

## JACKKNIFE ESTIMATORS FOR REDUCING BIAS IN ASSET ALLOCATION

Amit Partani  
David P. Morton

Graduate Program in Operations Research  
The University of Texas at Austin  
Austin, TX 78712, U.S.A.

Ivilina Popova

Albers School of Business and Economics  
Seattle University  
Seattle, WA 98122, U.S.A.

### ABSTRACT

We use jackknife-based estimators to reduce bias when estimating the optimal value of a stochastic program. Our discussion focuses on an asset allocation model with a power utility function. As we will describe, estimating the optimal value of such a problem plays a key role in establishing the quality of a candidate solution, and reducing bias improves our ability to do so efficiently. We develop a jackknife estimator that is adaptive in that it does not assume the order of the bias is known a priori.

### 1 INTRODUCTION

Monte Carlo simulation is used in assessing whether a candidate solution to a stochastic program is near optimal, when we cannot solve the stochastic program exactly. Optimizing a sample-mean estimator yields an optimistic bound, in expectation, on the original problem's optimal value, say,  $z^*$ . Restated, maximizing a sample mean yields a positively-biased estimate of  $z^*$ . Such bounds are analogous to optimistic bounds that arise, e.g., via Lagrangian or integrality relaxations in deterministic integer and non-linear programming. A weak relaxation bound yields a weak statement regarding the quality of a candidate solution. Similarly, when our estimate of the optimal value has large bias, it can significantly degrade our ability to establish that a candidate solution is near optimal. So, in this paper we seek to tighten an optimized sample-mean bound, reducing its bias using jackknife-based estimators. Unlike existing jackknife estimators, we do not assume the order of the bias is known when forming the estimator. Our approach is illustrated on an asset allocation model with a power utility function.

An overview of stochastic programming, and other approaches to problems that arise in finance, including pricing single instruments, time-static asset allocation, and time-dynamic asset-liability management is given in Ziemba and Mulvey (1998). Monte Carlo methods have seen extensive

application to pricing financial securities; see, e.g., Andersen and Broadie (2004), Fu et al. (2001), and Glasserman (2003). Asymptotic justification of replacing population means with sample-mean estimators in portfolio optimization is discussed in Jensen and King (1992). The type of simulation-based solution-quality assessments we make here have also been done in the context of portfolio optimization in Morton et al. (2003, 2006), but those papers do not include bias-reducing estimators.

We consider a portfolio allocation problem with  $m$  assets. These assets have random returns  $\xi = (\xi_1, \dots, \xi_m) \geq 0$ . The allocation problem selects the proportion,  $x = (x_1, \dots, x_m)$ , to invest in each asset to maximize expected utility. The random return of allocation  $x$  is  $\xi x \equiv \sum_{j=1}^m \xi_j x_j$ . We assume that  $\xi$ 's distribution is known, has finite second moments, does not depend on  $x$ , and that we can generate independent and identically distributed (i.i.d.) observations of  $\xi$ . If  $\xi_1$  takes value 1.5 and  $\xi_2$  takes value 0.75 then we have a 50% return on the first asset and a loss of 25% on the second asset. The asset allocation problem is

$$z^* = \max_{x \in X} Eu(x, \xi), \quad (1)$$

where  $X = \{x : \sum_{j=1}^m x_j = 1, x_j \geq 0, j = 1, \dots, m\}$  and  $u(x, \xi) = (\xi x)^\gamma - c \|x - x^t\|_2^2$ . The constraint set  $X$  requires all proportions sum to one and disallows shortselling.  $Eu(x, \xi)$  is the expected utility obtained by allocation  $x$ .

Our primary objective is to maximize the expected value of the power utility function of return,  $u_p(\cdot) = (\cdot)^\gamma$ , where  $\gamma \in (0, 1)$  captures the decision maker's aversion to risk. This is augmented by a secondary penalty term with  $0 \leq c \ll 1$ . This discourages deviation from the investor's existing target portfolio,  $x^t \in X$ , unless the power utility provides a reason to do so. If  $x^t$  is unknown or the secondary objective is not desired we set  $c = 0$ . The approach we describe applies to a wide variety of utility functions, but in this paper we restrict ourselves to this variant of the power utility. The power utility function exhibits increasing

absolute risk aversion (defined as  $R_A(z) = -u''_p(z)/u'_p(z)$ ) and constant relative risk aversion (defined as  $R_R(z) = R_A(z)z$ ). Under a utility function with constant relative risk aversion, the fraction of initial wealth that one invests in risky assets does not change as initial wealth is increased.

$Eu(x, \xi)$  is well defined for all  $x \in X$  since  $\xi$  is integrable. The feasible region of (1) is compact, and  $u^o(x) = Eu(x, \xi)$  is continuous on  $X$ . Hence, a finite optimal solution of (1) is achieved for some  $x^* \in X$ .

## 2 QUALITY-ASSESSMENT PROCEDURE

In this section, we discuss using Monte Carlo simulation to approximately solve (1) and review a simulation-based multiple replication procedure (MRP) used to assess the quality of a candidate solution. It is this MRP that our jackknife estimators of Section 3 aim to improve.

### 2.1 Monte Carlo Simulation Approach

Our ability to solve (1) rests, in part, on our ability to evaluate  $Eu(x, \xi)$ . Unless the distribution of  $\xi$  is particularly simple, it is usually impossible to evaluate  $Eu(x, \xi)$  exactly, even for a fixed allocation vector  $x \in X$ . When exact computation is not viable, it is natural to replace  $Eu(x, \xi)$  with a sampling-based estimator, which lends itself to computation. This paper considers estimators rooted in the sample mean of i.i.d. observations. With  $\xi^i, i = 1, \dots, n$ , i.i.d. as  $\xi$ , we can approximately solve (1) by solving

$$z_n^* = \max_{x \in X} \left[ \bar{u}_n(x) \equiv \frac{1}{n} \sum_{i=1}^n u(x, \xi^i) \right]. \tag{2}$$

Let  $x_n^*$  solve (2). Our utility function satisfies

$$u(x, \xi) \leq (\xi_1 + \dots + \xi_m)^\gamma, \forall x \in X. \tag{3}$$

Thus,  $u(x, \xi)$  is uniformly (in  $x \in X$ ) bounded by the right-hand side of (3), which is integrable since concavity of  $(\cdot)^\gamma$  implies  $E(\xi_1 + \dots + \xi_m)^\gamma \leq (E\xi_1 + \dots + E\xi_m)^\gamma$ . This, coupled with the fact that  $u^o(x) = Eu(x, \xi)$  is continuous and  $X$  is compact, is sufficient (Attouch and Wets 1990, Theorem 4.3) to ensure that i.i.d. sampling yields: (i)  $\lim_{n \rightarrow \infty} x_n^* = x^*$ , w.p.1, and (ii)  $\lim_{n \rightarrow \infty} z_n^* = z^*$ , w.p.1. Result (i) is often stated in the weaker form “all accumulation points of  $\{x_n^*\}$  solve (1), w.p.1.” However, the objective functions of (1) and (2) are strictly concave and  $X$  is compact, so  $x^*$  is unique and  $x_n^*$  is unique, given  $\xi^1, \dots, \xi^n$ . For more on such convergence results, as well as results concerning rates of convergence, see the survey of Shapiro (2003) or Dupačová and Wets (1988) and King and Rockafellar (1993). The following results are from Mak et al. (1999) and Norkin et al. (1998): (iii)

$Ez_n^* \geq z^*$ , and (iv)  $Ez_n^* \geq Ez_{n+1}^*$ . Result (iii) holds provided  $E\bar{u}_n(x) = Eu(x, \xi)$  and (iv) provided the sampling is i.i.d. These show that the optimal value of (2) has positive bias and that this bias shrinks as  $n$  grows.

### 2.2 Assessing Solution Quality

Given a feasible candidate solution to (1),  $\hat{x} \in X$ , it is natural to wish to assess  $\hat{x}$ 's quality. The candidate solution  $\hat{x}$  can be obtained by solving (2) or by any other means. Here, we measure quality via the optimality gap,  $\mu_{\hat{x}} = z^* - Eu(\hat{x}, \xi)$ . We view  $\mu_{\hat{x}}$  as an unknown parameter and we seek point and interval estimators for  $\mu_{\hat{x}}$ . We can estimate  $Eu(\hat{x}, \xi)$  by  $\bar{u}_n(\hat{x})$  and  $z^*$  by solving (2) for  $z_n^*$ . The former estimator is unbiased but that latter has positive bias. As a result, we obtain a positively-biased estimator of  $\mu_{\hat{x}}$  by  $G_n(\hat{x}) = z_n^* - \bar{u}_n(\hat{x})$ , i.e.,  $EG_n(\hat{x}) \geq \mu_{\hat{x}} \geq 0$ . So, if we can infer  $EG_n(\hat{x})$  is small then we can infer the same for  $\mu_{\hat{x}}$ , i.e., that  $\hat{x}$  is a high-quality solution. This leads us to seek an approximate one-sided  $\alpha$ -level confidence interval (CI) of the form

$$P(EG_n(\hat{x}) \leq \epsilon) \approx \alpha. \tag{4}$$

Here,  $\epsilon$  is a random CI width and we choose, e.g.,  $\alpha = 0.95$ . If (4) holds with equality then  $P(\mu_{\hat{x}} \leq \epsilon) \geq \alpha$ . The following procedure from Mak et al. (1998) forms i.i.d. replicates of  $G_n(\hat{x})$ . (We note that even though  $G_n(\hat{x})$ , in general, can be non-normal, in Bayraksan and Morton 2006 we develop an asymptotically-valid single-replication procedure.)

#### Multiple Replication Procedure (MRP<sup>o</sup>)

*Input:* Value  $\alpha \in (0, 1)$ , sample size  $n$ , replication size  $r$ , and candidate solution  $\hat{x} \in X$ .

*Output:* Approximate  $\alpha$ -level CI on  $\mu_{\hat{x}}$ .

1. For  $k = 1, \dots, r$ ,
  - Generate  $\xi^{k1}, \dots, \xi^{kn}$  i.i.d. as  $\xi$ .
  - Let  $\bar{u}_n(x) = \frac{1}{n} \sum_{i=1}^n u(x, \xi^{ki})$ .
  - Let  $G_n^k(\hat{x}) = \max_{x \in X} \bar{u}_n(x) - \bar{u}_n(\hat{x})$ .
2. Calculate gap estimate and sample variance by

$$\bar{G}_{rn}^o(\hat{x}) = \frac{1}{r} \sum_{k=1}^r G_n^k(\hat{x}), \text{ and}$$

$$s_o^2(\hat{x}) = \frac{1}{r-1} \sum_{k=1}^r (G_n^k(\hat{x}) - \bar{G}_{rn}^o(\hat{x}))^2.$$

3. Form CI for  $\mu_{\hat{x}}$  as,

$$\left[ 0, \bar{G}_{rn}^o(\hat{x}) + \frac{t_{r-1, \alpha} s_o(\hat{x})}{\sqrt{r}} \right].$$

Here,  $t_{r-1,\alpha}$  is the  $\alpha$  quantile of a  $t$  random variable with  $r - 1$  degrees of freedom. Validity of the MRP $^o$ 's CI on  $\mu_{\hat{x}} = z^* - Eu(\hat{x}, \xi)$  is due to the CLT

$$\sqrt{r} [\bar{G}_{rn}^o(\hat{x}) - EG_n(\hat{x})] \Rightarrow N(0, \sigma_o^2(\hat{x})) \text{ as } r \rightarrow \infty,$$

the fact that  $s_o^2(\hat{x})$  is a consistent estimator of  $\sigma_o^2(\hat{x}) = \text{var } G_n(\hat{x})$  and  $EG_n(\hat{x}) \geq \mu_{\hat{x}}$ . In terms of (4),  $\epsilon = \bar{G}_{rn}^o(\hat{x}) + t_{r-1,\alpha} s_o(\hat{x})/\sqrt{r}$ . Three factors, which we express in their population forms, contribute to this CI width: (i)  $z^* - Eu(\hat{x}, \xi)$ , i.e., the true optimality gap, (ii) the sampling error  $t_{r-1,\alpha} \sigma_o(\hat{x})/\sqrt{r}$ , and (iii) the bias which arises due to replacing  $z^*$  by  $z_n^*$ , i.e.,  $b(z_n^*) = Ez_n^* - z^*$ . It is our experience that bias often dominates the other factors, and in the remainder of this paper we propose estimators to reduce this bias.

### 3 JACKKNIFE ESTIMATORS

As indicated above, the bias of estimator  $z_n^*$  can significantly degrade our ability to assess the quality of a candidate asset allocation decision. Of course, one way to reduce this bias is to increase the sample size  $n$ , but this can be computationally expensive since the effort to solve the approximation problem (2) typically grows faster than linear with  $n$ . Furthermore, unless overly restrictive assumptions are made, it appears unlikely that an analytical expression for the asymptotic form of the bias can be derived. So, in this section we attempt to reduce bias using jackknife estimators.

#### 3.1 Generalized Jackknife Estimator

The jackknife estimator developed by Quenouille (1949) has become an important tool in simulation and data analysis for reducing bias. The standard jackknife estimator is used to eliminate  $O(n^{-1})$  bias and we briefly review its derivation as it motivates our adaptive estimator. Suppose  $g_n$  is an estimator of  $g_\mu$  based on  $n$  i.i.d. observations and assume  $b(g_n) = O(n^{-1})$ . More specifically, consider

$$Eg_n = g_\mu + a_1/n + a_2/n^2 + \dots \tag{5a}$$

$$Eg_m = g_\mu + a_1/m + a_2/m^2 + \dots, \tag{5b}$$

where  $m < n$ . We multiply (5a) by  $n$ , (5b) by  $m$  and subtract the latter from the former to obtain  $E[(ng_n - mg_m)/(n - m)] = g_\mu + O((nm)^{-1})$ . For reasons that become clear in Section 3.2, we assume  $n$  is even and use  $m = \frac{n}{2}$ . We view this derivation as solving the two equations in (5) for two unknowns, namely  $g_\mu$  and  $a_1$ , and this suggests estimating  $g_\mu$  by  $J_n = (2ng_n - ng_{\frac{n}{2}})/n$ , which has lower-order bias than  $g_n$ .

In Bayraksan et al. (2006), we show that the approximating problem (2) can have  $b(z_n^*) = O(n^{-p})$  for any

$p \in [1/2, \infty)$ . This motivates assuming

$$b(z_n^*) = Ez_n^* - z^* = a_1/n^p + o(n^{-p}). \tag{6}$$

Mimicking the above derivation of the standard jackknife estimator leads to the generalized jackknife estimator as developed by Gray and Schucany (1972)

$$J_n^p = \frac{(2n)^p z_n^* - n^p z_{\frac{n}{2}}^*}{(2n)^p - n^p}, \tag{7}$$

where  $z_n^*$  is the optimal value to (2) and  $z_{\frac{n}{2}}^*$  is the optimal value of (2) with half the observations deleted, i.e.,

$$z_{\frac{n}{2}}^* = \max_{x \in X} \frac{2}{n} \sum_{i=1}^{\frac{n}{2}} u(x, \xi^i).$$

Of course,  $\xi^1, \dots, \xi^n$  are i.i.d. and so  $z_{\frac{n}{2}}^*$  could be formed with any subset of the observations of size  $\frac{n}{2}$ . In the procedure below we replace  $z_{\frac{n}{2}}^*$  in (7) by the average of two observations based on  $\xi^1, \dots, \xi^{\frac{n}{2}}$  and  $\xi^{\frac{n}{2}+1}, \dots, \xi^n$ .

Unfortunately, we are unlikely to know the true order of the bias,  $p$ . We will form the generalized jackknife estimator using parameter  $q$  and reserve notation  $p$  to denote the true (and unknown) order of the bias. The following simple but useful result rests only on the monotonicity property  $Ez_n^* \leq Ez_{\frac{n}{2}}^*$ , and helps guide our choice of  $q$ .

**Theorem 1** *Let  $0 < q_1 < q_2$ . Then  $EJ_n^{q_1} \leq EJ_n^{q_2} \leq Ez_n^*$  and  $\lim_{q \rightarrow \infty} EJ_n^q = Ez_n^*$ .*

As  $q \rightarrow \infty$  for fixed  $n$ ,  $J_n^q \rightarrow z_n^*$ , w.p.1, i.e., large values of  $q$  effectively correspond to our original upper bound estimator and smaller values of  $q$  are more aggressive with respect to removing bias, with the risk that the corresponding ‘‘upper bound’’ estimator is smaller than  $z^*$ . The following multiple replication procedure with generalized jackknife estimators uses the above ideas except that the jackknife estimator is expressed in terms of the underlying gap estimator instead of  $z_n^*$ .

#### Generalized Jackknife MRP (MRP $^q$ )

*Input:* Value  $\alpha \in (0, 1)$ , sample size  $n$ , replication size  $r$ , jackknife parameter  $q$ , and candidate solution  $\hat{x} \in X$ .

*Output:* Approximate  $\alpha$ -level CI on  $\mu_{\hat{x}}$ .

1. For  $k = 1, \dots, r$ ,
  - Generate  $\tilde{\xi}^{k1}, \dots, \tilde{\xi}^{kn}$  i.i.d. as  $\xi$ .

- Let

$$\begin{aligned}
 G_n^k(\hat{x}) &= \max_{x \in X} \frac{1}{n} \sum_{i=1}^n u(x, \xi^{ki}) - \\
 &\quad \frac{1}{n} \sum_{i=1}^n u(\hat{x}, \xi^{ki}), \\
 G_{\frac{n}{2}}^{k,1}(\hat{x}) &= \max_{x \in X} \frac{2}{n} \sum_{i=1}^{\frac{n}{2}} u(x, \xi^{ki}) - \\
 &\quad \frac{2}{n} \sum_{i=1}^{\frac{n}{2}} u(\hat{x}, \xi^{ki}), \\
 G_{\frac{n}{2}}^{k,2}(\hat{x}) &= \max_{x \in X} \frac{2}{n} \sum_{i=\frac{n}{2}+1}^n u(x, \xi^{ki}) - \\
 &\quad \frac{2}{n} \sum_{i=\frac{n}{2}+1}^n u(\hat{x}, \xi^{ki}), \\
 \bar{G}_{\frac{n}{2}}^k(\hat{x}) &= \frac{1}{2} \left[ G_{\frac{n}{2}}^{k,1}(\hat{x}) + G_{\frac{n}{2}}^{k,2}(\hat{x}) \right].
 \end{aligned}$$

- Form

$$J_n^{qk} = \frac{(2n)^q G_n^k(\hat{x}) - n^q \bar{G}_{\frac{n}{2}}^k(\hat{x})}{(2n)^q - n^q}.$$

2. Calculate gap estimate and sample variance by

$$\bar{G}_{rn}^q(\hat{x}) = \frac{1}{r} \sum_{k=1}^r J_n^{qk}, \quad \text{and}$$

$$s_q^2(\hat{x}) = \frac{1}{r-1} \sum_{k=1}^r (J_n^{qk} - \bar{G}_{rn}^q(\hat{x}))^2.$$

3. Form CI for  $\mu_{\hat{x}}$  as,

$$\left[ 0, \bar{G}_{rn}^q(\hat{x}) + \frac{t_{r-1, \alpha} s_q(\hat{x})}{\sqrt{r}} \right].$$

The following quantity,  $\rho$ , compares the asymptotic decrease in the bias of the generalized jackknife estimator relative to the bias of the original estimator in the MRP<sup>o</sup> of Section 2.2:

$$\rho = \left| \lim_{n \rightarrow \infty} \lim_{r \rightarrow \infty} \frac{\bar{G}_{rn}^q(\hat{x}) - \mu_{\hat{x}}}{G_{rn}^o(\hat{x}) - \mu_{\hat{x}}} \right|.$$

The following result is adapted to our setting from Gray and Schucany (1972, Theorem 3.4).

**Theorem 2** Consider the gap estimator of the MRP<sup>q</sup>. Assume  $b(z_n^*) = a_1 n^{-p_1} + a_2 n^{-p_2} + o(n^{-p_2})$  where  $p_2 > p_1$  and  $a_1 \neq 0$ . Then,

- (i) if  $q = p_1$  then  $\rho = 0$ ;
- (ii) if  $q > \frac{p_1}{2}$  and  $q \neq p_1$  then  $0 < \rho < 1$ ;
- (iii) if  $q = \frac{p_1}{2}$  then  $\rho = 1$ ; and,
- (iv) if  $0 < q < \frac{p_1}{2}$  then  $1 < \rho < \infty$ .

Typically, when attempting to reduce bias via the generalized jackknife, one derives or postulates an asymptotic bias expansion (e.g., (5) or (6) or otherwise), and then employs the order- $p$  jackknife if the expansion is  $O(n^{-p})$ . However, we seek a one-sided CI on  $\mu_{\hat{x}}$  and in reducing bias in  $z_n^*$  we prefer to be conservative and err on the “high” side. Theorem 1 provides guidance in this regard. If  $b(z_n^*) = O(n^{-p})$  we prefer to use MRP<sup>q</sup> with  $q > p$ . Theorem 2 ensures we will achieve (asymptotic) bias reduction by using these conservative values of  $q$ . This approach is complicated by the fact that  $p$  is unknown, and so we address this issue next.

### 3.2 Adaptive Jackknife Estimator

The adaptive jackknife estimator is derived as follows. We assume  $b(z_n^*)$  satisfies (6). We let  $n$  be an integral multiple of 4 and form

$$\begin{aligned}
 E z_n^* &= z^* + a_1/n^{-p} + o(n^{-p}) \\
 E z_{\frac{n}{2}}^* &= z^* + a_1/(n/2)^{-p} + o(n^{-p}) \\
 E z_{\frac{n}{4}}^* &= z^* + a_1/(n/4)^{-p} + o(n^{-p}).
 \end{aligned}$$

We view these as three equations in three unknowns,  $z^*$ ,  $a_1$  and  $p$ . Solving for  $p$  and dropping the  $o(\cdot)$  terms, yields

$$p = \log_2 \left( \frac{E z_{\frac{n}{4}}^* - E z_{\frac{n}{2}}^*}{E z_{\frac{n}{2}}^* - E z_n^*} \right). \tag{8}$$

Our estimate of  $p$  replaces the expectations on the right-hand sides of (8) with sample means. We do this in the following manner: Given a sample  $\xi^1, \dots, \xi^n$ , we partition the sample into two subsamples of size  $n/2$  and, in turn, partition those subsamples into four subsamples of size  $n/4$ . We then form a single observation of  $z_n^*$ , average two i.i.d. estimates of  $z_{\frac{n}{2}}^*$ , and average four i.i.d. estimates of  $z_{\frac{n}{4}}^*$ . Repeating this with  $r$  i.i.d. replications we form estimators we denote  $\bar{z}_{rn}^*$ ,  $\bar{z}_{r\frac{n}{2}}^*$ , and  $\bar{z}_{r\frac{n}{4}}^*$ . These sample-mean estimators replace their population counterparts in (8) to yield

$$p_{rn}^a = \log_2 \left( \frac{\bar{z}_{r\frac{n}{4}}^* - \bar{z}_{r\frac{n}{2}}^*}{\bar{z}_{r\frac{n}{2}}^* - \bar{z}_{rn}^*} \right),$$

which is a nonlinear function,  $p_{rn}^a = f(Z_{rn})$ , of three sample means  $Z_{rn} = (\bar{z}_{r\frac{n}{4}}^*, \bar{z}_{r\frac{n}{2}}^*, \bar{z}_{rn}^*)$ . Let  $\hat{\Sigma}$  denote the standard sample covariance estimator of  $Z_{rn}$ . Using a first-order Taylor series expansion, we can estimate the variance

of  $p_{rn}^a$  as

$$s_p^2 = \nabla^T f(Z_{rn}) \hat{\Sigma} \nabla f(Z_{rn}).$$

We could correct for bias in  $p_{rn}^a$  using a second-order Taylor series, but we do not do so because when we tested this for problem instances described in the next section, the associated bias corrections for  $p$  were of negligible magnitude.

We could form an adaptive jackknife procedure by simply using  $q = p_{rn}^a$  in  $MRP^q$ . We do not do so for the reasons described at the end of Section 3.1, i.e., we seek a conservative procedure. We instead let

$$q = \max\{p_{rn}^a, \frac{1}{2}\} + t_{r-1, \beta} s_p. \tag{9}$$

We know from Theorem 1 that our bias reduction is less aggressive for larger values of  $q$  and so as  $\beta$  grows toward one, and the corresponding  $t$  quantile grows, our procedure becomes more conservative. With  $\beta$  at our disposal we examine the performance of a family of adaptive jackknife estimators in the next section. Of course, the freedom to choose a parameter such as  $\beta$  can be disconcerting to some. In this case, we recommend choosing  $\beta = \frac{1}{2} + \frac{\alpha}{2}$ . This choice inflates the value of  $\beta$  relative to  $\alpha$ , takes the correct value as  $\alpha$  approaches one and only allows choosing  $q \geq p_{rn}^a$ . Under relatively mild conditions (see, e.g., Shapiro 2003), we know  $b(z_n^*) = O(n^{-p})$  for  $p \geq 1/2$ , and hence we include the max operator in (9). We also note that  $s_p$  already includes the “ $r^{-1/2}$ ” since  $\hat{\Sigma}$  is the sample covariance of a vector of sample means. For more details on adaptive jackknife estimators, see Partani (2006).

#### 4 NUMERICAL RESULTS

In this section we compare the behavior of our estimators on the asset allocation model (1) with  $\gamma = 0.2$  and  $c = 0.002$ . We construct a portfolio out of 14 Exchange-Traded Funds (ETFs). These funds are designed to track certain market indices. Table 1 indicates the selected funds, the indices they track, the annualized means and standard deviations of return, and the corresponding information ratios (annual mean return divided by annual standard deviation) for the five years of monthly data we used from 1999 to 2004.

We modeled the 14-dimensional monthly return vector using a multivariate normal distribution. This assumption allows us to solve model (1) exactly, and therefore simulation is not needed. Our primary purpose for making this assumption is so we can better assess the performance of our estimators. That said, normal distributions appeared representative of our monthly return data. We performed 14 Kolmogorov–Smirnov goodness-of-fit tests (one for each marginal), and 13 of the 14 failed to reject the normal hypothesis at a 0.05-level of significance. The normal hy-

Table 1: 14 Exchange-Traded Funds (ETFs) with the ETF Ticker Symbol, the Index Being Tracked, Annualized Mean (M) and Standard Deviation (SD) of Return, and the Information Ratio (IR)

Ticker	Tracked Index	M(%)	SD(%)	IR
QQQ	NASDAQ	-7.9	40.4	-0.20
SPX	S&P 500	-4.3	16.1	-0.27
OEX	S&P 100	-5.1	17.5	-0.29
MDY	Midcap	6.3	17.9	0.35
RTY	Russell 3000	5.4	21.8	0.25
MSH	Morgan Stanley High Tech	-1.1	40.9	-0.03
TXX	CBOE Technology	0.9	41.2	0.02
IIX	AMEX Internet	-11.7	51.7	-0.23
BTK	AMEX Biotechnology	19.3	52.1	0.37
GOX	CBOE Gold	5.5	39.2	0.14
XAU	Gold & Silver	0.8	36.0	0.02
OSX	Oil	3.0	34.6	0.09
BKX	Bank Sector	-0.8	23.4	-0.04
DRG	Pharmaceutical	-4.1	17.5	-0.23

pothesis is more plausible for index-type funds than for stocks because the funds track indices that average many underlying stocks. As a result, heavy tails that can arise for individual stocks are less pronounced in such funds. Table 2 contains the correlation coefficients of the monthly returns. We note that normally-distributed returns are inconsistent with our earlier assumption that  $\xi \geq 0$ , i.e., with the normal distribution we can lose more money than we invest. However, the mean returns are large enough relative to the standard deviations that this probability is negligible (particularly for the funds selected during the course of the optimization algorithm).

Our underlying uniform random variates were generated by the combined multiple recursive generator of L’Ecuyer (1999). We generated scalar normal variates using the polar method and formed the vector normal variates using the standard approach with the Cholesky factor of the estimated covariance matrix. We solved instances of (2) calling the Ipopt nonlinear programming solver (Wächter and Biegler 2006). The computations were carried out on a 2.8 GHz Dell Xeon dual-processor computer with 1 GB of memory, although we only use one of the processors.

Table 3 shows  $x^*$  and the results of solving  $N$  i.i.d. replications of (2) with samples sizes ranging from  $n = 25$  to  $n = 800$ . (We used  $N = 8000$  for  $n = 25$ ,  $N = 4000$  for  $n = 50$  and  $N = 2000$  for  $n = 100, \dots, 800$ , and for Table 3 only we used  $c = 0$ .) The solution  $\bar{x}_n^*$  is the average of the optimal solutions  $x_n^*$  across the  $N$  replications. MSE indicates the empirical mean-square error of  $x_n^*$ , i.e.,  $\|x_n^* - x^*\|_2^2$  averaged over the  $N$  replications. The highest allocation in  $x^*$  is in the AMEX Biotech index, which had the best risk/return performance based on the historical returns. In particular, its information ratio (IR) is 0.37, as shown in Table 1. For comparison, the S&P 500 index lost money over our 5-year period. The next fund

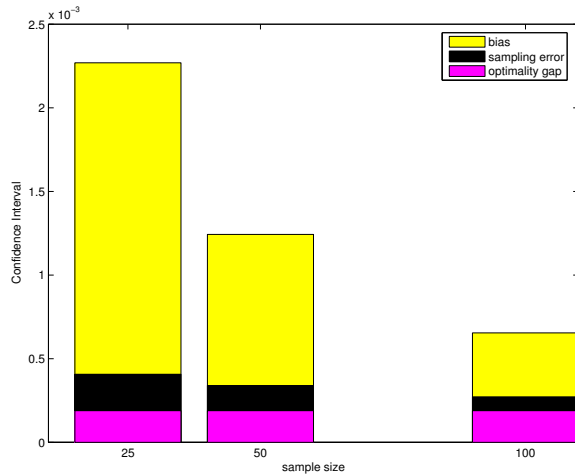


Figure 1: CI Width Versus Sample Size for the MRP<sup>o</sup>

with positive allocation is the CBOE Gold index, with an IR of 0.14. The final positive weight is the Midcap index with an IR of 0.35. The solution  $x^*$  makes sense from a diversification perspective since the correlation between the Biotech and Gold indices is almost zero and the correlation between the Biotech and the Midcap indices is positive. The highest allocation goes to the top funds, in terms of reward for the risk taken, but the allocation to the Gold index reduces the total risk of the portfolio. Assets QQQ, SPX and OEX always had zero allocation and are not listed in the table. The values of  $\bar{x}_n^*$  for small  $n$  might suggest the associated solutions  $x_n^*$  are more diverse than those for larger  $n$  and  $x^*$ . This is not the case, as shown by the average number of nonzero (with a zero tolerance of  $10^{-4}$ ) allocations in  $x_n^*$ , listed as NZ. As the sample size  $n$  grows, the solutions appear to be converging to  $x^*$ . The optimal solution  $x^*$  of (1) is unique, but Table 3 suggests that the set of near-optimal solutions is of significant size. Such solutions are obtained when we optimize the sample mean in place of the population mean. This is precisely the type of situation that leads to  $z_n^*$  having significant bias.

Figure 1 shows the confidence interval on  $\mu_{\hat{x}}$  generated by MRP<sup>o</sup> using  $r = 40$ ,  $\alpha = 0.95$ , and varying  $n$  from 25 to 50 to 100. We obtained  $\hat{x}$  by solving an instance of (2) with  $n = 400$ , and we obtained the existing investor’s portfolio,  $x^t$ , by solving a separate instance with  $n = 400$ . The figure is based on averaging the output of MRP<sup>o</sup> over  $N = 2000$  runs. The CI width is partitioned into the three factors discussed in Section 2.2, namely the optimality gap, the sampling error and the bias. Here, the bias estimate is formed by subtracting the known optimality gap from the average of the  $N = 2000$  point estimates  $\bar{G}_{rn}^o$ . We note that  $z^* = 1.0015$ . So, the  $\hat{x}$  we are using is suboptimal by about 0.02%, and with  $n = 100$  we are forming a 0.95-level CI on that optimality gap whose width is roughly 0.1% of  $z^*$ . We

can clearly see from the figure that bias dominates the CI width. This motivates use of the bias reduction techniques we have proposed.

We assess the performance of three optimality-gap point and interval estimators, denoted as follows: (i)  $D_o$ , the point estimate  $\bar{G}_{rn}^o$  and interval estimate of the MRP<sup>o</sup> in Section 2.2 in which we do not attempt to reduce bias; (ii)  $J_1$ , the standard jackknife estimator, i.e.,  $\bar{G}_{rn}^q$  and MRP<sup>q</sup> of Section 3.1 with  $q = 1$ ; and, (iii)  $D_\beta$ , the adaptive generalized jackknife estimator in which we choose  $q$  via (9) and use this in  $\bar{G}_{rn}^q$  and MRP<sup>q</sup>. For  $D_\beta$  we consider  $\beta$  ranging from 0.70 up to 0.99. Throughout we use  $n = 100$  and  $r = 40$ .

We begin by forming an empirical estimate of the mean-square error (MSE) of each estimator. We did so using  $N = 2000$  i.i.d. runs of each procedure (i)-(iii) above. The estimated MSE of  $D_o$  and  $J_1$  were  $4.5 \times 10^{-7}$ , and  $3.5 \times 10^{-8}$ , respectively. Figure 2(a) shows the MSE of  $D_\beta$  for various values of  $\beta$ , and also includes those of  $D_o$  and  $J_1$  for reference. Figures 2(b)-(d) show the negative and positive part of MSE and the probability the gap point estimate is below  $\mu_{\hat{x}}$ . Because of the nature of our point and interval estimators, we prefer estimators in which MSE<sup>-</sup> and this probability are small. Restated, in choosing between two estimators, we may prefer an estimator with slightly larger MSE if these other two measures are smaller. The first observation is that all our jackknife estimators significantly reduce MSE. The standard jackknife estimator performs very well with respect to MSE. This is not surprising considering the estimates of  $p$  we obtained via  $p_{rn}^a$  over the  $N = 2000$  replications were 0.80 with a standard error of 0.30. So,  $q = 1$  is arguably a reasonably conservative choice. That said, Figures 2(b)-(d) suggest that as  $\beta$  grows the MSE<sup>-</sup> and probability of having an invalid upper bound point estimate improve significantly while the relative increase in MSE<sup>+</sup> is modest.

Figure 3 shows the empirical coverage function of the interval estimators produced by our procedures (Schruben 1980). The original procedure, i.e., without bias reduction, produces an interval estimator with 100% coverage regardless of the value of  $\alpha$ . (Of course, as  $\alpha$  shrinks to zero this no longer holds but the smallest  $\alpha$  in the plot is 0.05.) Using  $\beta = 0.80$  yields an adaptive jackknife estimator that has undercoverage for large values of  $\alpha$ . The interval estimator of the standard jackknife and those associated with larger values of  $\beta = 0.90, 0.95$  and  $D_\alpha$  (which is based on the  $\beta = \frac{1}{2} + \frac{\alpha}{2}$  formula discussed at the end of Section 3.2) all appear to perform well with respect to coverage.

## 5 CONCLUSIONS

We have applied jackknife estimators to assess the quality of a candidate solution, measured by its optimality gap, for a static asset allocation model. And, we have developed a

Table 2: Correlation Coefficients for the Monthly ETF Returns

	QQQ	SPX	OEX	MDY	RTY	MSH	TXX	IIX	BTK	GOX	XAU	OSX	BKX	DRG
QQQ	1.00													
SPX	0.82	1.00												
OEX	0.84	0.99	1.00											
MDY	0.77	0.85	0.80	1.00										
RTY	0.80	0.68	0.63	0.85	1.00									
MSH	0.97	0.83	0.84	0.79	0.82	1.00								
TXX	0.97	0.82	0.83	0.75	0.80	0.98	1.00							
IIX	0.97	0.79	0.81	0.72	0.77	0.95	0.95	1.00						
BTK	0.69	0.37	0.38	0.55	0.74	0.65	0.64	0.64	1.00					
GOX	0.15	0.15	0.13	0.25	0.21	0.15	0.17	0.20	0.02	1.00				
XAU	0.09	0.13	0.10	0.24	0.19	0.10	0.12	0.13	0.03	0.96	1.00			
OSX	0.45	0.55	0.51	0.66	0.53	0.45	0.43	0.39	0.42	0.08	0.14	1.00		
BKX	0.19	0.61	0.57	0.50	0.15	0.24	0.19	0.18	-0.22	0.05	0.04	0.36	1.00	
DRG	0.05	0.41	0.43	0.19	-0.10	0.04	0.00	0.05	-0.14	-0.05	0.01	0.20	0.47	1.00

Table 3: Solutions  $x_n^*$  to (2) Averaged Over Multiple Replications

$n$	Average allocation $\bar{x}_n^*$											MSE	NZ
	MDY	RTY	MSH	TXX	IIX	BTK	GOX	XAU	OSX	BKX	DRG		
25	0.024	0.024	0.018	0.044	0.017	0.362	0.191	0.043	0.107	0.106	0.062	0.822	1.43
50	0.041	0.024	0.011	0.041	0.004	0.419	0.198	0.026	0.095	0.095	0.046	0.742	1.59
100	0.065	0.023	0.004	0.025	0.001	0.484	0.199	0.013	0.077	0.082	0.027	0.643	1.80
200	0.026	0.003	0.000	0.002	0.000	0.136	0.050	0.001	0.012	0.015	0.004	0.544	1.95
400	0.066	0.002	0.000	0.001	0.000	0.306	0.091	0.000	0.015	0.015	0.004	0.436	2.09
800	0.161	0.001	0.000	0.000	0.000	0.645	0.168	0.000	0.010	0.014	0.001	0.361	2.21
$x^*$	0.079	0	0	0	0	0.732	0.189	0	0	0	0	0	3

family of adaptive jackknife estimators, parameterized by  $\beta$ , that can be used to estimate the order of the bias when it is unknown. We considered a simple asset allocation model under normally-distributed returns so that we could solve it exactly and compute the true optimality gap to better assess the performance of our estimators. In our simplest procedure, we do not attempt to reduce bias, and in this case the bias dominates the width of our confidence intervals. All of the jackknife estimators we consider significantly decrease mean-square error by reducing bias. When one seeks a conservative point estimate for use in a one-sided confidence interval our adaptive jackknife with  $\beta = 0.90$ - $0.95$  provides significant improvement over neglecting bias entirely and may provide an attractive alternative to the standard jackknife.

**ACKNOWLEDGMENTS**

This research was partially supported by the National Science Foundation under Grant DMI-0217927.

**REFERENCES**

Andersen, L., and M. Broadie. 2004. A primal-dual simulation algorithm for pricing multi-dimensional American options. *Management Science* 50:1222–1234.

Attouch, H., and R.-B. Wets. 1990. Epigraphical processes: Laws of large numbers for random lsc functions. In *Séminaire d'Analyse Convexe*, Volume 20, 13.1–13.29. Department Des Sciences Mathématiques, Université Montpellier II.

Bayraksan, G., and D. Morton. 2006. Assessing solution quality in stochastic programs. *Mathematical Programming*. To appear.

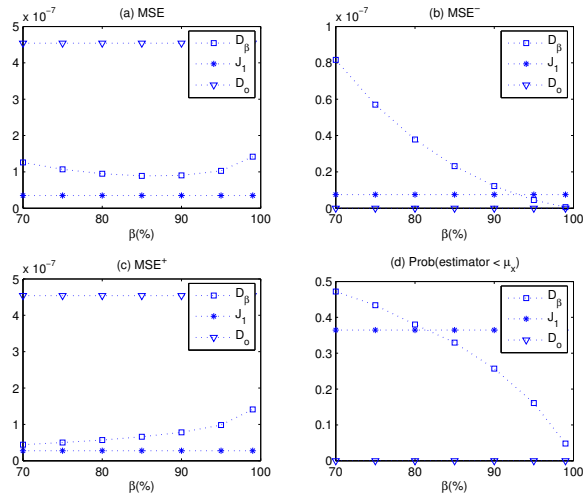


Figure 2: Empirical Mean-Square Error (MSE), Its Negative and Positive Parts, and the Probability the Point Estimate Is Below  $\mu_{\hat{x}}$

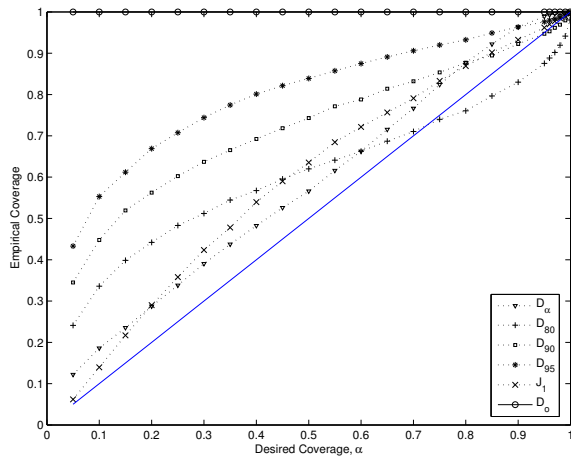


Figure 3: Coverage Function Plots Are Shown for the Interval Estimator without Bias Correction, the Standard Jackknife, the Adaptive Jackknife for  $\beta = 0.80, 0.90$  and  $0.95$ , and  $D_\alpha$ , Which Adjusts the Value of  $\beta$  According to  $\beta = \frac{1}{2} + \frac{\alpha}{2}$  in the Adaptive Estimator

Bayraksan, G., D. Morton, and A. Partani. 2006. Simulation-based optimality tests for stochastic programs. In *Planning Under Uncertainty: Stochastic Programming*, ed. G. Infanger and G. Dantzig. Kluwer Series on Advances in Mathematical Programming. Forthcoming.

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

Fu, M., S. Laprise, D. Madan, Y. Su, and R. Wu. 2001. Pricing American options: A comparison of Monte Carlo simulation approaches. *Journal of Computational Finance* 4:39–88.

Glasserman, P. 2003. *Monte Carlo methods in financial engineering*. Springer, New York.

Gray, H., and W. Schucany. 1972. *The generalized jackknife statistic*. Marcel Dekker, Inc., New York.

King, A., and D. Jensen. 1992. Linear-quadratic efficient frontiers for portfolio optimization. *Applied Stochastic Models and Data Analysis* 8:195–207.

King, A., and R. Rockafellar. 1993. Asymptotic theory for solutions in statistical estimation and stochastic programming. *Mathematics of Operations Research* 18:148–162.

L’Ecuyer, P. 1999. Good parameter sets for combined multiple recursive random number generators. *Operations Research* 47:159–164.

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

Morton, D., E. Popova, and I. Popova. 2006. Efficient fund of hedge funds construction under downside risk measures. *Journal of Banking and Finance* 30:503–518.

Morton, D., E. Popova, I. Popova, and M. Zhong. 2003. Optimizing benchmark-based utility functions. *Bulletin of the Czech Econometric Society* 10:1–18.

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

Partani, A. *Simulation-based procedures for efficient assessment of solution quality in stochastic programming*. Ph. D. thesis, The University of Texas at Austin. In progress.

Quenouille, M. 1949a. Approximate tests of correlation in time-series. *Journal of the Royal Statistical Society, Series B* 11:68–84.

Quenouille, M. 1949b. Problems in plane sampling. *Annals of Mathematical Statistics* 20:355–375.

Schruben, L. 1980. Coverage function of interval estimators of simulation response. *Management Science* 26:18–27.

Shapiro, A. 2003. Monte Carlo sampling methods. In *Stochastic Programming, Handbooks in Operations Research and Management Science*, ed. A. Ruszczyński and A. Shapiro. Elsevier, Amsterdam.

Wächter, A., and L. Biegler. 2006. On the implementation of a primal-dual interior point filter line search algorithm for large-scale nonlinear programming. *Mathematical Programming* 106:25–57.

Ziemba, W., and J. Mulvey. (Eds.) 1998. *Worldwide asset and liability modeling*. Cambridge University Press, Cambridge, United Kingdom.



## **AUTHOR BIOGRAPHIES**

**AMIT PARTANI** is a PhD student in the Graduate Program in Operations Research in the Mechanical Engineering Department at The University of Texas at Austin. His research interests include simulation-based approximations in stochastic programming and mixed integer optimization. His email address is <partani@mail.utexas.edu>.

**DAVID P. MORTON** is an Associate Professor in the Graduate Program in Operations Research in the Mechanical Engineering Department at The University of Texas at Austin. His research interests include computational stochastic programming, including simulation-based approximations in stochastic programming. His email address is <morton@mail.utexas.edu>, and his web page is <<http://www.me.utexas.edu/~orie/Morton.html>>.

**IVILINA POPOVA** is an Assistant Professor of Finance at Seattle University. She taught at Purdue University for 5 years prior to joining Albers School of Business and Economics. Before that she worked for Deutsche Asset Management and Koch Capital Markets. She is a member of INFORMS, ASA, AFA, SFA and WFA. Her email address is <popovai@seattleu.edu>, and her web page is <<http://fac-staff.seattleu.edu/popovai/web/>>.