NESTED SIMULATION FOR ESTIMATING PORTFOLIO LOSSES WITHIN A TIME HORIZON

Sandeep Juneja

L. Ramprasath

School of Technology and Computer Science
Tata Institute of Fundamental Research
Mumbai, 400005, INDIA

Institute for Financial Management and Research
24, Kothari Road, Nungambakkam
Chennai, 600034, INDIA

ABSTRACT

We consider the problem of estimating the probability that a stochastic process observed at discrete time intervals exceeds a specified threshold. We further assume that the value of this process at any time, along any realization, is a conditional expectation which is not known analytically but can be estimated via simulation. This leads to a nested simulation procedure. One application of this arises in risk management where our interest may be in the probability that a portfolio exceeds a threshold of losses at specified times. Here, if the portfolio consists of sophisticated derivatives, then as a function of the underlying security prices, the portfolio value at any time is a conditional expectation that may be evaluated via simulation. In our analysis, we note that conditional on the outer loop of simulation, our estimation problem is related to the large deviations based ordinal optimization framework, so that similar analysis may be used to efficiently allocate computational budget in portfolio evaluations at different times. We also propose a resource allocation methodology based on statistical hypothesis testing.

1 INTRODUCTION

In risk management applications in finance, tail risk measures such as value-at-risk (VaR) and conditional value-at-risk (cVaR) are extremely important performance evaluation measures. These depend upon and are closely related to the probability of large losses in a portfolio. In particular, the techniques that efficiently compute the latter provide direct insights into techniques that efficiently compute VaR or cVaR (Lan, Nelson, and Staum 2007; Gordy and Juneja 2009). In this article, we focus on estimating this probability of large losses in a portfolio of risk instruments.

The analysis in the existing literature typically focuses on evaluating the distribution of portfolio losses at a single time instant, for instance, one month or one year in future. Here we consider the problem of evaluating the probability that the maximum of the loss over multiple time periods exceeds a specified threshold. This is more relevant in many settings. For instance, a large bank may like to hold sufficient reserves so that with high probability it is protected against large losses within a specified time horizon, say one year. The losses may be compiled and computed periodically, say on a weekly or a monthly basis. The probability of large losses over multiple time horizons may also be thought of as a discrete approximation to the probability that the maximum loss within a given time horizon takes large values.

We consider a portfolio that includes sophisticated derivative instruments, whose price is a conditional expectation that depends on the value of the underlying securities. The loss from a portfolio at any time can therefore also be represented as a conditional expectation. Mathematically, our problem can be represented as follows: Suppose $\mathbf{Y} = (Y_i : i \leq K)$ denotes the process of underlying securities at times $(t_i : i \leq K)$. The loss amount at time interval t_i is a conditional expectation

$$L_i = E[X_i | \mathscr{F}_i],$$

where $\mathscr{F}_i$ denotes the information or the σ algebra associated with random variables $(Y_1, \ldots, Y_i)$ and the expectation is typically under the risk-neutral measure. We consider the problem of estimating

$$P\left(\max_{i \leq K} L_i > u\right).$$

Here, each L_i is not known in a closed form, but samples of X_i can be generated. To keep the analysis simple, in this article for the most part we assume that conditional on $\mathbf{Y}$, each X_i has a Gaussian distribution. This may be approximately achieved by batching of samples when necessary. Our large deviations analysis (discussed below) allows general distributions. The evaluation of the portfolio value in our framework is a nested simulation. In the outer loop the sample paths of the underlying securities $\mathbf{Y}$ are generated. Then, in the inner loop each portfolio value L_i is estimated via simulation.

One issue of interest is the division of computational budget between the outer loop and the inner loop. This issue has been addressed in a related settings by Lee (1998) and by Gordy and Juneja (2009). Another issue of interest is the allocation of computational budget to estimate different L_i in each inner loop. Clearly, it makes sense to allocate more resources to time periods i, where estimate of L_i is close in a probabilistic sense to the loss threshold u. We address this in a general framework by modeling it as a problem of ascertaining whether there exists a population in a finite collection, whose mean exceeds a given threshold. We assume that while the means from each population are unknown, we can generate independent samples from each population via simulation.

This problem is addressed using two separate approaches. One approach is based on large deviations based ordinal optimization ideas developed by Glynn and Juneja (2004) (also see Chen, He and Fu (2006)). Using these ideas, we come up with an asymptotically optimal budget allocation that maximizes the probability of making a correct decision via simulation, when the distributional information of each population is known.

Szechtman and Yucesan (2008) tackle a related problem also using large deviations based ideas. They identify for each population in a finite collection whether its mean exceeds a given threshold. Our problem differs somewhat from theirs and our results differ in that we observe degenerate asymptotically optimal allocations. Specifically, we see that in certain settings it is asymptotically optimal to allocate all computer budget to a single population. It is noteworthy that in practice the distribution information of each population has to be estimated from generated samples leading to certain performance degradation due to estimation errors.

We also introduce another hypothesis testing based heuristic for efficiently allocating computational budget to different populations to control the probability of making a false decision. In this setting we assume that each population sample has a Gaussian distribution. This of course is almost always true if batching of samples is allowed. The idea here is to continue to allocate computational effort to those time periods, where there is a statistically significant chance of making an erroneous conclusion. We test the efficacy of the proposed procedures on a small numerical example.

The outline of this article is as follows: Section 2 evaluates the bias and the mean squared error of the nested simulation. The large deviations based allocation of the samples in the inner loop is studied in section 3. Section 4 proposes an alternative sequential algorithm for the inner loop, using the hypothesis testing framework from Statistics and finally section 5 discusses the results of a simulation experiment.

2 BUDGET ALLOCATION IN NESTED SIMULATION

In this section we first discuss the bias involved in estimating $P(\max_{i \leq K} L_i > u)$. Since u is not a parameter that is varied in our analysis, without loss of generality, we set it to zero and consider $P(\max_{i \leq K} L_i > 0)$, which we denote by α. Then we discuss how the computation effort may be allocated between the inner and the outer loop to minimize the resultant mean square error.

2.1 Bias Evaluation

Suppose that each L_i is estimated by the average $\bar{X}_i(m)$ of m samples $\mathbf{X_i}(m) = (X_i(1), \ldots, X_i(m))$. We generate these samples so that conditional on $\mathbf{Y}_i = (Y_j : 1 \leq j \leq i)$, the samples $X_i(j)$ are all mutually independent. Let $\sigma_i^2(\mathbf{Y}_i)$ denote the conditional variance of X_i conditioned on $\mathbf{Y}_i$. In this section, conditioned on $\mathbf{Y}_i$, each X_i is assumed to have a Gaussian distribution with mean $L_i = E(X_i|\mathbf{Y}_i)$ and variance $\sigma_i^2(\mathbf{Y}_i)$.

We estimate α by estimating

$$\hat{\alpha}(m) = P\left(\max_{i \leq K} \bar{X}_i(m) > 0\right).$$

via simulation.

Our first result is that the bias

$$|\hat{\alpha}(m) - \alpha| = K^2 O(m^{-1}). \tag{1}$$

We argue this heuristically giving the flavor of the steps involved without spelling out all the technical assumptions needed. Let

$$A_{ij} = \left\{ L_i > \max_{k \neq i} L_k, \, \bar{X}_j(m) > \max_{k \neq j} \bar{X}_k(m) \right\}.$$

We assume that $(L_i : i \leq K)$ have a joint multi-variate probability density function, so that $\cup_{i,j} A_{i,j}$ is a probability 1 event. Then the bias $|\hat{\alpha}(m) - \alpha|$ equals,

$$\sum_{i,j} \left| E\left[I(A_{i,j}) \left(I(\max_{i \leq K} L_i > 0) - I(\max_{i \leq K} \bar{X}_i(m) > 0) \right) \right] \right|,$$

where $I(B)$ is an indicator of event B.

This is decomposed into two terms

$$\sum_{i=j} \left| E\left[I(A_{i,j}) \left(I(\max_{i \leq K} L_i > 0) - I(\max_{i \leq K} \bar{X}_i(m) > 0) \right) \right] \right| \tag{2}$$

and

$$\sum_{i \neq j} \left| E\left[I(A_{i,j}) \left(I(\max_{i \leq K} L_i > 0) - I(\max_{i \leq K} \bar{X}_i(m) > 0) \right) \right] \right|. \tag{3}$$

The terms in (2) can be re-expressed as

$$\sum_{i \leq K} \left| E\left[I(A_{i,i}) \left(I(L_i > 0) - I(\bar{X}_i(m) > 0) \right) \right] \right|$$

and this is bounded from above by

$$\sum_{i \leq K} \left| E\left[I(L_i > 0) - I(\bar{X}_i(m) > 0) \right] \right|.$$

Let $\tilde{L}_i = \dfrac{L_i}{\sigma_i(\mathbf{Y}_i)}$ and let $f_{\tilde{L}_i}$ denote its pdf and $f_{\tilde{L}_i}^{(1)}$ denote its derivative. Then, from Lee (1998),

$$E\left(I(L_i > 0) - I(\bar{X}_i(m) > 0) \right) = -\frac{f_{\tilde{L}_i}^{(1)}(0)}{2m} + o(m^{-1}).$$

Therefore, (2) is bounded from above by term $KO(m^{-1})$.

Now consider the term

$$\left| E\left[I(A_{i,j}) \left(I(\max_{i \leq K} L_i > 0) - I(\max_{i \leq K} \bar{X}_i(m) > 0) \right) \right] \right|$$

for $i \neq j$. We will show that this term is bounded from above by an $O(m^{-1})$ term. First rewrite the above term as,

$$\left| E\left[I(A_{i,j}) \left(I(L_i > 0) - I(\bar{X}_j(m) > 0) \right) \right] \right|.$$

This is bounded from above by

$$EI(A_{i,j})I(L_i > 0, \bar{X}_j(m) < 0) \tag{4}$$

plus

$$EI(A_{i,j})I(L_i < 0, \bar{X}_j(m) > 0). \tag{5}$$

The term (4) can be bounded from above by

$$E\left[I(L_i > 0, L_i > L_j, \bar{X}_i(m) < \bar{X}_j(m) < 0)\right].$$

Similarly, (5) is bounded from above by

$$E\left[I(\bar{X}_j(m) > 0, \bar{X}_j(m) > \bar{X}_i(m), L_j < L_i < 0)\right].$$

We now show that (4) can be bounded from above by an $O(m^{-1})$ term. The fact that (5) is $O(m^{-1})$ can then be similarly argued, and then the bias result (1) follows.

To see that (4) is $O(m^{-1})$, note that it is bounded from above by

$$E\left[I(L_i > 0, \bar{X}_i(m) < \bar{X}_j(m) < 0)\right]. \tag{6}$$

We now argue that this is asymptotically similar to

$$\frac{1}{4m} f_{L_i,L_j}(0,0) E[\sigma_i^2(\mathbf{Y_i})|L_i = 0, L_j = 0] + O(m^{-3/2}),$$

where f_{L_i,L_j} denotes the joint pdf of (L_i, L_j).

We may express $\bar{X}_i(m)$ as $L_i + W_i/\sqrt{m}$, where conditioned on $\mathbf{Y}_i$, W_i has a Gaussian distribution with mean zero and variance $\sigma_i(\mathbf{y}_i)^2$, and its distribution is independent of m. Let $f_{L_i,W_i,\bar{X}_j(m)}$ denote the joint pdf of $(L_i, W_i, \bar{X}_j(m))$. Then, (6) equals

$$\int_{-\infty}^0 \int_0^{-w_i/\sqrt{m}} \int_{l_i+w_i/\sqrt{m}}^0 f_{L_i,W_i,\bar{X}_j(m)}(l_i, w_i, x) dx dl_i dw_i.$$

Under mild conditions, using Taylor series expansion of $f_{L_i,W_i,\bar{X}_j(m)}(l_i, w_i, x)$ at point $f_{L_i,W_i,\bar{X}_j(m)}(0, w_i, 0)$, this can be seen to equal

$$(1 + O(m^{-1/2})) \int_{-\infty}^0 f_{L_i,W_i,\bar{X}_j(m)}(0, w_i, 0) \int_0^{-w_i/m} \int_{l_i+w_i/\sqrt{m}}^0 dx dl_i dw_i.$$

This in turn equals

$$\int_{-\infty}^0 f_{L_i,W_i,\bar{X}_j(m)}(0, w_i, 0) \frac{w_i^2}{2m} dw_i + O(m^{-3/2}). \tag{7}$$

Recall that $\bar{X}_j(m) = L_j + W_j/\sqrt{m}$ and we will now show

$$f_{L_i,W_i,\bar{X}_j(m)}(0, w_i, 0) = f_{L_i,W_i,L_j}(0, w_i, 0) + O(m^{-1}). \tag{8}$$

This is because,

$$f_{L_i,W_i,L_j+W_j/\sqrt{m}}(0, w_i, 0) = \int_{\Re} f_{L_i,W_i,L_j|W_j=w}(0, w_i, -w_j/\sqrt{m}) f_{W_j}(w) dw \tag{9}$$

where $f_{L_i,W_i,L_j|W_j=w}(l_i, w_i, l_j)$ denotes the conditional pdf of (L_i, W_i, L_j) conditioned on $W_j = w$, and f_{W_j} denotes the pdf of W_j. Then, using the Taylor series expansion of the integrand in (9), and noting the $EW_j|Y_j = 0$, under mild conditions, (8) follows.

Then, (7) equals

$$f_{L_i,L_j}(0,0) \frac{1}{2m} E[W_i^2 I(W_i \le 0)|L_i = 0, L_j = 0] + O(m^{-3/2}).$$

This in turn equals

$$f_{L_i,L_j}(0,0)\frac{1}{4m}E[\sigma_i^2(\mathbf{Y_i})|L_i=0,L_j=0]+O(m^{-3/2}),$$

since

$$E[W_i^2 I(W_i \leq 0)|L_i=0,L_j=0] = E[W_i^2 I(W_i \leq 0)|\mathbf{Y_i}|L_i=0,L_j=0] = \frac{1}{2}E[\sigma_i^2(\mathbf{Y_i})|L_i=0,L_j=0].$$

2.2 Minimizing Mean Square Error

Recall that our aim is to estimate $\alpha = P(\max_{i \leq K} L_i > 0)$. We estimate this by estimating $\hat{\alpha}(m) = P(\max_{i \leq K} \bar{X}_i(m) > 0)$ via simulation.

The nested procedure for estimating $\hat{\alpha}(m)$ is as follows:

1. In the outer loop we generate n independent samples of the process $\mathbf{Y}$.
2. For each run of the outer loop, in the inner loop we generate samples of $\bar{X}_i(m)$ for $i \leq K$ and set our output to $I(\max_{i \leq K} \bar{X}_i(m) > 0)$.
3. The simulation output is the average of the independent outputs from each inner loop.

It is clear that the variance of the simulation output equals

$$\frac{1}{n}\hat{\alpha}(m)(1-\hat{\alpha}(m)).$$

In the analysis in the previous subsection we established that for a fixed K, the bias

$$|\hat{\alpha}(m)-\alpha| = O(m^{-1}).$$

A more refined analysis can show that the RHS is asymptotically similar to c/m where c is a constant independent of m. Then, as in Lee (1998) and Gordy and Juneja (2008), we can look for optimal n and m that minimize the mean square error of the resulting simulation estimator.

This mean square error equals the variance plus the squared bias, that is,

$$\frac{1}{n}\hat{\alpha}(m)(1-\hat{\alpha}(m))+(\hat{\alpha}(m)-\alpha)^2.$$

This in turn equals

$$\frac{1}{n}\alpha(1-\alpha)+c^2/m^2+O(m^{-1})/n+o(m^{-2}).$$

If we assume that the average computational effort in the outer loop is a constant γ and the average computational effort in the inner loop is $m\beta$, then for large values of n, the overall computational effort is close to $n(\gamma+m\beta)$. A reasonable optimization problem is to look for m and n that minimize the mean square error assuming that that the total computational budget $\Gamma = n(\gamma+m\beta)$ is large but fixed.

Then, it can be seen that optimal number of outer loops is

$$n^* = \left(\frac{\alpha(1-\alpha)}{2\beta^2 c^2}\right)^{1/3}\Gamma^{2/3}+o(\Gamma^{2/3}),$$

and optimal size for the inner loop is

$$m^* = \left(\frac{2c^2}{\alpha(1-\alpha)\beta}\right)^{1/3}\Gamma^{1/3}+o(\Gamma^{1/3}).$$

3 LARGE DEVIATIONS BASED ALLOCATION

In this section, we focus on the computational effort allocation problem in the inner loop of the nested simulation, where the allocation is made to minimize the probability of making a false identification. The problem in this section is analyzed more generally. We consider the setting where there are K populations. Each population is characterized by its probability distribution which is not explicitly known, but from which i.i.d. samples can be generated. These samples are mutually independent across populations. Our aim is to correctly identify if any of the populations has a mean greater than zero.

We first conduct the analysis assuming that each population has a general distribution, using the large deviations framework. We determine the allocations that minimize the asymptotic probability of false identification, under the assumption that the distributions associated with each population are known. This corresponds to maximizing the large deviations rate function associated with the probability of false identification. We then briefly discuss issues related to the fact that population distribution is not known but is estimated from the generated samples.

3.1 Mathematical Framework and Analysis

Let $F_i(\cdot)$ denote the distribution function of population i, μ_i the associated mean, $\Lambda_i(\cdot)$ the log-moment generating function and $\Lambda_i^*(\cdot)$ its Fenchel-Legendre transform, i.e.,

$$\Lambda_i^*(x) = \sup_\theta (\theta x - \Lambda_i(\theta)).$$

We suppose that total m simulation samples are available to us to allocate to different populations and we want to allocate to minimize the probability of false identification. Specifically, suppose that we assign $p_i m$ samples to population i (assume that $p_i m$ is an integer to avoid undue technicalities), $p_i > 0$, $\sum_{i=1}^K p_i = 1$. We look to identify the set of $\mathbf{p} = (p_1, \ldots, p_K)$ that minimize the asymptotic probability of false identification. Let $(X_{i1}, X_{i2}, \ldots)$ denote i.i.d. samples from population i and let $\bar{X}_i(n) = \frac{1}{n} \sum_{i=1}^n X_{ik}$ denote the associated sample mean. Two separate cases need to be considered.

3.1.1 All μ_i are Negative

In this setting, the probability of false identification equals

$$P(\max_{i \leq K} \bar{X}_i(p_i m) > 0).$$

We need to identify its large deviations rate function. To this end, using Boole's inequality we have,

$$\max_{i \leq K} P(\bar{X}_i(p_i m) > 0) \leq P(\max_{i \leq K} \bar{X}_i(p_i m) > 0) \leq K \max_{i \leq K} P(\bar{X}_i(p_i m) > 0). \tag{10}$$

From Cramer's Theorem (see for e.g. Dembo and Zeitouni (1998)), it follows that

$$\lim_{m \to \infty} \frac{1}{m} \log P(\bar{X}_i(p_i m) > 0) = -p_i \Lambda_i^*(0).$$

From this and (10), it follows that

$$\lim_{m \to \infty} \frac{1}{m} \log P(\max_{i \leq K} \bar{X}_i(p_i m) > 0) = -\min_{i \leq K} p_i \Lambda_i^*(0).$$

Thus, asymptotically our problem becomes

$$\max_{p_i > 0, \sum_{i=1}^K p_i = 1} \min_{i \leq K} p_i \Lambda_i^*(0).$$

The solution to this is,

$$p_i^* = \frac{1/\Lambda_i^*(0)}{\sum_{j=1}^K 1/\Lambda_j^*(0)} \quad \text{for } i = 1, \ldots, K,.$$

Therefore the populations that are most likely to cross the zero threshold are given more computational budget. When each population is Gaussian distributed with mean μ_i and variance σ_i^2, then

$$\Lambda_i^*(x) = \frac{(x-\mu_i)^2}{\sigma_i^2} \qquad \text{and} \qquad p_i^* = \frac{\sigma_i^2/\mu_i^2}{\sum_{j=1}^K \sigma_j^2/\mu_j^2}.$$

3.1.2 Some μ_i are Positive

In this case the probability of false identification corresponds to

$$P(\max_{i\leq K} \bar{X}_i(p_im) \leq 0) = \prod_{i=1}^K P(\bar{X}_i(p_im) \leq 0).$$

In general

$$\lim_{m\to\infty} \frac{1}{m} \log P(\bar{X}_i(p_im) \leq 0) = -p_i\Lambda_i^*(0),$$

if $\mu_i > 0$ and

$$\lim_{m\to\infty} \frac{1}{m} \log P(\bar{X}_i(p_im) \leq 0), = 0$$

if $\mu_i \leq 0$. Let B denote the set of populations for which $\mu_i > 0$. Therefore,

$$\lim_{m\to\infty} \frac{1}{m} \log P(\max_{i\leq K} \bar{X}_i(p_im) \leq 0) = -\sum_{i\in B} p_i\Lambda_i^*(0).$$

In particular, if in the set B there is a unique i with the largest value of $\Lambda_i^*(0)$, then the asymptotically optimal allocation corresponds to giving full computational budget to that population. In the Gaussian settings, the population with $\mu_i > 0$ and the highest μ_i^2/σ_i^2 gets all the computational budget.

The bias result from section 2 can be extended to this optimal allocation design and one can show that the same convergence rate for the mean squared error is applicable here.

3.1.3 Implementation Issues

In practice, the distribution and the log-moment generating function of each population is not known and is estimated from data. Thus, the allocation derived above is subject to estimation errors. One may consider an iterative procedure, where there is a pilot phase where each population gets equal number of samples. From this, the associated distributions are estimated and optimal allocations ascertained. In the next phase, a small fraction of the budget is allocated to each population even if that is not suggested by the optimal allocation (to compensate for errors in distribution estimation) but apart from this correction, the allocations are based on the optimal calculations in the previous phase. This process is repeated till the computational budget gets exhausted.

It is easy to show that as long as every population gets sufficient computation budget, the resulting allocations, using empirical log moment generating functions as proxies for the true log moment generating functions in our calculations, converge to the optimal allocations (this is shown in a related setting by Glynn and Juneja (2004)). Szechtman and Yucesan (2008) develop a provably convergent stochastic approximation algorithm that can also be adapted to our settings.

It is interesting to note in our settings that the allocations derived form the above scheme appear to be directionally correct even when the estimate of the underlying distribution may be wrong. This is easy to see in the Gaussian setting. Consider the case where after running simulations for each population, mean and variance for population i are estimated as $\hat{\mu}_i$ and $\hat{\sigma}_i^2$. Further suppose that all $\hat{\mu}_i$ are negative. Then, based on the proposed scheme, greater budget is given to population i with larger $\hat{\sigma}_i^2/\hat{\mu}_i^2$ as any reasonable scheme should, to correctly resolve whether this true population mean is below 0 or above it.

Similarly, if any of the $\hat{\mu}_i$ is positive, then it is reasonable to give most of the computational budget in a given phase to a population with positive sample mean and highest value of $\hat{\mu}_i^2/\hat{\sigma}_i^2$, as again it is important to resolve quickly whether the true mean of this population is positive.

4 HYPOTHESIS TESTING BASED ALLOCATION SCHEME

In this section we propose a hypothesis testing based algorithm for making a correct identification of existence of a population with mean greater than 0. Here we assume that each population has a Gaussian distribution. As is well known, this can be approximately achieved by sample batching.

The essential idea is to fix a bias or the error level ε that we are willing to tolerate in our identification procedure. Then suppose that initially a pilot run is conducted where equal number m_0 samples are allocated to each population. Let $\hat{\mu}_i$ and $\hat{\sigma}_i^2$ denote the estimated mean and variance for population i. Two cases can occur:

All $\hat{\mu}_i \leq 0$: In this case suppose as a null hypothesis that there exists a design with true mean greater than zero. We test to see if this can be rejected at significance level ε. Conservatively, we test the null hypothesis that each mean $\mu_i = 0$ versus the alternate hypothesis that $\mu_i < 0$ for each i at level ε/K. That is, for each population i we test whether

$$\frac{\hat{\mu}_i}{\hat{\sigma}_i} < \frac{1}{\sqrt{m_0}} t_{m_0-1,\varepsilon/K}, \tag{11}$$

where, $t_{m_0-1,\varepsilon/K}$ satisfies the equation

$$P(T < t_{m_0-1,\varepsilon/K}) = \varepsilon/K$$

and T is a rv with t-distribution with $m_0 - 1$ degrees of freedom. All i for which inequality (11) does not hold are allocated more samples. The new sample amount $m_1 - m_0$ is chosen for such a population i so that

$$\frac{\hat{\mu}_i}{\hat{\sigma}_i} = \frac{1}{\sqrt{m_1}} t_{m_1-1,\varepsilon/K}.$$

In particular, m_1 is roughly proportional to $\left(\frac{\hat{\mu}_i}{\hat{\sigma}_i}\right)^2$ as in the large deviations based algorithm. On the other hand, if all populations satisfy (11), then we stop with the deduction that no population mean exceeds 0.

There exists non-empty set B: $\hat{\mu}_i > 0$, for $i \in B$:

Here we consider the null hypothesis that all $\mu_i \leq 0$. Again, conservatively we test the null hypothesis that $\mu_i = 0$ for all i, with the alternate hypothesis that there exists a population with a positive mean. If there exist a population $i \in B$ so that

$$\frac{\hat{\mu}_i}{\hat{\sigma}_i} > \frac{1}{\sqrt{m_0}} t_{m_0-1,\varepsilon}, \tag{12}$$

then we can reject the null hypothesis at the ε significance level and the simulation stops with the decision that there exists a population with mean exceeding zero. If not then we assign additional samples $m_1 - m_0$ to the population $i \in B$ with the largest $\frac{\hat{\mu}_i}{\hat{\sigma}_i}$ so that

$$\frac{\hat{\mu}_i}{\hat{\sigma}_i} = \frac{1}{\sqrt{m_1}} t_{m_1-1,\varepsilon}.$$

Note again the similarity to the large deviations based algorithm.

After the additional samples are allocated, then the estimated sample means and variances can be updated and iteratively this procedure can be repeated until one of the two stopping rules is met. In the case where some of the true means are very close to zero, we may observe that this procedure takes a long time to terminate. in that case overall limit on the computational budget may be imposed so that once that limit is met, the decision is made based on current sample averages. Thus, if there exists a sample average that exceeds zero, our output for this experiment is that the true mean is greater than

zero, else we decide otherwise. A more detailed analysis of the proposed scheme will be provided in an expanded version of this paper.

5   NUMERICAL RESULTS

We now illustrate the need for a better algorithm than the naive equal allocation scheme through a simple numerical example, taken from Lan and Nelson and Staum (2007). Consider the following situation: At time 0, we sell a put option with strike $110 and maturity 3 months on a stock, whose initial price is $100. We assume the Black Scholes model for the stock price with volatility = 15% and the risk free rate = 6%. We sell the put option at the Black Scholes price and look at the riskiness of the portfolio, evaluated every week till the expiry of the option. Thus the portfolio is simply the short put option and the proceeds of the sale invested at the risk free interest rate.

Since our goal is to arrive at better allocation strategies compared to the equal allocation scheme, for the inner loop, we make the following numerical comparisons: for a fixed path in the outer loop, the portfolio values are computed at the end of every week using Black Scholes formula, for 3 months (i.e., 12 weeks). This way we arrive at the indicator whether the minimum portfolio value (equivalent of maximum loss) over these 12 time points is above or below a fixed threshold. This indicator value is now estimated using Monte Carlo simulations at each of the 12 time points using simulation to estimate the price of a put instead of using the Black Scholes formula. Finally, the frequency of correct identification using 10,000 trials is compared for equal and optimal allocation schemes.

In particular, we chose an outer path such that only one of the 12 portfolio losses was slightly above the threshold and others were below. So the true indicator value was 1, but equal allocation could correctly identify only 60% of the cases, which leaves a lot of room for improvement. Figure 1 shows the probabilities of correct identification as we increase the computing budget for the inner loop from 480 to 4800. The results show a marked improvement for the optimal allocation scheme over the naive equal allocation scheme.
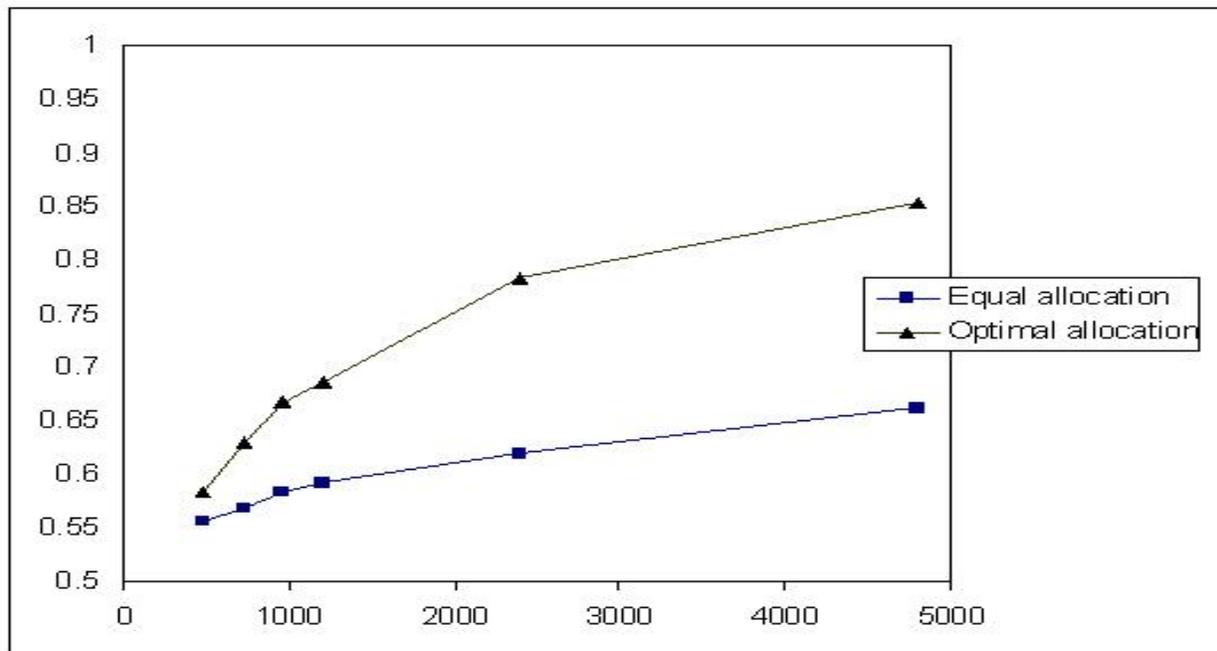

Figure 1: Probability of correct identification (*y* axis) versus computing budget (*x* axis) for equal and optimal allocation procedures

The optimal allocation scheme used here was based on the large deviation heuristics, where a initial sample of 20 was taken at each time point. The optimal allocation of the remaining budget was decided based on this sample using the Gaussian assumption, as detailed in section 3. We haven't implemented the sequential algorithm in this paper and we hope to present it at the conference. These results indicate that there is still room for significant improvements over the naive equal allocation scheme.

REFERENCES

Chen, C. H., D. He, and M. Fu. 2006. Efficient dynamic simulation allocation in ordinal optimization. *IEEE Transactions on Automatic Control* 51 (12): 2005–2009.

Dembo, A., and O. Zeitouni. 1992. *Large deviations techniques and applications.* Boston: Jones and Bartlett.

Glynn, P. W., and S. Juneja. 2004. A large deviations perspective on ordinal optimization. In *Proceedings of the 2004 Winter Simulation Conference*, ed. R. G. Ingalls, M. D. Rossetti, J. S. Smith, and B. A. Peters, 577–585. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Gordy, M., and S. Juneja. 2009. Nested simulation in portfolio risk measurement, forthcoming in *Management Science*.

Lan, H., B. L. Nelson and J. Staum. 2007. A confidence interval for tail conditional expectation via two-level simulation. In *Proceedings of the 2007 Winter Simulation Conference*, ed. S. G. Henderson, B. Biller, M. H. Hsieh, J. Shortle, J. D. Tew, and R. R. Barton, 949–957. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Lee, S. H. 1998. *Monte Carlo expectation of conditional expectation quantiles.* Ph.D. thesis, Department of Operations Research, Stanford University, Palo Alto, California.

Szechtman, R., and E. Yucesan. 2008. A new perspective on feasibility determination. In *Proceedings of the 2008 Winter Simulation Conference*, ed. S. J. Mason, R. R. Hill, L. Monch, O. Rose, T. Jefferson, and J. W. Fowler, 273–280. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

AUTHOR BIOGRAPHIES

SANDEEP JUNEJA is Associate Professor in the School of Technology and Computer Science at Tata Institute of Fundamental Research. He received his Ph.D. from Stanford University. His research interests include applied probability, computational finance, simulation and rare-event analysis. He serves on the editorial boards of Mathematics of Operations Research and ACM TOMACS. His email address is <juneja@tifr.res.in>.

L. RAMPRASATH is Assistant Professor at Institute for Financial Management and Research. He received his Ph.D. in Statistics and Probability from Rutgers, The State University of New Jersey. His research interests include static hedging strategies for exotic options, efficient simulation techniques in financial engineering and statistical inference. His email address is <rprasath@ifmr.ac.in>.