February 23, 2000

Variance Reduction Techniques for Monte Carlo Estimates of Value at Risk

by

Brian Fuglsbjerg

Financial Research Department,
SimCorp A/S

Oslo Plads 12,
DK-2100 Copenhagen O,
Denmark.

E-mail: Brian.Fuglsbjerg@simcorp.com.
Tel: (+45) 35 44 89 67.
Fax: (+45) 35 44 88 11.

Preliminary Version

Comments Welcome

Variance Reduction Techniques for Monte Carlo Estimates of Value at Risk

Abstract: The sampling variance of Monte Carlo estimates of Value at Risk is reduced using Importance Sampling and Control Variates. Both Importance Sampling and Control Variates are based on second-order approximations of the change in portfolio value, and the variance reduction technique yields very high variance reductions when this approximation is reasonable. This involves portfolios where Delta's are large, relative to higher order derivatives. The methods improve the variance reduction relative to the Importance Sampling and Stratified Sampling approach suggested by [Glasserman et al., 1999a].

JEL classification code: C63, G18.

1 Introduction

Value at Risk (VaR) is a popular metric for measuring and managing portfolio risk. VaR is the maximal portfolio loss on a given, fixed portfolio, which can be observed in a given period of time at a prespecified confidence level.

$$P(\Delta V \leq -\text{VaR}) = 1 - \alpha,$$

where $\Delta V = V(0) - V(t)$ is the change in portfolio value over the VaR horizon. Typically the confidence level α is chosen to be 95% or 99%, and the time horizon to be one day or two weeks.

Monte Carlo is basically the only alternative to the traditional, simple methods. The simplest of these methods is the Delta-Normal method. The basic underlying assumptions for this method are that the risk factors, i.e. all sources of risk in the economy are multivariate normal distributed, and that the portfolio is well approximated by a linear function of the risk factors (for an explanation of Delta-Normal VaR in full detail, please consult e.g. [Duffie and Pan, 1997, J.P.Morgan, 1996, Jorion, 1997]) This simple approximation tends to be far too impecise, in particular for portfolios with high option content or at high confidence levels (99% and higher), but also for portfolios where the Delta is small relative to Gamma, such as swap portfolios and Delta hedged portfolios. The linear portfolio assumption can easily be extended to an assumption about quadratic dependence on the risk factors if the assumption of the risk factors multivariate normality is maintained (for an introduction to the mathematics of Delta-Gamma-Normal VaR, please consult e.g. [Rouvinez, 1997]). However, since Delta-Gamma-Normal VaR is based on a local approximation, it is generally imprecise for large moves in the risk factors, and it may therefore lead to large errors in estimated VaR at high confidence levels (99% and above).[1]

The main problem in using simulation for VaR problems is the amount of time it takes to compute accurate estimates - in particular when the portfolio consists of many assets and/or when the confidence level is high. The problem with applying standard simulation techniques to VaR-like settings is that many simulation paths are wasted in the sense that they are concentrated around current portfolio value, rather than being concentrated around the wanted VaR value. Therefore a very large number of simulations are needed to achieve a sufficiently high precision on the estimate[2] and variance reduction methods can potentially reduce sampling variance significantly. In this paper we devise one powerful way of estimating VaR using Monte Carlo.

This paper is very similar to [Cardenas et al., 1999] and [Glasserman et al., 1999a] in the sense that all papers pursue the same strategy, of finding a way to minimize sampling variance on Monte Carlo VaR, so that estimates can be computed in a reasonable amount

[1] One additional simple alternative to Monte Carlo is the so-called Historical Simulation, where VaR is estimated using observed historical portfolio returns (see [Beder, 1995, Jorion, 1997]).

[2] See [Pritsker, 1996] for a discussion of the needed number of simulations to achieve a certain accuracy at a given confidence level.

of time. All papers propose variance reduction methods for simulation of Value at Risk based on the quadratic approximation of the portfolio value. [Cardenas et al., 1999] apply Control Variates (CV) and a very simple Stratification (two strata) to reduce variance. The CV is simply a second-order Taylor approximation of the change in portfolio value, so the local approximation they use becomes less correlated with the change in portfolio value as the confidence level increases. A less correlated control variate has less explanatory power of the change in portfolio value and the result is a lower variance reduction, and the efficiency of this variance reduction scheme is therefore decreasing as the confidence level increases. In Section 2 we discuss this problem further and suggest a simple improvement. [Glasserman et al., 1999a] propose to use a particular Importance Sampling scheme in conjunction with Stratified Sampling to reduce variance. Both Importance Sampling and Stratified Sampling appear to reduce variance significantly, and the form of Importance Sampling (IS) they propose seems to be particularly well-suited to reduce variance for VaR-like problems. The most promising feature of the Importance Sampling as the variance reduction method is that the variance ratios tend to *increase* as the confidence levels increase. This method will also be discussed in more detail later.

The purpose of this paper is to present a different variance reduction technique for VaR evaluations. The basic idea is simple: IS shifts the mean and the covariance of the risk factors, and the variance reduction method we propose benefits from the change of mean. We combine IS with CV based on a second-order Taylor expansion around the IS mean. The result is that we in general increase the correlation between the control variate and the change in portfolio value substantially, relative to the original second-order Taylor expansion, leading to higher variance reductions. We furthermore apply stratified sampling combined with IS, where we stratify on the quadratic approximation around the shifted mean, and again variance is reduced.

The resulting variance reductions are impressive: we compute variance ratios for 7 portfolios of stock options, originally proposed by [Glasserman et al., 1999a]. For these portfolios we are able to reduce variance by at least a factor 28 at the 95% confidence level, a factor 50 at the 99% level, and up to a factor 1400 at 99.5% level. In all cases the variance reductions are at least as high as the reductions that can be achieved by implementing Importance Sampling and Stratification on the quadratic approximation as suggested by [Glasserman et al., 1999a]. The extra variance reduction comes at a cost of one additional computation of the second-order Taylor approximation (Delta's and Gamma's), but the variance ratios are quite high, relative to stratifying on the original Taylor expansion, so the extra computational effort seems worthwhile for these portfolios.

We then consider three Delta hedged options portfolios, where the options are hedged with the underlying security. These portfolios are such that the quadratic approximation is a poor approximation of the change in portfolio value, and it is demonstrated that the high variance ratios for the portfolios proposed by [Glasserman et al., 1999a] are by no means guaranteed; in fact simulation variance might even increase if the approximation is very poor. This result is not surprising, considering all variance reduction methods rely heavily on the quadratic approximation, and that the portfolios are constructed such that this approximation is poor.

In the remaining Sections of the paper we take a further look at the details of the variance reduction techniques we employ. First, in Section 2 we discuss the technical issues in the use of CV and IS and in the combination of the two methods; then, in Section 3 we investigate the numerical performance of the proposed variance reduction methods; finally, in Section 4 we summarize and propose future extensions.

2 Variance reduction techniques

Instead of computing Value at Risk we consider the very closely related problem of calculating the exceedance probability $p(v) = P(\Delta V \leq v) = E(\mathbf{1}_{\{\Delta V \leq v\}})$, where ΔV is the change in portfolio value on the fixed portfolio over the period $[0, t]$, and v is a reference point.[3]

Obviously, standard Monte Carlo – without variance reductions – is a most inefficient way of estimating the tail probability since the major part of the simulation paths does not influence the estimate. Since the estimate with n sample paths is $\frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{\Delta V(x_i) \leq v\}}$, where x_i is a draw from the underlying risk factor distribution, we can compute the variance on the estimator as $\frac{1}{n}(1-p)p$, which is very high compared to the mean value p, especially when p is small, say 1% or less.

Importance sampling (IS) is a standard way of reducing variance for simulations of rare events [Fishman, 1995], and the benefits of IS are getting increasingly bigger as v gets smaller and the wanted exceedance probability $p(v)$ approaches 0. Instead of sampling from the original distribution of the risk factors, the risk factor outcomes are sampled from an alternative distribution, and the estimate is corrected with the likelihood ratio corresponding to this change of measure. Let the density of the original distribution be $f(x)$, and let the new IS distribution have probability density function $f^*(x)$. The Radon-Nikodym derivative corresponding to the change of measure is simply the ratio of the densities $l(x) = \frac{f(x)}{f^*(x)}$ and the exceedance probability is now computed as $E^*(\mathbf{1}_{\{\Delta V \leq v\}} l(x))$. E^* is the expectation corresponding to the density f^*. The MC estimator is $\frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{(-\infty,v)}(\Delta V(x_i)) l(x_i)$, where the indicator function $\mathbf{1}_{(-\infty,v)}(y)$ returns if $y \in (-\infty, v)$ and 0 otherwise, and the x_i's are IID from the IS distribution F^*. Although there is no guarantee that IS reduces variance, the method very often turns out to be extremely efficient because the Radon-Nikodym derivative turns out to be smaller than 1 in the sample area, i.e. where $\Delta V \leq v$. The number of sample paths influencing the estimate is therefore increased, and the intuition is therefore that the sampling variance is reduced.

[3] It should be clear that VaR can be computed through a simple iterative procedure finding successive exceedance probabilities - obviously using the already simulated portfolio values. The problem of estimating is therefore equivalent for all practical purposes, but exceedance probability is merely an expectation of an indicator function, so the problem is better suited for simulation.

The variance of the IS estimate is

$$\begin{aligned}\mathrm{var}(\hat{p}(v)) &= \frac{1}{n}\big(E^*\big[\mathbf{1}_{(-\infty,v)}(\Delta V(x))l(x)^2\big] - p(v)^2\big)\\ &= \frac{1}{n}\big[\int \mathbf{1}_{(-\infty,v)}(\Delta V(x))l(x)f(x)dx - p(v)^2\big].\end{aligned}$$

Choosing sampling density $f^*(x)$ proportional to $\mathbf{1}_{(-\infty,v)}(\Delta V(x))f(x)$ yields an IS estimate with zero variance. This zero variance density unfortunately requires prior knowledge of the set $\{x|\Delta V(x) \leq v\}$ which is unknown in general. It also requires knowledge of the scaling parameter that makes $f^*(x)$ a density. This scaling turns out to be the unknown exceedance probability $p(v)$. Therefore it is impossible to apply the zero variance IS distribution, so one can alternatively restrict the allowed IS distributions to be in a parameterized group of densities, and attempt to minimize the second moment of the estimate within this set of distributions. In Section 2.2 we show in more detail how we use a second-order Taylor approximation of the change in portfolio value to select an IS distribution.

Like IS, Control Variates (CV) may be an efficient way of reducing the sampling variance. Suppose there is a random variable $H(x)$, highly correlated with $\mathbf{1}_{(-\infty,v)}(\Delta V(x))$, which has a known mean value. An unbiased estimate of $p(v)$ is then given by

$$\hat{p}(v) = \frac{1}{n}\big[\sum_{i=1}^{n} \mathbf{1}_{(-\infty,v)}(\Delta V(x_i)) - \alpha\,(H(x_i) - E(H(x_i))\big],$$

and the variance of this estimate is simply

$$\frac{1}{n}(\sigma_V^2 + \alpha^2\sigma_H^2 - 2\alpha\sigma_{VH}),$$

where σ_V is the standard deviation of the portfolio value estimator, σ_H is the standard deviation of the control variate, and σ_{HV} is the covariance between the two estimators. If the covariance and the standard deviation on the control variate are known, it is possible to compute the optimal regression coefficient, α, that minimizes the estimation variance. In general, these are unknown and there are two potential ways to apply CV.

- Perform a regression to obtain the optimal coefficient α
- Set α to 1 and hope that the regression coefficient is close to 1.

In the first case one should be aware of the errors-in-variables problem associated with a simultaneous estimation of α and $p(v)$. If the same sample is used for estimating both variables the estimators are statistically dependent, and the resulting estimate of $p(v)$ is therefore generally biased. Presampling of the regression coefficient is the most obvious way to avoid this bias, but [Avramidis and Wilson, 1990] present a method where the bias is avoided without presampling.[4]

[4]For many practical purposes, where the portfolio is changed only moderately from day to day, one can use the last estimate of the regression coefficient for today's estimate of Value at Risk.

In the latter case, the loss in variance reduction of just setting $\alpha = 1$ is not large if the regression coefficient is close to 1. To exemplify: suppose the variances of the two estimates are similar and that the regression parameter is 0.9. The variance ratio from using the correct regression coefficient is found to be $\frac{1}{(1-\rho^2)} = 5.26$, where ρ is the correlation between the two stochastic variables, whereas assuming $\alpha = 1$ results in variance ratio $\frac{1}{2*(1-\rho)} = 5.00$, so the relative benefit of using the optimal ratio is not very big. On the other hand, one should be aware that the variance on $p(v)$ may increase using CV if the quadratic approximation is poor relative to standard Monte Carlo; however, this is also the case for [Cardenas et al., 1999] and [Glasserman et al., 1999a], since those methods are based on the quadratic approximation, which in this case is a poor approximation. The benefit of CV is rapidly decreasing as the correlation decreases, and the potential of variance reduction using CV is thus highly dependent on the control variables available. An absolute correlation below 0.7 will result in variance ratios less than 2. It is therefore extremely important that a good approximation is chosen.

2.1 Quadratic approximation as control variate

A quadratic function of normal distributed variables is known to be a linear combination of non-central χ^2 variables and a normal distributed variable. Furthermore, the characteristic function of the linear combination is known so the distribution can be computed through inversion of the characteristic function and integration of the density. Following the procedure of [Davies, 1980], the resulting exceedance probability $P(Q(x) < v)$ can be computed numerically with narrow precision bounds. A quadratic approximation $Q(x)$ of $\Delta V(x)$ is therefore a natural candidate for a control variate, and for small moves in the risk factor the second-order Taylor approximation is obviously highly correlated with ΔV. For more extreme moves in the risk factors, as it is the case for computation of small exceedance probabilities, the correlation declines, and this finding leads [Glasserman et al., 1999a] to the conjecture that CV based on a quadratic approximation are of little use for variance reduction in the case of simulation of VaR/small exceedance probabilities. On the other hand, [Cardenas et al., 1999] apply the second-order Taylor approximation, together with a simple stratification/IS strategy, and that method does seem to reduce variance although no variance ratios are formally computed. In this section we go through the details of variance reduction, using the quadratic approximation as CV. We will ultimately apply CV together with IS.

Performing a quadratic expansion of the change in portfolio value around μ leads to the following expression:

$$Q(\mu; x) \equiv \Delta V(\mu) + \theta \Delta t + \delta'(x - \mu) + \frac{1}{2}(x - \mu)'\Gamma(x - \mu) \approx \Delta V(x),$$

where θ, δ, and Γ depend on μ. If μ is chosen to be the mean of x and x has covariance matrix $C'C$, $(x - \mu)$ can be rewritten as $C'z$, where z is a vector of IID normal distributed variables. Letting Λ be the diagonal matrix of eigenvalues to the matrix $\frac{1}{2}C\Gamma C'$, with corresponding matrix of eigenvectors U, i.e. such that

$$\frac{1}{2}C\Gamma C' = U'\Lambda U,$$

Replacing C by $D = CU$ we find that $D'z$ has the same distribution as $x - \mu$, and if the quadratic approximation is rewritten in terms of these standard normals, z, we get

$$Q(\mu; x) = \Delta V(\mu) + \theta\Delta t + \delta' D' z + z'\Lambda z.$$

It is now a simple task to rewrite this quadratic form as a linear combination of non-central χ^2 distributions and one normal distributed variable (see [Rouvinez, 1997]), and the approximate exceedance probability is easily computed by inversion of the characteristic function, using e.g. [Davies, 1980].

$$p_{cv}(v) = P(\delta' D' z + z'\Lambda z \leq v - \Delta V(\mu) - \theta\Delta t).$$

Sampling of the control variate is equally easy as the quadratic function is very cheap to evaluate, compared to a large portfolio of financial securities. Since the cost of applying CV is small - the largest extra cost is associated with computing Delta and Gamma sensitivities, but the number of evaluations of the portfolio needed to find these sensitivities are modest, compared to the number of necessary evaluations to simulate with low variance. Even a moderate variance reduction can therefore justify the use of Control Variates.

2.2 Importance sampling based on the quadratic approximation

Exponential Shifting is a common way of performing variance reduction for simulation of exceedance probabilities (see e.g. [Fishman, 1995]). In [Glasserman et al., 1999a] and [Glasserman et al., 1999b] this variance reduction method is applied in combination with stratified sampling for VaR, where the change of measure is based on the quadratic approximation of the change in portfolio value. With exponential shifting, both the mean and variance of the underlying risk factors are changed, but in a very simple way, so that the likelihood ratio remains simple. In [Glasserman et al., 1999c] minor modifications of this basic IS method are discussed, but neither of the modifications appear to have significant implications for the variance ratios, so in this study we limit IS to be the change of measure suggested in [Glasserman et al., 1999a].

Suppose the risk factors are normally distributed with mean 0 and covariance matrix $C'C$. Performing a quadratic expansion around 0 using the same manipulations as in the previous Section yields

$$Q(0; z) = \theta\Delta t + \delta' D' z + z'\Lambda z.$$

The second moment of the estimator using one sample is simply

$$E^*(\mathbf{1}_{(-\infty,v)}(\Delta V(z)) l^2(z)) \approx E^*(\mathbf{1}_{(-\infty,v)}(Q(0; z)) l^2(z)).$$

Remark: The method proposed by [Cardenas et al., 1999] can be seen as an attempt to use the zero variance IS density corresponding to the quadratic approximation of ΔV.

Choosing $f^*(x) \propto f(x)\mathbf{1}_{(-\infty,v)}Q(0;z)$ is possible, so, if the quadratic approximation is sufficiently good, very large variance reductions can be achieved. If, on the other hand, the quadratic approximation is not sufficiently accurate, large errors may arise from using $f^*(x)$ since no sampling is performed in the set $\{z \, |Q(0;z) > v\}$. As a consequence [Cardenas et al., 1999] choose a modification of $f^*(x)$, where the entire state space is sampled and they furthermore allow for change of the sampling distribution if the quadratic approximation is too poor.

The last term above can be bounded using the inequality $\mathbf{1}_{(-\infty,v)}(y) \leq \exp(\vartheta(v-y))$, where $\vartheta \geq 0$. The change of measure is chosen so that the upper bound for the second moment of the quadratic approximation is minimized. The particular change of measure is performed by changing the covariance of z from I to $B_\vartheta = (I + 2\vartheta\Lambda)^{-1}$ and the mean of z from 0 to $\mu_\vartheta = -\vartheta B_\vartheta D\delta$. This change of sampling distribution, suggested by [Glasserman et al., 1999a], implies that the likelihood ratio is exponentially affine in Q.

$$l(z) = \exp(\vartheta Q(0;z) + \psi(\vartheta)),$$

where

$$\psi(\vartheta) = \frac{1}{2}\sum_{i=1}^{m}\frac{(\vartheta(\delta' D)_i)^2}{1+2\vartheta\Lambda_{ii}} - \frac{1}{2}\log(1+2\vartheta\Lambda_{ii})$$

is the logarithm of the moment generating function of Q evaluated at $-\vartheta$. The bound can then simply be expressed as

$$\begin{aligned} E^*(\mathbf{1}_{(-\infty,v)}(Q(0;z))l^2(z)) &= E^*(\exp(2\vartheta Q(0;x) + 2\psi(\vartheta))\mathbf{1}_{(-\infty,v)}(Q(0;z)) \\ &\leq \exp(2\vartheta v + 2\psi(\vartheta)). \end{aligned}$$

ϑ is then simply chosen to minimize the upper bound for the second moment. Note, the likelihood ratio tends to be small, whenever the quadratic approximation is small, so exponential shifting appears to be a good change of measure if the quadratic approximation is reasonably accurate.

The additional per sample cost of applying importance sampling is not high compared to the cost of evaluating the portfolio. The only additional cost is the cost of evaluating the likelihood ratio, and this is a very fast operation. IS should therefore be applied to reduce variance, even if the variance ratios are moderate.

2.3 Importance Sampling and Control Variates

For both of the proposed methods the additional per sample cost is moderate, compared to the cost of evaluating the portfolio value, but there are further costs of reducing simulation variance. For Control Variates there is a cost associated with computing the sensitivities, computing eigenvalues and eigenvectors, and inverting the characteristic function. For IS there are setup costs in connection with computing sensitivities, eigenvalues and -vectors, and finding the optimal change of measure. Since the sampling variance of raw Monte Carlo is very high when the exceedance probability is close to 0, the overhead costs associated with setting up the simulation and computing the estimate are insignificant, compared to the large costs associated with the repeated revaluations of the portfolio. In this light, combining several variance reduction methods is a straightforward way of minimizing the number of revaluations necessary to achieve a given variance on the estimate.

[Glasserman et al., 1999a] combine IS with Stratified Sampling, resulting in variance reductions up to a factor of 450, in their test portfolios. IS alone appears to provide most of the variance reduction. In their sample portfolios a large part of the total variance reductions are produced by IS alone, and the event that the level is exceeded is not rare under the IS measure, therefore a suitable approximation of the change in portfolio value might be able to produce a large *additional* variance reduction.

The quadratic approximation used in this paper differ in one important aspect from the approximation in [Cardenas et al., 1999] and [Glasserman et al., 1999a]. They use a quadratic expansion around 0 as their control variate. However, for estimates of extreme loss probabilities, only large changes in risk factors tend to influence the estimate, and the correlation is therefore not close to 1. This obviously leads to decreasing variance reductions for decreasing exceedance probabilities, which the results in [Glasserman et al., 1999a] confirm. However, since Exponential Shifting changes the sample mean such that approximately half of the simulated changes in portfolio values are below the reference point v, a Taylor expansion around this mean is suspected to have a much higher correlation with the change in portfolio value.

After selecting the sampling distribution of z, the distribution of the risk factors $x = D'z \sim N_m(D'\vartheta B_\vartheta D\delta, D'B_\vartheta D)$. Performing a second-order Taylor expansion around the shifted mean for x yields a new approximation of the change in value

$$\Delta V \approx \Delta V(\hat{\mu}) + \hat{\theta}\Delta t + \hat{\delta}'(x - \hat{\mu}) + \frac{1}{2}(x - \hat{\mu})'\hat{\Gamma}(x - \hat{\mu}).$$

Since the situation is completely similar to the one in Section 2.1, we can compute eigenvalues and -vectors for the new quadratic approximation, and represent the quadratic form as follows

$$\hat{Q}(\hat{\mu}; \hat{z}) = \hat{\theta}\Delta t + \hat{\delta}'\hat{D}'\hat{z} + \hat{z}'\hat{\Lambda}\hat{z}.$$

Sampling and numerical inversion of the characteristic function for this control variate is now fully straightforward. We simulate the control variate under the IS distribution. This ensures that many simulation paths influence the control variate, such that its sampling variance is reduced. Furthermore, the control variate, sampled under the IS distribution is expected to have a higher correlation with the change in portfolio value sampled under the IS distribution. The true value of the control variate is computed under the original measure. There is additional overhead associated with this variance reduction method, since the sensitivities have to be recomputed. The rationale for using this extra overhead is that the needed number of revaluations of the portfolio to achieve a given accuracy is reduced, so the total simulation time is reduced.

2.4 Importance Sampling and Stratified Sampling

The Taylor expansion around the shifted mean, can be used for stratification, in exactly the same way as it is used in [Glasserman et al., 1999a]. Only, stratification is performed in the new quadratic form, which is expected to a better approximation of the change in portfolio value. The stratification, and hence the variance reduction is therefore expected to be better. The extra overhead is the computation of sensitivities, but this is again justified for even modest variance reductions. For a detailed discussion of IS and Stratified sampling, please refer to [Glasserman et al., 1999a]. We present performance results for the described variance reduction methods in the following Section.

3 Numerical results

In this section we compute variance ratios for the variance reduction methods for a variety of portfolios. Our economy is set up similarly to the one in [Glasserman et al., 1999a], and to begin with we compute variance ratios for portfolios from Table 3 in that paper. We have ten underlying risk factors and, as they do, we consider portfolios of stock options. All underlying assets have an initial value of 100, an annual volatility of 30%, and are uncorrelated in portfolios 1 - 7 . Furthermore, we have a riskless, continuously compounded short rate of 5%.

The portfolios are:

1. **0.5 year ATM:** Short ten ATM[5] calls and short five ATM puts, with 0.5 years to maturity.
2. **0.1 year ATM:** As Portfolio 1, but with 0.1 years to maturity.
3. **Zero delta:** Short ten ATM calls, with 0.5 years to maturity, and a short put position, such that the initial delta is zero.
4. **0.25 OTM:** Ten short calls, struck 10% above ATM strike, and short ten puts also 10% OTM, all with maturity 0.25 years.
5. **0.25 ITM:** Ten short calls and puts with maturity 0.25 years. The calls and puts are 10% ITM.
6. **One dominant eigenvalue:** Same as portfolio 3, but with number of options on asset 1 increased by a factor 10.
7. **Linear eigenvalues:** Same as portfolio 3, but with number of options on asset i increased by a factor i, $i = 1, \dots, 10$.

[5]Here ATM corresponds to options with strike level equal to the forward value of the underlying.

Each simulation is carried out by sampling 120,000 outcomes of the risk factors, and variance ratios are estimated from 100 resimulations of the exceedance probability. The results are presented in Tables 1 and 2, below. For each portfolio we compute variance ratios for 3 different exceedance probabilities - approximately 5%, 1%, and 0.5%[6]. Variance ratios are computed for the following variance reduction methods: Importance Sampling alone, (IS), Control Variates alone - one where the CV is approximated around 0 (CV - 0) and one where the CV is approximated around the IS mean (CV - μ), the combination of the two IS and CV - one where the regression variable α is set to 1 (ISCV), and one where the optimal regression variable is presampled (ISCV - α), the combination of IS and stratification on the quadratic approximation - one where the stratification is performed on the approximation around 0 (ISS - 0) and one where the IS mean is used (ISS - μ). We finally provide estimates of the regression parameter (α). For the two alternative stratifications we use the same number of samples as [Glasserman et al., 1999a], that is, we stratify the sample of 120,000 simulations in 40 equiprobable subsamples with 3,000 simulations in each subsample. In Table 1 we report variance ratios for the simplest methods, where only one variance reduction method is applied.

Table 1 documents that both IS and CV can reduce simulation variance substantially. If the mean of the IS is used for the second-order approximation ($\text{CV} - \mu$), the variance ratios applied increase relative to standard CV ($\text{CV} - 0$), and they do not decrease as drastically as standard CV variance ratios when the exceedance probability decreases, and CV is apparently able to decrease variance rather much, even at very high confidence levels. For high exceedance probabilities, e.g. 5% CV reduce variance more than IS, whereas IS performs better for low exceedance probabilities. In Table 2 we present variance ratios for all methods that apply two variance reduction methods simultaneously. We also present estimates of the optimal regression coefficient.

All methods in Table 2 yield very high variance reductions, in comparison with the setting where only one variance reduction method is applied. The IS and Stratification method proposed by [Glasserman et al., 1999a] clearly underperforms the other methods investigated here, simply because the second-order approximation applied for stratification is too imprecise. Both ISCV and $\text{ISS} - \mu$ produce very high variance ratios for all portfolios. In many cases variance ratios are at least a factor 2 - 3 higher for these methods than for ISS $- 0$.

The lowest overall variance ratio for the seven portfolios has increased from 6 for the ISS $- 0$ method to 28 for the ISCV methods, and to 40 for $\text{ISS} - \mu$. The two ISCV methods perform almost equally well, $\text{ISCV} - \alpha$ is on average slightly better than $\text{ISCV} - 0$, but since the regression coefficient is close to 1 the improvement is merely marginal. ISCV is the best overall method, whenever the quadratic approximation is very good, i.e. when the regression coefficient is very close to 1. Portfolio 1 and Portfolio 5 are examples of portfolios where the quadratic approximation is very close to the change in portfolio value. In case the quadratic approximation is slightly less precise, ISS - μ appears to be better

[6] Some probabilities are chosen slightly above/below these levels to make this setting as similar to [Glasserman et al., 1999a] as possible.

	$P(\Delta V < v)$	IS	CV − 0	CV − μ
Portfolio 1	5.3%	7.8	41.5	74.8
	1.0%	28.0	17.4	46.0
	0.5%	61.9	20.6	45.3
Portfolio 2	5.0%	7.3	6.0	7.4
	1.1%	17.7	5.5	4.8
	0.3%	62.4	4.8	5.1
Portfolio 3	4.7%	8.2	9.1	11.5
	1.1%	22.8	7.6	8.4
	0.5%	33.4	6.0	7.6
Portfolio 4	5.0%	6.3	8.3	14.1
	1.1%	34.4	6.8	19.4
	0.5%	39.7	3.2	8.4
Portfolio 5	5.0%	6.2	23.2	24.2
	1.1%	24.0	14.5	22.2
	0.5%	58.2	13.3	26.8
Portfolio 6	5.0%	4.5	33.4	28.1
	1.2%	6.8	6.7	8.1
	0.5%	23.5	5.0	5.5
Portfolio 7	5.0%	7.4	18.3	16.6
	1.0%	17.5	7.0	8.0
	0.5%	37.2	4.1	6.3

Table 1: Table of variance ratios for simulation of exceedance probabilities, where only one variance reduction technique is applied. The methods applied are Importance Sampling (IS) using exponential shifting, Control Variates based on a quadratic expansion around 0 (CV − 0), and Control Variates based on the Importance Sampling mean (CV − μ).

than the other methods. The lowest variance ratios are observed for Portfolio 2, where the short option maturity implies that the second-order approximation is relatively imprecise.

The key to the high variance ratios in Tables 1 and 2 is that the quadratic approximation is relatively good – even for the portfolios with zero Delta[7]. This is indicated by the high regression coefficients for the Control Variates, which for the portfolios above are at least 0.9. In Table 3 we present variance ratios for a Delta-hedged portfolio, where the hedge is performed using the underlying, rather than the options themselves. It turns out that the quadratic approximation is much worse for these portfolios. Three example portfolios are investigated:

8. **0.5 year ATM, Delta hedged:** Short 10 ATM call options, short 5 ATM put options, all options maturing in 0.5 years and a long position in the underlying to create a zero Delta portfolio.

[7]The portfolios with zero Delta are concave functions, so it is no surprise that the quadratic approximation is reasonably good.

	$P(\Delta V < v)$	ISCV	ISCV $-\ \alpha$	ISS $-$ 0	ISS $-\ \mu$	α
Portfolio 1	5.3%	198	189	98	173	1.00
	1.0%	592	496	126	330	1.00
	0.5%	1415	1320	286	795	1.00
Portfolio 2	5.0%	28	30	27	40	0.96
	1.1%	56	56	28	83	0.93
	0.3%	139	126	87	194	0.91
Portfolio 3	4.7%	38	44	56	44	0.96
	1.1%	82	76	135	126	0.93
	0.5%	120	114	167	161	0.90
Portfolio 4	5.0%	45	57	25	77	0.97
	1.1%	197	167	62	204	0.96
	0.5%	198	236	79	381	0.96
Portfolio 5	5.0%	94	79	21	86	0.98
	1.1%	279	268	59	181	0.98
	0.5%	597	478	154	556	0.97
Portfolio 6	5.0%	58	55	6	77	1.00
	1.2%	52	54	17	74	0.95
	0.5%	65	82	33	99	0.90
Portfolio 7	5.0%	48	47	16	61	0.97
	1.0%	70	77	27	102	0.92
	0.5%	101	94	33	93	0.90

Table 2: Table of variance ratios for variance reduction methods where two techniques are applied simultaneously. We apply IS combined with CV, where we set the regression parameter to 1 (ISCV). For comparison we present variance ratios for ISCV with α chosen optimally (ISCV $-\ \alpha$), the suggested variance reduction method in [Glasserman et al., 1999a] Importance Sampling and Stratification on the quadratic expansion around 0 (ISS $-$ 0), and finally the combination of IS and Stratification based on an expansion around the IS mean (ISS $-\ \mu$). We furthermore present the optimal regression parameters (α).

9. **0.25 year OTM, Delta hedged:** Short 10 OTM call options, short 5 OTM put options, which are Delta hedged with the underlying. The options mature in 0.25 years, and are 10% out of the money.

10. **0.25 year ATM/OTM, Delta hedged:** Short 10 ATM call options, short 10 OTM put options, delta hedged with the underlying. 0.25 years to option maturity.

These portfolios exemplify how Delta hedged portfolios may be poorly approximated by the second-order approximation, and how this poor approximation results in low variance ratios. Table 3 reports how the poor approximations result in low variance ratios. In fact, variance ratios might be below 1 for the proposed methods if the approximation is too poor. Since the Delta is zero is the IS mean zero, and the two methods based on IS and Stratification, ISS $-\ \mu$ and ISS $-$ 0 are identical.

	$P(\Delta V < v)$	α	IS	ISCV	ISCV $-\ \alpha$	ISS $-\ \mu$
Portfolio 8	5.0%	0.35	5.6	1.2	3.4	7.4
	1.0%	0.46	14.2	9.6	19.6	20.0
	0.5%	0.51	20.2	13.2	19.9	24.6
Portfolio 9	5.0%	0.79	4.8	7.3	6.7	10.1
	1.0%	0.70	16.6	18.7	22.9	27.4
	0.5%	0.67	44.4	47.3	53.6	67.2
Portfolio 10	5.1%	0.17	0.3	0.3	0.2	0.3
	1.1%	0.20	0.1	0.1	0.1	0.1
	0.5%	0.19	0.1	0.1	0.1	0.1

Table 3: Variance ratios for Delta hedged portfolios, where the portfolios are hedged with the underlying security. Results are presented for Importance Sampling combined with Control Variates with and without taking the optimal regression coefficient into consideration (ISCV $-\ \alpha$ and ISCV), and Importance Sampling combined with Stratification on the second-order approximation around the IS sampling mean (ISS $-\ \mu$).

First, we note that the regression coefficients are reduced substantially for all test portfolios. The regression coefficients are between 0.17 and 0.79 for the three portfolios under consideration. For Portfolio 9, where the second-order approximation catches most of the variation in portfolio value, the variance ratios are relatively large; however, they are not nearly as impressive as the variance ratios for Portfolios 1 $-$ 7. The variance ratios are smaller for Portfolio 8 where the regression coefficient is approximately 0.5, and for Portfolio 10 variance ratios are below 1, hence variance reduction based on the second-order approximation tends to *increase* variance.

4 Conclusion

Substantial variance reductions on Monte Carlo estimates of exceedance probabilities/Value at Risk can be achieved using Control Variates and Importance Sampling, or alternatively, using the Importance Sampling technique in combination with Stratification on a quadratic approximation around the IS mean. The methods are particularly useful when Delta is large relative to Gamma, and the second-order approximation is therefore particularly good. When Delta is small compared to Gamma – e.g. for Delta hedged portfolios or for ATM swap portfolios, the variance ratios are much smaller, and simulation variance may even increase relative to standard Monte Carlo if the quadratic approximation is poor.

The assumption of multivariate normality is a central assumption in this paper; however, this assumption is critical for a large portion of the financial risk factors since many of these risks are heavy-tailed or skewed. The variance reduction methods described here could easily be extended to jump-diffusion setting where the jumps are Poisson distributed and jump sizes are normally distributed, [Duffie and Pan, 1999] apply [Davies, 1980] inversion of the characteristic function in this particular setting, and simulation of both the Poisson distribution and the normal distributions is straightforward. The characteristic function of a multivariate t-distribution with common degrees of freedom is also known, and it is easy to generate deviates from the multivariate t-distribution, see e.g. [Embrechts et al., 1999], but it is subject to future research to determine whether the multivariate t-distribution is a good description of all risk factors.

The results in Table 3 demonstrate that variance reductions based on a local second-order approximation might increase variance. Adaptive simulation of VaR is a potential way of obtaining variance reductions even for portfolios, where the local approximation is very poor. [Glasserman et al., 1999b] demonstrate how stratification can be improved if presampling is applied, at least for the portfolios they examine. Improvements using pilot simulation can possibly be achieved for other variance reduction methods and this is subject to further research.

References

[Avramidis and Wilson, 1990] Avramidis, A. N. and Wilson, J. R. (1990). Control variates for stochastic network simulation. *Proc. 1990 Winter Simulation Conference*, pages 323 – 332.

[Beder, 1995] Beder, T. S. (1995). Var: Seductive but dangerous. *Financial Analyst Journal*, pages 12–24.

[Cardenas et al., 1999] Cardenas, J., Fruchard, E., Picron, J-F. Reyes, C., Walters, K., and Yang, W. (1999). Monte carlo within a day. *Risk*, february:55–59.

[Davies, 1980] Davies, R. B. (1980). The distribution of a linear combination of χ^2 random variables. *Applied Statistics*, 29:323–333.

[Duffie and Pan, 1997] Duffie, D. and Pan, J. (1997). An overview over value at risk. *Journal of Derivatives*, pages 7–49.

[Duffie and Pan, 1999] Duffie, D. and Pan, J. (1999). Analytical value-at-risk with jump-diffusions. Working paper, Graduate School of Business, Stanford University.

[Embrechts et al., 1999] Embrechts, P., McNeil, A. J., and Straumann, D. (1999). Correlation and dependency in risk management. Working paper, ETH, Zürich.

[Fishman, 1995] Fishman, G. S. (1995). *Monte Carlo: Concepts, Algorithms, and Applications*. Springer.

[Glasserman et al., 1999a] Glasserman, P., Heidelberger, P., and Shahabuddin, P. (1999a). Importance sampling and stratification for value-at-risk. Working paper, IBM research division.

[Glasserman et al., 1999b] Glasserman, P., Heidelberger, P., and Shahabuddin, P. (1999b). Stratification issues in estimating value-at-risk. Working paper, IBM research division.

[Glasserman et al., 1999c] Glasserman, P., Heidelberger, P., and Shahabuddin, P. (1999c). Variance resuction techniques for estimating value-at-risk. Working paper, IBM research division.

[Jorion, 1997] Jorion, P. (1997). *Value at Risk*. McGraw-Hill.

[J.P.Morgan, 1996] J.P.Morgan (1996). *RiskMetrics technical document, 4th ed.* J. P. Morgan.

[Pritsker, 1996] Pritsker, M. (1996). Evaluating Value at Risk Methodologies: Accuracy versus Computational Time. University of Pensylvania, The Wharton School, Financial Institutions Center, Working Paper Series, 96-48.

[Rouvinez, 1997] Rouvinez, C. (1997). Going greek with var. *Risk*, 10(2):57–65.