpha)(\hat{x}_{AV}-D_i)^+ + \alpha(D_i-\hat{x}_{AV})^+ \Big| \hat{x}_{AV}\right]\right]$$

$$= Var_{\hat{x}_{AV}}\left[\frac{(1-\alpha)(\hat{x}_{AV}-(1-\hat{x}_{AV}))}{N}\right.$$

$$+ (1-\alpha)\left(\frac{\lceil\alpha N\rceil - N/2 - 1}{N}\right)\frac{1}{\hat{x}_{AV}-1/2}\int_{1/2}^{\hat{x}_{AV}}((\hat{x}_{AV}-z)+(\hat{x}_{AV}-1+z))dz$$

$$+ \left.\left(\frac{N-\lceil\alpha N\rceil}{N}\right)\frac{1}{1-\hat{x}_{AV}}\int_{\hat{x}_{AV}}^{1}\left[\alpha(z-\hat{x}_{AV})+(1-\alpha)(\hat{x}_{AV}-1+z)\right]dz\right] \tag{4.9}$$

$$= Var_{\hat{x}_{AV}}\left[\frac{(-2\alpha N + N + \lceil\alpha N\rceil)\hat{x}_{AV} - \lceil\alpha N\rceil + \alpha N}{2N}\right] \tag{4.10}$$

The second term on the right side of (4.1) becomes:

$$\mathbb{E}_{\hat{x}_{AV}}\left[Var\left[\frac{1}{N}\sum_{i=1}^{N}(1-\alpha)(\hat{x}_{AV}-D_i)^+ + \alpha(D_i-\hat{x}_{AV})^+ \Big| \hat{x}_{AV}\right]\right]$$

$$= E_{\hat{x}_{AV}}\left[(N-\lceil\alpha N\rceil)\frac{(1-\hat{x}_{AV})^2}{12N^2}\right]$$

$$= \frac{N-\lceil\alpha N\rceil}{12N^2}\left(Var[\hat{x}_{AV}] + E[\hat{x}_{AV}]^2 - 2E[\hat{x}_{AV}] + 1\right) \tag{4.11}$$

27

Combining (4.1) with (4.7), (4.8), (4.10), and (4.11) gives $Var[z^*_{\text{MP}_{N(D_1,...,D_N)}}]$.

## Latin Hypercube Sampling (LH)

Under Latin Hypercube sampling, the $i^{th}$ demand value $D_i$ is uniformly distributed on $[(i-1)/N, i/N]$, and we compute $Var[z^*_{\text{MP}_{N(D_1,...,D_N)}}]$ directly:

$$
\begin{aligned}
Var[&z^*_{\text{MP}_{N(D_1,...,D_N)}}] \\
&= Var\left[\frac{1}{N}\sum_{i=1}^{N}(1-\alpha)(\hat{x}_{LH} - D_i)^+ + \alpha(D_i - \hat{x}_{LH})^+\right] \\
&= \frac{1}{N^2}Var\left[\sum_{i=1}^{\lceil\alpha N\rceil-1}(1-\alpha)(D_{\lceil\alpha N\rceil} - D_i) + \sum_{i=\lceil\alpha N\rceil+1}^{N}\alpha(D_i - D_{\lceil\alpha N\rceil})\right] \\
&= \frac{1}{N^2}Var\left[(\lceil\alpha N\rceil - 1 + \alpha - \alpha N)^2 Var[D_{\lceil\alpha N\rceil}] + (1-\alpha)^2\sum_{i=1}^{\lceil\alpha N\rceil-1} Var[D_i] + \alpha^2\sum_{i=\lceil\alpha N\rceil+1}^{N} Var[D_i]\right] \\
&= \frac{\lceil\alpha N\rceil(\lceil\alpha N\rceil - 2\alpha N - 1) + \alpha N(\alpha N - \alpha + 2)}{12N^4}.
\end{aligned}
$$

Table 5: Instance: **20** Sampling Experimental Results

| Sampling Method | $N$ | Expected Perceived Cost | | | Expected Actual Cost | | |
|---|---|---|---|---|---|---|---|
| IS | 50 | 253515.8454 | ± | 336.8149 | 254353.3385 | ± | 11.4169 |
| AV | 50 | 254203.1410 | ± | 136.6106 | 254351.0447 | ± | 9.3779 |
| LH | 50 | 254215.2169 | ± | 118.9508 | 254339.0586 | ± | 6.9906 |
| IS | 100 | 254231.3360 | ± | 273.0644 | 254337.9853 | ± | 8.6909 |
| AV | 100 | 254331.8362 | ± | 101.0049 | 254340.0692 | ± | 6.7947 |
| LH | 100 | 254231.5488 | ± | 74.8851 | 254328.8424 | ± | 5.0718 |
| IS | 500 | 254278.1028 | ± | 145.2945 | 254321.9324 | ± | 3.8601 |
| AV | 500 | 254341.3324 | ± | 42.3029 | 254320.6661 | ± | 3.3999 |
| LH | 500 | 254316.1232 | ± | 39.2008 | 254316.1177 | ± | 2.2652 |
| IS | 1000 | 254286.1589 | ± | 90.7834 | 254318.2198 | ± | 3.1933 |
| AV | 1000 | 254269.6213 | ± | 28.9737 | 254316.9865 | ± | 1.9282 |
| LH | 1000 | 254291.9907 | ± | 23.4488 | 254315.4244 | ± | 2.0295 |
| IS | 5000 | 254345.5616 | ± | 91.7814 | 254314.9654 | ± | 3.8992 |
| AV | 5000 | 254326.1383 | ± | 36.9586 | 254310.5078 | ± | 9.3001 |
| LH | 5000 | 254316.2795 | ± | 52.2501 | 254317.2299 | ± | 5.5867 |
| IS | 10000 | 254287.1841 | ± | 108.3197 | 254314.8506 | ± | 5.7181 |
| AV | 10000 | 254305.5082 | ± | 30.9672 | 254313.8079 | ± | 6.0928 |
| LH | 10000 | 254313.7684 | ± | 27.3360 | 254315.3482 | ± | 5.4347 |
| IS | 50000 | 254337.0962 | ± | 34.4324 | 254312.2884 | ± | 5.7661 |
| AV | 50000 | 254304.9466 | ± | 10.0011 | 254315.8427 | ± | 6.0459 |
| LH | 50000 | 254319.3118 | ± | 10.0143 | 254314.9763 | ± | 5.6158 |

Table 6: Instance: **fleet** Sampling Experimental Results

| Sampling Method | $N$ | Expected Perceived Cost | | | Expected Actual Cost | | |
|---|---|---|---|---|---|---|---|
| IS | 50 | -8485.6497 | ± | 75.3407 | -8429.2251 | ± | 9.9445 |
| AV | 50 | -8482.8612 | ± | 51.1750 | -8447.9424 | ± | 6.3037 |
| LH | 50 | -8474.2256 | ± | 29.3190 | -8442.4515 | ± | 0.4595 |
| IS | 100 | -8503.0462 | ± | 55.3921 | -8458.7741 | ± | 4.5802 |
| AV | 100 | -8484.0698 | ± | 27.1713 | -8463.3326 | ± | 3.8583 |
| LH | 100 | -8469.5738 | ± | 21.0250 | -8470.7080 | ± | 2.1371 |
| IS | 500 | -8501.2678 | ± | 22.4279 | -8474.6087 | ± | 1.5976 |
| AV | 500 | -8494.8387 | ± | 13.1892 | -8476.1345 | ± | 0.9500 |
| LH | 500 | -8486.5738 | ± | 9.5603 | -8477.9379 | ± | 0.6748 |
| IS | 1000 | -8492.4291 | ± | 14.2548 | -8477.0147 | ± | 1.1371 |
| AV | 1000 | -8481.4679 | ± | 9.9610 | -8478.3750 | ± | 0.6165 |
| LH | 1000 | -8482.8327 | ± | 5.6025 | -8478.0745 | ± | 0.5679 |
| IS | 5000 | -8471.2606 | ± | 24.4252 | -8478.2766 | ± | 1.8874 |
| AV | 5000 | -8479.9744 | ± | 10.3676 | -8478.3836 | ± | 2.1887 |
| LH | 5000 | -8475.0336 | ± | 8.0231 | -8478.2846 | ± | 1.2494 |
| IS | 10000 | -8464.0984 | ± | 14.8897 | -8478.5433 | ± | 1.4171 |
| AV | 10000 | -8482.4463 | ± | 8.2533 | -8479.6614 | ± | 1.8777 |
| LH | 10000 | -8479.1817 | ± | 2.9475 | -8479.0759 | ± | 0.9389 |
| IS | 50000 | -8480.4352 | ± | 6.7220 | -8479.3574 | ± | 1.1808 |
| AV | 50000 | -8478.9275 | ± | 5.7735 | -8478.2568 | ± | 1.4731 |
| LH | 50000 | -8478.4742 | ± | 2.1767 | -8478.4915 | ± | 0.6477 |


Table 7: Instance: **gbd** Sampling Experimental Results

| Sampling Method | $N$ | Expected Perceived Cost | | | Expected Actual Cost | | |
|---|---|---|---|---|---|---|---|
| IS | 50 | 1662.2925 | ± | 31.0094 | 1661.3689 | ± | 2.0441 |
| AV | 50 | 1649.5162 | ± | 20.3102 | 1659.0115 | ± | 1.5372 |
| LH | 50 | 1655.2688 | ± | 0.7773 | 1655.6289 | ± | 0.0011 |
| IS | 100 | 1633.8430 | ± | 16.9782 | 1658.4205 | ± | 1.3132 |
| AV | 100 | 1649.9295 | ± | 12.1617 | 1656.6244 | ± | 0.3580 |
| LH | 100 | 1655.6283 | ± | 0.0001 | 1655.6283 | ± | 0.0001 |
| IS | 500 | 1650.6305 | ± | 7.7778 | 1655.9825 | ± | 0.1749 |
| AV | 500 | 1654.7028 | ± | 7.0706 | 1655.7696 | ± | 0.1016 |
| LH | 500 | 1655.6283 | ± | 0.0001 | 1655.6283 | ± | 0.0001 |
| IS | 1000 | 1652.6645 | ± | 5.6258 | 1655.7667 | ± | 0.1011 |
| AV | 1000 | 1654.5455 | ± | 4.4724 | 1655.6793 | ± | 0.0629 |
| LH | 1000 | 1655.6283 | ± | 0.0001 | 1655.6283 | ± | 0.0001 |
| IS | 5000 | 1664.4786 | ± | 6.6078 | 1655.6294 | ± | 0.0041 |
| AV | 5000 | 1658.2503 | ± | 5.8087 | 1655.6278 | ± | 0.0000 |
| LH | 5000 | 1655.6284 | ± | 0.0000 | 1655.6285 | ± | 0.0000 |
| IS | 10000 | 1658.0433 | ± | 7.1031 | 1655.6278 | ± | 0.0000 |
| AV | 10000 | 1656.9337 | ± | 3.9622 | 1655.6278 | ± | 0.0000 |
| LH | 10000 | 1655.6285 | ± | 0.0000 | 1655.6285 | ± | 0.0000 |
| IS | 50000 | 1655.7737 | ± | 2.2905 | 1655.6278 | ± | 0.0000 |
| AV | 50000 | 1655.7645 | ± | 1.5347 | 1655.6278 | ± | 0.0000 |
| LH | 50000 | 1655.6285 | ± | 0.0000 | 1655.6285 | ± | 0.0000 |

Table 8: Instance: **LandS** Sampling Experimental Results

| Sampling Method | $N$ | Expected Perceived Cost | | | Expected Actual Cost | | |
|---|---|---|---|---|---|---|---|
| IS | 50 | 225.5090 | ± | 1.8534 | 225.7178 | ± | 0.0195 |
| AV | 50 | 225.2518 | ± | 0.4522 | 225.6885 | ± | 0.0153 |
| LH | 50 | 225.4986 | ± | 0.1069 | 225.6718 | ± | 0.0096 |
| IS | 100 | 224.7467 | ± | 1.3817 | 225.6777 | ± | 0.0150 |
| AV | 100 | 225.6273 | ± | 0.2776 | 225.6705 | ± | 0.0104 |
| LH | 100 | 225.6038 | ± | 0.0708 | 225.6446 | ± | 0.0037 |
| IS | 500 | 225.8406 | ± | 0.7661 | 225.6447 | ± | 0.0040 |
| AV | 500 | 225.5915 | ± | 0.0830 | 225.6354 | ± | 0.0025 |
| LH | 500 | 225.6350 | ± | 0.0299 | 225.6315 | ± | 0.0017 |
| IS | 1000 | 225.7752 | ± | 0.5853 | 225.6379 | ± | 0.0023 |
| AV | 1000 | 225.5982 | ± | 0.0353 | 225.6335 | ± | 0.0020 |
| LH | 1000 | 225.6277 | ± | 0.0242 | 225.6311 | ± | 0.0016 |
| IS | 5000 | 225.1825 | ± | 0.2695 | 225.6271 | ± | 0.0039 |
| AV | 5000 | 225.5988 | ± | 0.0472 | 225.6312 | ± | 0.0039 |
| LH | 5000 | 225.6333 | ± | 0.0287 | 225.6289 | ± | 0.0043 |
| IS | 10000 | 225.8143 | ± | 0.4435 | 225.6312 | ± | 0.0053 |
| AV | 10000 | 225.6274 | ± | 0.0282 | 225.6283 | ± | 0.0050 |
| LH | 10000 | 225.6400 | ± | 0.0130 | 225.6329 | ± | 0.0020 |
| IS | 50000 | 225.6558 | ± | 0.1345 | 225.6314 | ± | 0.0054 |
| AV | 50000 | 225.6315 | ± | 0.0159 | 225.6308 | ± | 0.0054 |
| LH | 50000 | 225.6244 | ± | 0.0069 | 225.6297 | ± | 0.0059 |

Table 9: Instance: **snip** Sampling Experimental Results

| Sampling Method | $N$ | Expected Perceived Cost | | | Expected Actual Cost | | |
|---|---|---|---|---|---|---|---|
| IS | 50 | 86.7569 | ± | 0.8627 | 88.9335 | ± | 0.3354 |
| AV | 50 | 87.1225 | ± | 0.9321 | 88.8618 | ± | 0.2891 |
| LH | 50 | 88.0046 | ± | 0.3398 | 88.1592 | ± | 0.0679 |
| IS | 100 | 86.6793 | ± | 0.7723 | 88.5825 | ± | 0.1328 |
| AV | 100 | 87.7957 | ± | 0.7176 | 88.2950 | ± | 0.0908 |
| LH | 100 | 88.1680 | ± | 0.1924 | 88.2457 | ± | 0.0878 |
| IS | 500 | 87.9032 | ± | 0.3421 | 88.1996 | ± | 0.0594 |
| AV | 500 | 87.9938 | ± | 0.2066 | 88.1439 | ± | 0.0207 |
| LH | 500 | 88.1638 | ± | 0.0950 | 88.1489 | ± | 0.0260 |
| IS | 1000 | 87.9584 | ± | 0.2275 | 88.1297 | ± | 0.0071 |
| AV | 1000 | 88.0234 | ± | 0.2022 | 88.1239 | ± | 0.0046 |
| LH | 1000 | 88.1152 | ± | 0.0516 | 88.1238 | ± | 0.0042 |
| IS | 5000 | 88.1936 | ± | 0.2808 | 88.1244 | ± | 0.0128 |
| AV | 5000 | 88.0438 | ± | 0.3386 | 88.1273 | ± | 0.0156 |
| LH | 5000 | 88.1572 | ± | 0.0941 | 88.1225 | ± | 0.0098 |
| IS | 10000 | 88.0736 | ± | 0.2214 | 88.1234 | ± | 0.0121 |
| AV | 10000 | 88.1650 | ± | 0.1555 | 88.1215 | ± | 0.0094 |
| LH | 10000 | 88.1285 | ± | 0.0461 | 88.1199 | ± | 0.0122 |
| IS | 50000 | 88.1425 | ± | 0.1308 | 88.1303 | ± | 0.0144 |
| AV | 50000 | 88.1134 | ± | 0.0784 | 88.1241 | ± | 0.0081 |
| LH | 50000 | 88.1053 | ± | 0.0253 | 88.1201 | ± | 0.0062 |

Table 10: Instance: **ssn** Sampling Experimental Results

| Sampling Method | $N$ | Expected Perceived Cost | | | Expected Actual Cost | | |
|---|---|---|---|---|---|---|---|
| IS | 50 | 4.5957 | $\pm$ | 0.5450 | 14.0593 | $\pm$ | 0.3396 |
| AV | 50 | 4.8005 | $\pm$ | 0.5866 | 14.1261 | $\pm$ | 0.3873 |
| LH | 50 | 9.0945 | $\pm$ | 0.4887 | 10.2504 | $\pm$ | 0.0791 |
| IS | 100 | 6.9280 | $\pm$ | 0.4918 | 12.0752 | $\pm$ | 0.1766 |
| AV | 100 | 6.9832 | $\pm$ | 0.6046 | 12.2632 | $\pm$ | 0.2061 |
| LH | 100 | 9.8333 | $\pm$ | 0.3454 | 10.2244 | $\pm$ | 0.0592 |
| IS | 500 | 9.0208 | $\pm$ | 0.2139 | 10.3871 | $\pm$ | 0.0483 |
| AV | 500 | 9.0206 | $\pm$ | 0.2221 | 10.3596 | $\pm$ | 0.0464 |
| LH | 500 | 9.9076 | $\pm$ | 0.1322 | 10.0161 | $\pm$ | 0.0217 |
| IS | 1000 | 9.4323 | $\pm$ | 0.1508 | 10.1897 | $\pm$ | 0.0289 |
| AV | 1000 | 9.1977 | $\pm$ | 0.1538 | 10.1629 | $\pm$ | 0.0274 |
| LH | 1000 | 9.7371 | $\pm$ | 0.1102 | 10.0186 | $\pm$ | 0.0115 |
| IS | 5000 | 9.7719 | $\pm$ | 0.2473 | 9.9522 | $\pm$ | 0.0175 |
| AV | 5000 | 9.6999 | $\pm$ | 0.1787 | 9.9363 | $\pm$ | 0.0242 |
| LH | 5000 | 9.9226 | $\pm$ | 0.0910 | 9.9153 | $\pm$ | 0.0245 |
| IS | 10000 | 9.8007 | $\pm$ | 0.1407 | 9.9236 | $\pm$ | 0.0200 |
| AV | 10000 | 9.8716 | $\pm$ | 0.1064 | 9.9397 | $\pm$ | 0.0255 |
| LH | 10000 | 9.9125 | $\pm$ | 0.1610 | 9.8988 | $\pm$ | 0.0278 |
| IS | 50000 | 9.8224 | $\pm$ | 0.0654 | 9.9174 | $\pm$ | 0.0181 |
| AV | 50000 | 9.8916 | $\pm$ | 0.1016 | 9.8948 | $\pm$ | 0.0296 |
| LH | 50000 | 9.9036 | $\pm$ | 0.0441 | 9.8879 | $\pm$ | 0.0219 |

Table 11: Instance: **storm** Sampling Experimental Results

| Sampling Method | $N$ | Expected Perceived Cost | | | Expected Actual Cost | | |
|---|---|---|---|---|---|---|---|
| IS | 50 | 15498.1040 | $\pm$ | 14.1757 | 15499.0977 | $\pm$ | 0.1084 |
| AV | 50 | 15497.8955 | $\pm$ | 2.5719 | 15498.8736 | $\pm$ | 0.0467 |
| LH | 50 | 15498.5191 | $\pm$ | 0.3829 | 15498.8026 | $\pm$ | 0.0367 |
| IS | 100 | 15500.5301 | $\pm$ | 8.2347 | 15498.9087 | $\pm$ | 0.0708 |
| AV | 100 | 15499.8339 | $\pm$ | 1.3689 | 15498.8177 | $\pm$ | 0.0404 |
| LH | 100 | 15499.0448 | $\pm$ | 0.2643 | 15498.7357 | $\pm$ | 0.0189 |
| IS | 500 | 15496.6425 | $\pm$ | 3.8410 | 15498.7721 | $\pm$ | 0.0246 |
| AV | 500 | 15498.6479 | $\pm$ | 0.4858 | 15498.7355 | $\pm$ | 0.0065 |
| LH | 500 | 15498.7878 | $\pm$ | 0.1381 | 15498.7281 | $\pm$ | 0.0056 |
| IS | 1000 | 15499.7652 | $\pm$ | 2.5354 | 15498.7369 | $\pm$ | 0.0110 |
| AV | 1000 | 15498.7001 | $\pm$ | 0.1510 | 15498.7281 | $\pm$ | 0.0062 |
| LH | 1000 | 15498.7242 | $\pm$ | 0.0711 | 15498.7248 | $\pm$ | 0.0057 |
| IS | 5000 | 15499.2379 | $\pm$ | 2.9851 | 15498.7285 | $\pm$ | 0.0138 |
| AV | 5000 | 15498.7462 | $\pm$ | 0.1444 | 15498.7343 | $\pm$ | 0.0171 |
| LH | 5000 | 15498.7515 | $\pm$ | 0.1100 | 15498.7322 | $\pm$ | 0.0170 |
| IS | 10000 | 15499.1428 | $\pm$ | 2.1525 | 15498.7328 | $\pm$ | 0.0158 |
| AV | 10000 | 15498.7847 | $\pm$ | 0.1868 | 15498.7416 | $\pm$ | 0.0148 |
| LH | 10000 | 15498.7181 | $\pm$ | 0.0702 | 15498.7186 | $\pm$ | 0.0192 |
| IS | 50000 | 15498.8456 | $\pm$ | 1.5608 | 15498.7153 | $\pm$ | 0.0215 |
| AV | 50000 | 15498.7431 | $\pm$ | 0.0784 | 15498.7348 | $\pm$ | 0.0090 |
| LH | 50000 | 15498.7214 | $\pm$ | 0.0364 | 15498.7208 | $\pm$ | 0.0192 |
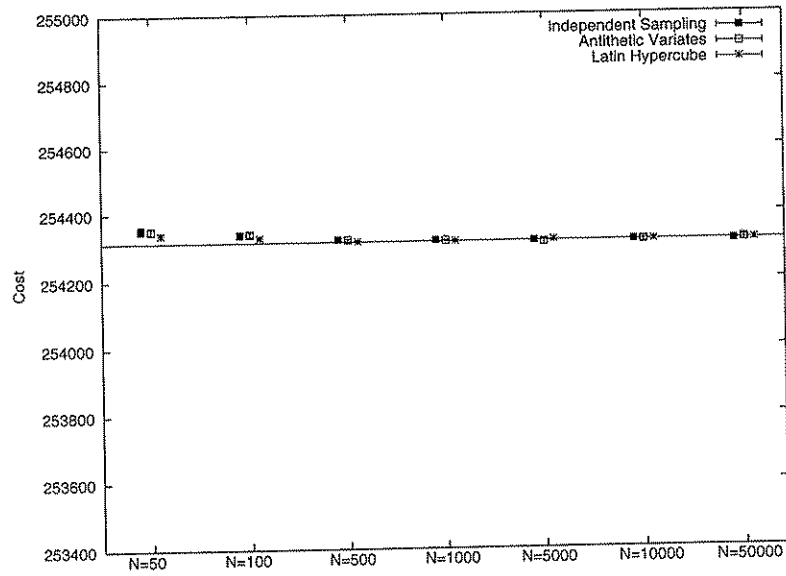
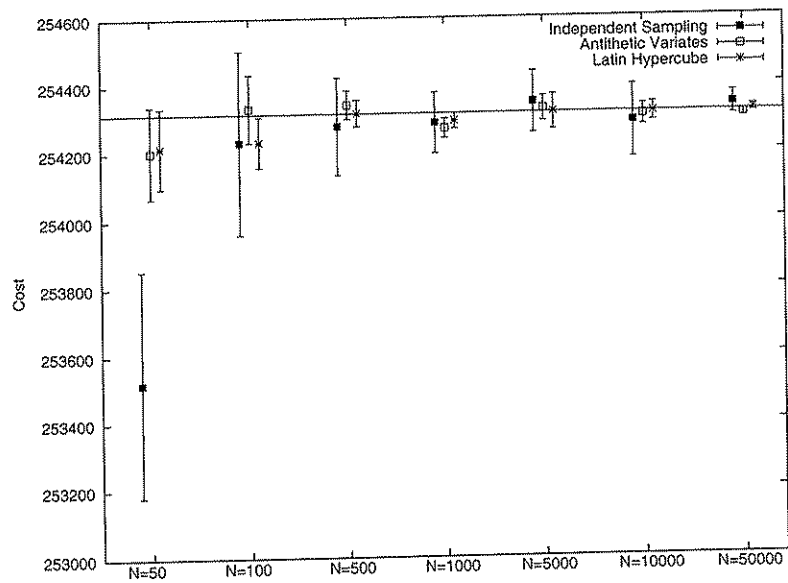Figure 10: Expected Actual Cost Estimates for **20**



Figure 11: Expected Perceived Cost Estimates for **20**
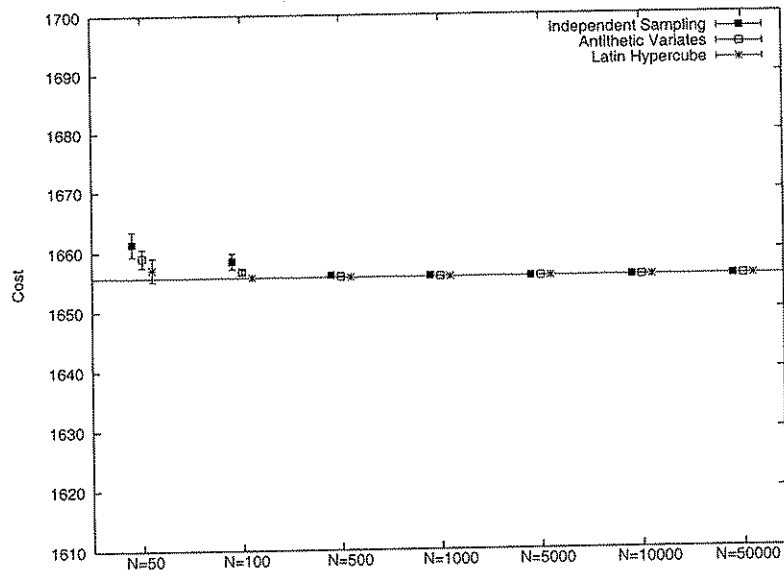
Figure 12: Expected Actual Cost Estimates for **gbd**
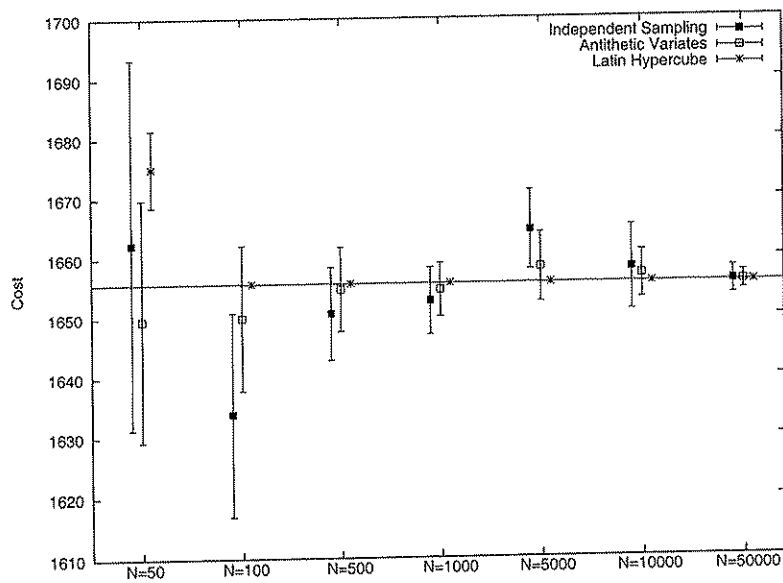


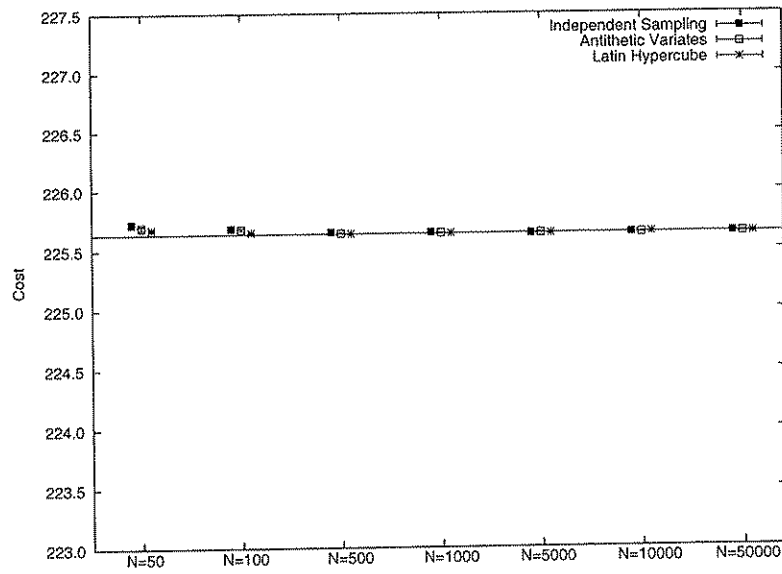Figure 13: Expected Perceived Cost Estimates for **gbd**

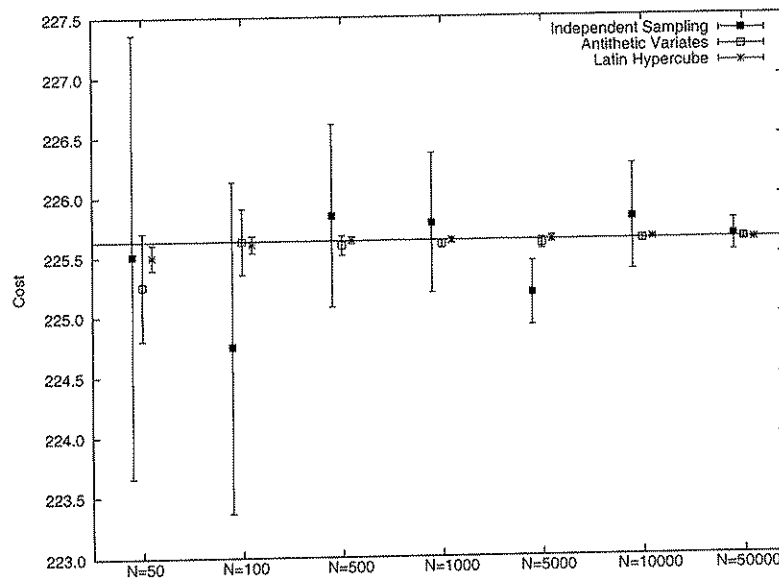Figure 14: Expected Actual Cost Estimates for **LandS**



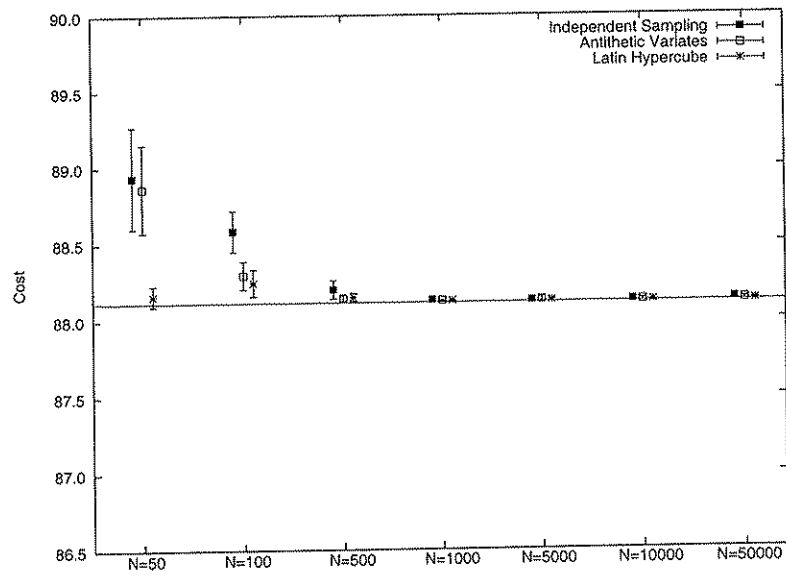Figure 15: Expected Perceived Cost Estimates for **LandS**

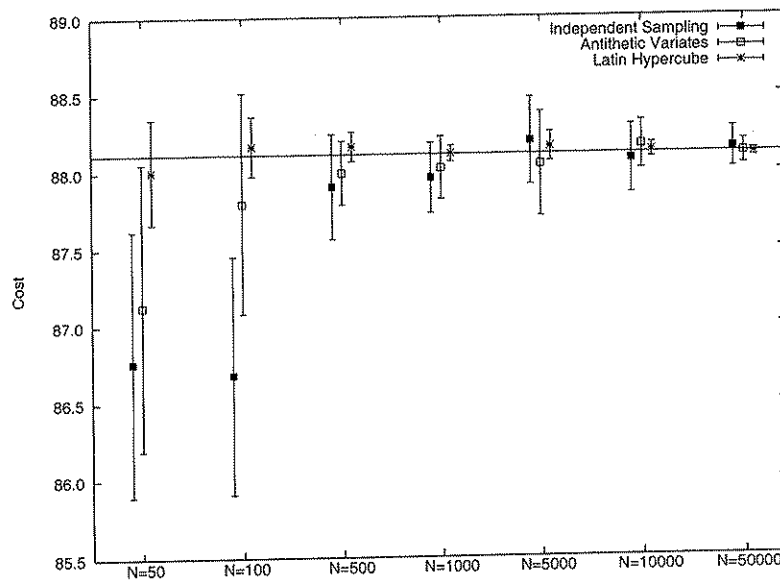Figure 16: Expected Actual Cost Estimates for **snip**



Figure 17: Expected Perceived Cost Estimates for **snip**
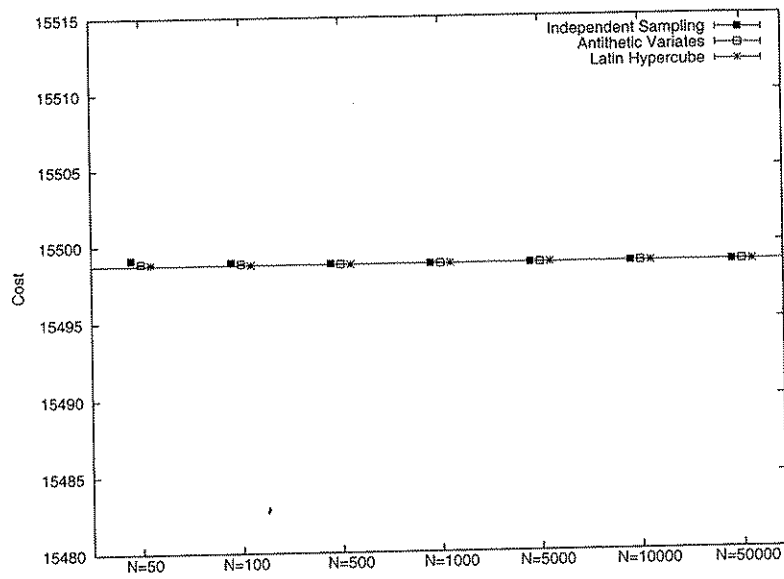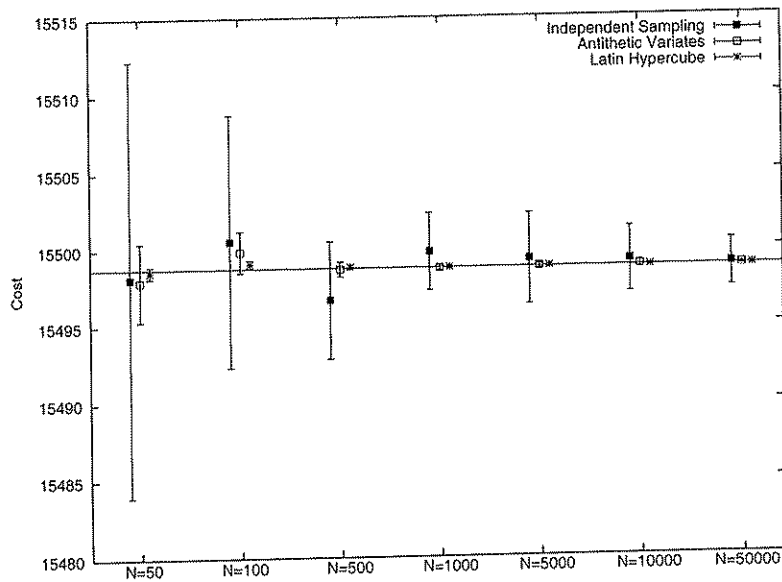
Figure 18: Expected Actual Cost Estimates for **storm**



Figure 19: Expected Perceived Cost Estimates for **storm**