# A Large Deviations Analysis of Quantile Estimation with Application to Value At Risk

Xing Jin and Michael C. Fu
R.H. Smith School of Business
University of Maryland
College Park, MD 20742-1815

October 2001

**Abstract**

Quantile estimation has become increasingly important, particularly in the financial industry, where Value-at-Risk has emerged as a standard measurement tool for controlling portfolio risk. In this paper we apply the theory of large deviations to analyze various simulation-based quantile estimators. First, we show that the coverage probability of the standard quantile estimator converges to one *exponentially fast* with sample size. Then we introduce a new quantile estimator that has a *provably faster* convergence rate. Furthermore, we show that the coverage probability for this new estimator can be *guaranteed* to be 100% with sufficiently large, *but finite*, sample size. Numerical experiments on a VaR example illustrate the potential for dramatic variance reduction.

## 1 Introduction

The use of quantiles as primary measures of performance has gained prominence recently, especially in the financial industry, where Value-at-Risk (VaR) has emerged as a standard tool to measure and control the risk of trading portfolios. In terms of statistics, VaR is nothing more than a quantile of a portfolio's potential profit and loss over a given time period. Quantiles provide additional or alternative information about the distribution of the performance measure of interest, say $Y$, where the $r$-th quantile of $Y$ is defined by $P[Y \leq \xi_r] = r$ for prespecified $r$ $(0 < r < 1)$. The most well-known quantile is the median, where $r = 0.50$. If, for instance, $Y$ is the delay experienced by a customer in a queueing system, then 50% of the customers experience delays less than the median $\xi_{0.50}$, but 5% of the customers experience delays longer than $\xi_{0.95}$. Thus, quantiles are clearly useful in describing tail behavior.

Our setting is that of complex stochastic systems where simulation is required. Variance reduction techniques are crucial for improving the efficiency of simulation, and there is a huge body of literature dedicated towards this goal, but it is almost exclusively directed towards the expected value of an output random variable. However, output analysis for the estimation of quantiles differs significantly from estimation of means. The limited literature relevant to our work is summarized as follows. Perhaps the earliest work is Hsu and Nelson (1990) and Hesterberg and Nelson (1998), who applied control variates to obtain variance reduction in simulation-based quantile estimation. Most closely related to our work is Avramidis and Wilson (1998), who employed correlation-induction techniques to improving quantile estimation. In all of these, the traditional approach to evaluating the performance of the estimator is to invoke the central limit theorem to estimate the variance of the estimator. In other words, the goal is to minimize the confidence interval half-length, given the confidence level. This approach, however, suffers some drawbacks. First of all, the variance of the limiting normal distribution is just an asymptotic variance of the estimator, and it is often difficult to verify the assumptions underlying the central limit theorem for a complex system. Secondly, it may be cumbersome to determine the run length (or replication counts) of simulation in order to achieve a prespecified precision in terms of variance, whereas in practice, an institution with risk exposure of its portfolios may be more concerned with the probability that the estimation error of a quantile is within its tolerance.

Based on these observations, we propose a new approach to dealing with the performance of quantile estimates. Applying the theory of large deviations, we show that the probability of the standard quantile estimator falling in a neighborhood of the true quantile, which we call the **coverage probability**, converges to one *exponentially fast* with increasing sample size. Then, modifying the correlation-induced Latin Hypercube Sampling (LHS) estimator of Avramidis and Wilson (1998), we propose a new quantile estimator, for which the probability of belonging to a neighborhood of the true quantile (which we call the *coverage probability*) is one for sufficiently large, *but finite*, sample size. Furthermore, in special cases, an *exact* (here meaning non-asymptotic) upper bound for the variance of the estimator can be obtained, and it is shown that the convergence rate is $O(1/n)$, as opposed to the usual Monte Carlo $O(1/\sqrt{n})$ rate. Finally, we apply the estimator to VAR estimation in a typical financial model. Numerical experiments demonstrate substantial variance reduction compared with independent sampling and with the estimator of Avramidis and Wilson (1998).

To be more specific, let $F(\cdot)$ denote the (unknown) cumulative distribution function. In terms of the inverse c.d.f., the quantile is given by $\xi_r = F^{-1}(r)$, where

$$F^{-1}(u) = \min\{t : F(t) \geq u\} \text{ for all } u \in (0,1).$$

A natural estimator for $\xi_r$ is the direct-simulation estimator

$$\widehat{\xi}(n) = \min\{t : F_n(t) \geq u\}, \tag{1}$$

where $F_n(t)$ is the empirical discrete c.d.f. based on sample $\{Y_i, i = 1, ..., n\}$. In terms of order statistics $Y_{(1)} \leq Y_{(2)} \leq ... \leq Y_{(n)}$ obtained by sorting the observations $\{Y_i, i = 1, ..., n\}$ in ascending order, $F_n(t)$ is defined as follows:

$$F_n(t) = \begin{cases} 0, & \text{if } t < Y_{(1)}, \\ \frac{i}{n}, & \text{if } Y_{(i)} \leq t < Y_{(i+1)} \text{ and } 1 \leq i \leq n-1, \\ 1, & \text{if } Y_{(n)} \leq t. \end{cases} \tag{2}$$

Avramidis and Wilson (1998) discuss two kinds of quantiles estimators, single-sample and multiple-sample estimators. The single-sample estimator is just the direct-simulation estimator given by (1), whereas the multiple-sample estimator is obtained by computing the sample mean using $k$ independent single-sample quantile estimates based on sample size $m = n/k$ ($n$, the overall number of samples, is chosen to be an integral multiple of $k$ such that $m$ is an integer). Let $\widehat{\xi}^{(i)}(m)$ denote the $i$th direct-simulation estimate, $i = 1, ..., k$. Then the direct-simulation multiple-sample estimator is simply the mean of the single-sample estimators:

$$\bar{\xi}(k,n) = k^{-1} \sum_{i=1}^{k} \widehat{\xi}^{(i)}(m). \tag{3}$$

Avramidis and Wilson (1998) improve upon the standard estimator by inducing negative correlation between each pair of $\{\widehat{\xi}^{(i)}(m), i = 1, ..., k\}$. In other words, negative correlation is induced between corresponding quantile estimators in different samples while mutual independence is maintained within each sample. Establishing a central limit theorem for their estimator, they obtain the canonical Monte Carlo $1/\sqrt{n}$ convergence rate.

By applying the theory of large deviations, we show that for all of the estimators discussed thus far,

$$P\left[\left|\widehat{\xi}(n) - \xi_r\right| \geq \varepsilon\right] \leq \exp[-nI_n(\varepsilon)],$$

where $\varepsilon$ is the tolerance and $I_n(\varepsilon)$ is the decay rate, which we use as an indication of quality of the estimator, in addition to the usual measure of estimator variance. We then replace the empirical c.d.f. given by (2) with a different form (see (9) in Section 4). We show that the property of negative dependence, which holds for Latin Hypercube Sampling, guarantees a larger $I_n(\varepsilon)$ than for i.i.d samples. We propose an estimator such that for $\varepsilon > 0$, there exists a *finite* $n$ such that the coverage probability is 100%, e.g.,

$$P\left[\left|\widehat{\xi}(n) - \xi_r\right| < \varepsilon\right] = 1!$$

Furthermore, instead of the usual $1/\sqrt{n}$ convergence rate ($1/n$ decrease in variance), our analysis allows us to establish a $1/n$ rate of convergence ($1/n^2$ decrease in variance) for one dimensional case.

In sum, our work makes the following new contributions to simulation-based quantile estimation:

- In addition to being the first application of the theory of large deviations to this important setting, the theory allows us to obtain *stronger results* under *weaker assumptions* than have been obtained by using the usual asymptotic central limit theorem analysis.

- We analyze the *coverage probability* in addition to the estimator variance as a measure of performance, and establish an *exponential convergence rate*, for which the property of negative dependence ensures a faster convergence rate over independent sampling.

- We introduce a new estimator that has *provably better theoretical properties* and shows *significantly better empirical performance.*

The rest of this paper is organized as follows. In the next section, we present some preliminary results. First, we introduce the notion of negative dependence, and establish some properties of random vectors satisfying this property. Then we derive some large deviations results for negatively dependent sequences. As an important example, we review Latin Hypercube Sampling. In Section 3, the main results using the large deviations theory for quantile estimation are established. The new quantile estimator is presented in Section 4, along with the analysis of its theoretical convergence properties. Section 5 contains the numerical experiments for theVaR finance example. The more technical details of some of the proofs are included in the appendix.

# 2 Preliminary Results

## 2.1 Negative Dependence

In this section we present the notation of negative dependence, a generalization of the notation of negatively quadrant dependence defined by Lehmann (1966); see also Nelsen (1999). The pair $(X, Y)$ or its (joint) distribution $F$ is *negatively quadrant dependent* if

$$P[X \leq x, Y \leq y] \leq P[X \leq x]P[Y \leq y] \text{ for all } x, y. \tag{4}$$

Mutual independence is the case where equality holds in (4).

**Definition 2.1:** The random variables $X_i, i = 1, ..., n$ are called *negatively dependent* if the following two inequalities hold for all $x_1, ..., x_n$

$$P[X_1 \leq x_1, ..., X_n \leq x_n] \leq P[X_1 \leq x_1] \cdot ... \cdot P[X_n \leq x_n], \tag{5}$$

$$P[X_1 \geq x_1, ..., X_n \geq x_n] \leq P[X_1 \geq x_1] \cdot ... \cdot P[X_n \geq x_n]. \tag{6}$$

Obviously, letting $X_1 = X,\ X_2 = Y, x_i = \infty$ for $i = 3, ..., n$ in (5) gives (4). Thus any pair from a set of negatively dependent random variables are negatively quadrant dependent.

The real-valued functions $f_1(\cdot), ..., f_n(\cdot)$ of $d$ arguments are *concordant* for the $i$th coordinate if, considered as functions of the $i$th coordinate (with all other coordinates held fixed), they are monotone in the same direction, i.e., either all non-decreasing or all non-increasing.

**Lemma 2.1:** If $X_i, i = 1, ..., n$ are negatively dependent with $E|X_i| < \infty, i = 1, ..., n$, then

$$E(X_i X_j) \leq E(X_i)E(X_j), \text{for } i \neq j, i, j = 1, ..., n, \tag{7}$$

with equality holding if $X_i, i = 1, ..., n$ are independent. In particular, (7) implies

$$Cov[X_i, X_j] = E(X_i X_j) - E(X_i)E(X_j) \leq 0.$$

Furthermore, if $X_i, i = 1, ..., n$ are non-negative and $E(X_1 \cdot ... \cdot X_n)$ are finite, then

$$E(X_1 \cdot ... \cdot X_n) \leq E(X_1) \cdot ... \cdot E(X_n). \tag{8}$$

**Proof:** The proof of (7) can be found in Lehmann (1966). We now prove (8). By noticing that $X_i, i = 1, ..., n$ are non-negative,

$$E(X_1 \cdot ... \cdot X_n) = E \int_0^\infty ... \int_0^\infty I(x_1, X_1) \cdot ... \cdot I(x_n, X_n) dx_1 ... dx_n,$$

where $I(x, X) = 1$ if $x \leq X$ and $= 0$ otherwise. Since $E(X_1 \cdot ... \cdot X_n)$ is assumed finite, we can exchange expectation and integral, which gives

$$\begin{aligned}
E(X_1 \cdot ... \cdot X_n) &= \int_0^\infty ... \int_0^\infty E[I(x_1, X_1) \cdot ... \cdot I(x_n, X_n)] dx_1 ... dx_n \\
&= \int_0^\infty ... \int_0^\infty P[X_1 \geq x_1, ..., X_n \geq x_n] dx_1 ... dx_n \\
&\leq \int_0^\infty ... \int_0^\infty P[X_1 \geq x_1] \cdot ... \cdot P[X_n \geq x_n] dx_1 ... dx_n \\
&= E(X_1) \cdot ... \cdot E(X_n),
\end{aligned}$$

the inequality following from (6), completing the proof of (8). □

**Lemma 2.2:** Let the random vectors $X^{(j)} = (X_1^{(j)}, ..., X_n^{(j)}), j = 1, ..., d$, be mutually independent and random variables $X_1^{(j)}, ..., X_n^{(j)}$ be negatively dependent for each $j = 1, ..., d$. Suppose real-valued functions $f_1(\cdot), ..., f_n(\cdot)$ of $d$ arguments are concordant for each coordinate, then $f_i(X_i^{(1)}, ..., X_i^{(d)}), i = 1, ..., n$ are negatively dependent.

**Proof:** See Appendix.

From Lemma 2.2, if the quantile estimators $\{\widehat{\xi}^{(i)}(m), i = 1, ..., k\}$ are negatively dependent, then quantile estimators in different samples are negatively correlated and hence, the overall variance of quantile estimator $\bar{\xi}(k, n)$ is reduced.

## 2.2 Latin Hypercube Sampling

Monte Carlo integration is often used to estimate integrals over multidimensional domains. This integration has the advantage of very general applicability and error estimation based on the central limit theorem. It is commonly done via independent sampling. Stratification can be used to increase the precision of the estimate. McKay et al. (1979) introduce Latin Hypercube Sampling (LHS) as a method of stratifying on all the input dimensions simultaneously.

Suppose the joint distribution of the random vector of parameters $X = (X_1, ..., X_d)$ is given by $F$. For now, we will assume that the components of $X$ are independent. Denote by $G_k$ the cumulative distribution function $X_k$, and let $X_{jk}$ be the $k$th component of $X^{(j)}$, the $j$th simulated value of $X$. Let $n$ be the size of sample. Then, $X_{jk}$ is generated by

$$X_{jk} = G_k^{-1}\left(\frac{\pi_k(j) - U_{jk}^*}{n}\right), \text{ for } j = 1, ..., n \text{ and } k = 1, ..., d,$$

where

(a) $\pi_1(\cdot), ..., \pi_d(\cdot)$ are permutations of the integers $\{1, ..., n\}$ that are randomly sampled with replacement from the set of $n!$ such permutations, with $\pi_k(j)$ denoting the $j$th element in the $k$th randomly sampled permutation; and

(b) $\{U_{jk}^*;\ j = 1, ..., n\ ,\ k = 1, ..., d\}$ are i.i.d. $U(0, 1)$, i.e., uniformly distributed random numbers on $[0, 1]$, sampled independently of $\pi_1(\cdot), ..., \pi_d(\cdot)$.

Since $\pi_k(\cdot)$ is a random permutation of the integers $\{1, ..., n\}$ and $U_{jk}^*$ is a uniform distributed random number on $[0, 1]$ sampled independently of $\pi_k(\cdot)$, it is not difficult to see the random variable

$$U_k^{(j)} = \frac{\pi_k(j) - U_{jk}^*}{n} \tag{9}$$

is still uniformly distributed on $[0, 1]$. And, moreover, since $\pi_k(\cdot)$ is a random permutation of the integers $\{1, ..., n\}$, every subinterval (stratum) of the form $(l - 1/n, l/n]$ for $l = 1, ..., n$ contains exactly one of the random numbers $U_k^{(j)}$, $j = 1, ..., n$, realizing a stratification.

We conclude this section by validating the negative dependence of LHS random numbers produced according to (9). Since $\{\pi_k(j); j = 1, ..., n\ ,\ k = 1, ..., d\}$ are integers, it suffices to show (5) and (6) hold with

$x_1, ..., x_n$ taking integer values. Without loss of generality, suppose $x_1 \leq ... \leq x_n$. Note that if

$$P[\pi_k(1) \leq x_1, ..., \pi_k(n) \leq x_n] = 0,$$

$$P[\pi_k(1) \leq x_1, ..., \pi_k(n) \leq x_n] \leq P[\pi_k(1) \leq x_1\} \cdot ... \cdot P\{\pi_k(n) \leq x_n];$$

if

$$P[\pi_k(1) \leq x_1, ..., \pi_k(n) \leq x_n] > 0,$$

$$\begin{aligned} P[\pi_k(1) &\leq x_1, ..., \pi_k(n) \leq x_n] = \frac{x_1(x_2-1) \cdot ... \cdot (x_k - k + 1)}{k!} \\ &\leq \frac{x_1 x_2 \cdot ... \cdot x_k}{k^k} = P[\pi_k(1) \leq x_1\} \cdot ... \cdot P\{\pi_k(n) \leq x_n], \end{aligned}$$

where the last inequality follows from the following inequalities:

$$\frac{x_i - i + 1}{n - i + 1} \leq \frac{x_i}{n}, \text{ for } i = 1, ..., n.$$

Consequently, (5) can be derived by taking conditional probability on $U^*_{jk}$, $j = 1, ..., n$ since $\{U^*_{jk};\ j = 1, ..., n\}$ are i.i.d. $U(0,1)$ sampled independently of $\pi_k(\cdot)$. (6) can be proved in an analogous manner.

## 2.3 Auxiliary Results from Large Deviations Theory

To establish our main results, we apply the large deviations principle, which yields an exponential convergence rate under appropriate conditions. In order to motivate our discussion, we briefly outline some background from the theory of large deviations (Bucklew 1990, Dembo and Zeitouni 1998, Deuschel and Stroock 1989).

Consider a random variable $Y$ with mean $\mu = E[Y]$. Its moment generating function $M(\lambda) = E[\exp(\lambda Y)]$ is viewed as an extended valued function, i.e., it can take value $+\infty$. It holds that $M(\lambda) > 0$ for all $\lambda \in \mathcal{R}, \mathcal{M}(\prime) = \infty$, and the domain $\{\lambda : M(\lambda) < +\infty\}$ of the moment generating function is an interval containing zero. The conjugate function

$$I(z) = \sup_{\lambda \in \mathbf{R}} \{\lambda z - \Lambda(\lambda)\}$$

of the logarithmic moment generating function $\Lambda(\lambda) = \log M(\lambda)$, is called the rate function of $Y$.

Consider an i.i.d. sequence $Y_1, ..., Y_n$ of replications of random variable $Y$ and let $S_n/n = n^{-1} \sum_{i=1}^n Y_i$ be the corresponding sample average. If $M(\lambda)$ exists in a neighborhood $(-\varepsilon, \varepsilon)$ of $\lambda = 0$ for some $\varepsilon > 0$, then

$$\begin{aligned} -\inf_{z \in int(\Gamma)} I(z) &\leq \liminf_{n \to \infty} \frac{\log(P[S_n/n \in \Gamma])}{n} \\ &\leq \limsup_{n \to \infty} \frac{\log(P[S_n/n \in \Gamma])}{n} \leq -\inf_{z \in cl(\Gamma)} I(z), \end{aligned}$$

where $int(\Gamma)$ and $cl(\Gamma)$ denote the interior and the topological closure, respectively, of the set $\Gamma$, and $I(z)$ is the corresponding rate function given by

$$\begin{aligned} I(z) &= \sup_{\lambda \in (-\varepsilon, \varepsilon)} \left\{ \lambda z - \lim_{n \to \infty} \frac{\log(E[\exp[\lambda S_n])}{n} \right\} \\ &= \sup_{\lambda \in (-\varepsilon, \varepsilon)} \{\lambda z - \Lambda(\lambda)\}, \end{aligned}$$

the second equality following from $E[\exp[\lambda S_n] = (E[\exp[\lambda Y])^n$, because the sequence $Y_1, ..., Y_n$ is i.i.d.

In particular, when $\Gamma = \{x : x \in [\mu - \varepsilon, \mu + \varepsilon]\}$, then

$$- \inf_{z \in int(\Gamma)} I(z) = - \inf_{z \in cl(\Gamma)} I(z)$$

and

$$\lim_{n \to \infty} \frac{\log(P[S_n/n \in \Gamma])}{n} = - \inf_{z \in cl(\Gamma)} I(z).$$

Thus, the rate function $I(z)$ characterizes the asymptotic behavior of $S_n/n$ as an estimator of $\mu = EY$. This is the key motivation for us to use the rate function as performance measure, and it will be shown that variance reduction techniques can enhance the convergence rate, as shown in the following generalization of a result in Fu and Jin (2001) needed for this purpose.

**Lemma 2.3:** Consider a sequence of negatively dependent and identically distributed random variables $\{Y_n, n \geq 1\}$ with moment generating function $M(\lambda) = E[\exp(\lambda Y_1)]$. Let $S_n = \sum_{i=1}^{n} Y_i$. If $M(\lambda)$ exists in a neighborhood $(-\varepsilon, \varepsilon)$ of $\lambda = 0$ for some $\varepsilon > 0$, then

$$P[S_n/n \geq x] \leq e^{-n\Delta_+(x,n)}, \ \forall x > E(Y_1), \tag{10}$$

$$P[S_n/n \leq x] \leq e^{-n\Delta_-(x,n)}, \ \forall x < E(Y_1), \tag{11}$$

where

$$\begin{aligned} \Delta_+(x,n) &= \sup_{0 \leq \lambda \leq \varepsilon} \left( \lambda x - \frac{\log E\{\exp[\lambda S_n]\}}{n} \right) \\ &\geq \sup_{0 \leq \lambda \leq \varepsilon} (\lambda x - \log E\{\exp[\lambda Y_1]\}) > 0, \end{aligned} \tag{12}$$

and

$$\begin{aligned} \Delta_-(x,n) &= \sup_{-\varepsilon \leq \lambda \leq 0} \left( \lambda x - \frac{\log E\{\exp[\lambda S_n]\}}{n} \right) \\ &\geq \sup_{-\varepsilon \leq \lambda \leq 0} (\lambda x - \log E\{\exp[\lambda Y_1]\}) > 0. \end{aligned} \tag{13}$$

Conversely, if $E[|Y_1|] < \infty$ and for any $x > E(Y_1)$, there exists $\alpha(x) > 0$ such that

$$P[S_n/n \geq x] \leq e^{-n\alpha(x)}, \tag{14}$$

and for any $x < E(Y_1)$, there exists $\beta(x) > 0$ such that

$$P[S_n/n \leq x] \leq e^{-n\beta(x)}, \tag{15}$$

then the moment generating function $M(\lambda)$ exists in a neighborhood $(-\varepsilon, \varepsilon)$ of $\lambda = 0$ for some $\varepsilon > 0$.

**Proof:** *Proof of Sufficiency*: Consider $x > E(Y_1)$. Noticing that for any $\lambda \geq 0$,

$$\begin{aligned} P\left[S_n/n \geq x\right] &\leq E\{\exp[\lambda(S_n - nx)]\} \\ &= \exp\left[-n\left(\lambda x - \frac{\log E\{\exp[\lambda S_n]\}}{n}\right)\right], \end{aligned}$$

hence

$$
\begin{aligned}
P\left[S_n/n \geq x\right] &\leq \exp\left[\min_{0\leq\lambda\leq\varepsilon}\left[-n\left(\lambda x - \frac{\log E\{\exp[\lambda S_n]\}}{n}\right)\right]\right] \\
&= \exp\left[-n \sup_{0\leq\lambda\leq\varepsilon}\left(\lambda x - \frac{\log E\{\exp[\lambda S_n]\}}{n}\right)\right].
\end{aligned}
$$

From Lemmas 2.1 and 2.2, $\{\exp[\lambda Y_n], n \geq 1\}$ are negatively dependent and

$$
E\{\exp[\lambda S_n]\} = E\{\prod_{i=1}^{n}\exp[\lambda Y_i]\} \leq \prod_{i=1}^{n} E\{\exp[\lambda Y_i]\} = (E\{\exp[\lambda Y_1]\})^n,
$$

implying

$$
\begin{aligned}
P\left[S_n/n \geq x\right] &\leq \exp\left[-n \sup_{0\leq\lambda\leq\varepsilon}\left(\lambda x - \log E\{\exp[\lambda Y_1]\}\right)\right] \\
&= \exp\left[-n \sup_{0\leq\lambda\leq\varepsilon}\left(\lambda x - \Lambda(\lambda)\right)\right],
\end{aligned}
$$

where $\Lambda(\lambda) = \log E\{\exp[\lambda Y_1]\}$.

Furthermore, by a Taylor's series expansion,

$$
\begin{aligned}
\sup_{0\leq\lambda\leq\varepsilon}\left(\lambda x - \Lambda(\lambda)\right) &= \sup_{0\leq\lambda\leq\varepsilon}\left(\lambda x - (\Lambda(0) + \Lambda'(0)\lambda + o(\lambda))\right) \\
&= \sup_{0\leq\lambda\leq\varepsilon}\left(\lambda(x - EY_1) + o(\lambda)\right) > 0,
\end{aligned}
$$

completing the proof of (10).

By noticing that for any $\lambda \leq 0$,

$$
P[S_n/n \leq x] \leq E\{\exp[\lambda(S_n - nx)]\},
$$

(11) can be proved in exactly the same manner. We now turn to proving the necessity portion.

*Proof of Necessity*: Suppose (14) and (15) are true. Without loss of generality, assume that $EY_1 = 0$. Thus, for any integer $m > 0$, on the one hand, by (14),

$$
P\left\{Y_1 \geq \frac{3m}{2}, \frac{\sum_{i=2}^{m} Y_i}{m-1} \geq -\frac{1}{2}\right\} \leq P\left\{\frac{\sum_{i=1}^{m} Y_i}{m} \geq 1\right\} \leq e^{-m\alpha(1)}. \tag{16}
$$

On the other hand, by (15),

$$
\begin{aligned}
P\left[Y_1 \geq \frac{3m}{2}, \frac{\sum_{i=2}^{m} Y_i}{m-1} \geq -\frac{1}{2}\right] &= P\left[Y_1 \geq \frac{3m}{2}\right] - P\left[Y_1 \geq \frac{3m}{2}, \frac{\sum_{i=2}^{m} Y_i}{m-1} \leq -\frac{1}{2}\right] \\
&\geq P\left[Y_1 \geq \frac{3m}{2}\right] - P\left[\frac{\sum_{i=2}^{m} Y_i}{m-1} \leq -\frac{1}{2}\right] \\
&\geq P\left[Y_1 \geq \frac{3m}{2}\right] - e^{-(m-1)\beta(-\frac{1}{2})}. \qquad (17)
\end{aligned}
$$

Combining (16) and (17) gives

$$
P\left[Y_1 \geq \frac{3m}{2}\right] \leq e^{-m\alpha(1)} + e^{-(m-1)\beta(-\frac{1}{2})}.
$$

Consequently, for any $\lambda \in [0, 2\min\{\alpha(1), \beta(-\frac{1}{2})\}/3)$,

$$\begin{aligned} & \sum_{m=1}^{\infty} \exp\left[\frac{3m\lambda}{2}\right] P\left[Y_1 \geq \frac{3m}{2}\right] \\ \leq & \sum_{m=1}^{\infty} \left(\exp\left[-\left(\alpha(1) - \frac{3\lambda}{2}\right) m\right] + \exp\left[\beta(1) - \left(\beta(1) - \frac{3\lambda}{2}\right) m\right]\right) < \infty, \end{aligned}$$

implying

$$E\{\exp[\lambda Y_1]\} < \infty$$

for any $\lambda \in [0, 2\min\{\alpha(1), \beta(-\frac{1}{2})\}/3)$. Likewise, it can be shown that

$$E\{\exp[\lambda Y_1]\} < \infty$$

for any $\lambda \in [0, -2\min\{\alpha(1), \beta(-\frac{1}{2})\}/3)$, completing the necessity portion of the proof. □

# 3 Large Deviation Results for Quantiles

As before, $F(\cdot)$ will denote the (unknown) c.d.f. of the output random variable $Y$. In this section, we will focus on the single-sample direct-simulation estimator $\widehat{\xi}(n)$. We will establish large deviation results of $\widehat{\xi}(n)$ for the cases in which the c.d.f. is continuous everywhere and where discontinuities are allowed.

## 3.1 Continuous Distributions

In this section, we are in a position to show that $\widehat{\xi}(n)$ converges to $\xi_r$ exponentially fast in probability as $n$ goes to infinity. The following is our main result.

**Theorem 3.1:** If the distribution function $F(\cdot)$ is strictly increasing and $\{Y_n, n \geq 1\}$ are negatively dependent, then

$$P\left[\left|\widehat{\xi}(n) - \xi_r\right| \geq \varepsilon\right] \leq e^{-n\Delta_+(\varepsilon,n)} + e^{-n\Delta_-(\varepsilon,n)}, \forall \varepsilon > 0, \tag{18}$$

where

$$\begin{aligned} \Delta_+(\varepsilon, n) &= \exp\left[-n \sup_{-\infty<\lambda\leq 0}\left(\lambda r - \frac{\log E\{\exp[\lambda \sum_{i=1}^{n} I(Y_i \leq \xi_r + \varepsilon)]\}}{n}\right)\right], \\ \Delta_-(\varepsilon, n) &= \exp\left[-n \sup_{0\leq\lambda<\infty}\left(\lambda r - \frac{\log E\{\exp[\lambda \sum_{i=1}^{n} I(Y_i \leq \xi_r - \varepsilon)]\}}{n}\right)\right]. \end{aligned}$$

And, moreover, the rate is enhanced by negatively dependence in the sense that

$$\Delta_+(\varepsilon, n) \geq \sup_{-\infty<\lambda\leq 0} (\lambda r - \log E\{\exp[\lambda I(Y \leq \xi_r + \varepsilon)]\}) > 0, \tag{19}$$

$$\Delta_+(\varepsilon, n) \geq \sup_{0\leq\lambda<\infty} (\lambda r - \log E\{\exp[\lambda I(Y \leq \xi_r - \varepsilon)]\}) > 0, \tag{20}$$

where the right-hand "sup" quantiles are the rates for i.i.d. samples.

**Proof:** From the definition of $\widehat{\xi}$,

$$P\left[\widehat{\xi}(n) - \xi_r \leq -\varepsilon\right] = P\left[F_n^{-1}(r) \leq \xi_r - \varepsilon\right]$$

$$
\begin{aligned}
&= P\left[F_n(\xi_r - \varepsilon) \geq r\right] \\
&= P\left[\frac{\sum_{i=1}^n I(Y_i \leq \xi_r - \varepsilon)}{n} \geq r\right] \\
&\leq \exp\left[-n \sup_{0 \leq \lambda < \infty}\left(\lambda r - \frac{\log E\{\exp[\lambda \sum_{i=1}^n I(Y_i \leq \xi_r - \varepsilon)]\}}{n}\right)\right], \qquad (21)
\end{aligned}
$$

the last inequality following from Lemma 2.3.

In exactly the same manner, we can show

$$
P\left[\widehat{\xi}(n) - \xi_r \geq \varepsilon\right] \leq \exp\left[-n \sup_{-\infty < \lambda \leq 0}\left(\lambda r - \frac{\log E\{\exp[\lambda \sum_{i=1}^n I(Y_i \leq \xi_r + \varepsilon)]\}}{n}\right)\right]. \qquad (22)
$$

Consequently, by combining (21) and (22),

$$
\begin{aligned}
P\left[\left|\widehat{\xi}(n) - \xi_r\right| \geq \varepsilon\right] &= P\left[\widehat{\xi}(n) - \xi_r \geq \varepsilon\right] + P\left[\widehat{\xi}(n) - \xi_r \geq \varepsilon\right] \\
&\leq \exp\left[-n \sup_{-\infty < \lambda \leq 0}\left(\lambda r - \frac{\log E\{\exp[\lambda \sum_{i=1}^n I(Y_i \leq \xi_r + \varepsilon)]\}}{n}\right)\right] \\
&\quad + \exp\left[-n \sup_{0 \leq \lambda < \infty}\left(\lambda r - \frac{\log E\{\exp[\lambda \sum_{i=1}^n I(Y_i \leq \xi_r - \varepsilon)]\}}{n}\right)\right],
\end{aligned}
$$

completing the proof of (18).

Next, noticing that the indicator functions $I(Y_i \leq \xi_r - \varepsilon)$ and $I(Y_i \leq \xi_r - \varepsilon)$ are decreasing functions with respect to $X_i$ for each $i$, we have as in Lemma 2.1, by Lemma 2.2,

$$
\sup_{-\infty < \lambda \leq 0}\left(\lambda r - \frac{\log E\{\exp[\lambda \sum_{i=1}^n I(Y_i \leq \xi_r + \varepsilon)]\}}{n}\right) \geq \sup_{-\infty < \lambda \leq 0}\left(\lambda r - \log E\{\exp[\lambda I(Y \geq \xi_r + \varepsilon)]\}\right) > 0,
$$

$$
\sup_{0 \leq \lambda < \infty}\left(\lambda r - \frac{\log E\{\exp[\lambda \sum_{i=1}^n I(Y_i \leq \xi_r - \varepsilon)]\}}{n}\right) \geq \sup_{0 \leq \lambda < \infty}\left(\lambda r - \log E\{\exp[\lambda I(Y \leq \xi_r - \varepsilon)]\}\right) > 0,
$$

the positiveness following from that the distribution function $F(\cdot)$ is strictly increasing, so $E[I(Y \leq \xi_r - \varepsilon)] = F(\xi_r - \varepsilon) < F(\xi_r) = r$ and $E[I(Y \leq \xi_r + \varepsilon)] = F(\xi_r + \varepsilon) > F(\xi_r) = r$. □

**Remark**: In order to estimate the convergence rate, the multiple-sample estimator $\bar{\xi}(k, n)$ defined by (3) can be taken to be the true quantile. Given $\varepsilon > 0$, let $\#\{i : \widehat{\xi}^{(i)}(m) \notin [\xi_r - \varepsilon, \xi_r + \varepsilon], i = 1, ..., k\}$ denote the number of $\widehat{\xi}^{(i)}(m)$'s which are not in the interval $[\xi_r - \varepsilon, \xi_r + \varepsilon]$. Then, the convergence rate corresponding to the interval $[\bar{\xi}(k, n) - \varepsilon, \bar{\xi}(k, n) + \varepsilon]$ can be estimated by

$$
\frac{-\log\left[\#\{i : \widehat{\xi}^{(i)}(m) \notin [\bar{\xi}(k, n) - \varepsilon, \bar{\xi}(k, n) + \varepsilon], i = 1, ..., k\}/k\right]}{m}.
$$

## 3.2 Distributions With Discontinuities

In the last subsection, we investigated the asymptotic behavior of quantiles estimation for the random variable with continuous distribution functions. In financial industrial there is a growing necessity to deal with random variables with discontinuous distribution. Examples are portfolios of non-traded loans (purely discrete distributions) or portfolios containing derivatives (mixtures of continuous and discrete distributions). A random variable with a discrete distribution has a non-unique quantile. It is well known that in case of a non-unique quantile (i.e., $a_r = \inf\{x|F(x) = r\} \neq b_r = \sup\{x|F(x) = r\}$), the quantity $Y_{\lfloor nr \rfloor, n}$ does not converge to $\xi_r$. This follows immediately from the following result given by Feldman and Tucker (1966).

**Lemma 3.1:** Suppose $a_r = \inf\{x|F(x) = r\} \neq b_r = \sup\{x|F(x) = r\}$). Then $Y_{\lfloor nr \rfloor, n}$ obeys the oscillatory effect with respect to the interval $[a_r, b_r]$, i.e., $P[Y_{\lfloor nr \rfloor, n} \leq a_r$ i.o.]=$P[Y_{\lfloor nr \rfloor, n} \geq b_r$ i.o.] $= 1$, where i.o. means "infinitely often".

Feldman and Tucker (1966) also designed consistent estimates of a quantile of a distribution function when the quantile is not unique. Their results are summarized by next two propositions.

**Lemma 3.2:** If for some selected $K > 0, \delta > 0$, the sequence of integers $\{L(n)\}$ satisfies (i) $0 < \lfloor nr \rfloor - L(n) \leq Kn^{\frac{1}{2}+\varepsilon}$ for some $\varepsilon, 0 < \varepsilon < \frac{1}{2}$, and (ii) $\lfloor nr \rfloor - L(n) \geq (1+\delta)(2n \log\log n/2)^{\frac{1}{2}}$, then $Y_{L(n),n} \to \xi_r$ a.s.

**Lemma 3.3:** A necessary and sufficient consition that $Y_{L(n),n} \to \xi_r$ in probability is that (i) $L(n)/n \to r$ as $n \to \infty$ and (ii) $n^{\frac{1}{2}}(r - L(n)/n) \to \infty$ as $n \to \infty$.

Motivated by the above lemmas, we introduce the estimator

$$\widetilde{\xi}(n) = F_n^{-1}\left(\frac{L(n)}{n}\right) = Y_{L(n),n},$$

where

$$L(n) = \lfloor nr - (1+\delta)(2n \log\log n/2)^{\frac{1}{2}} \rfloor.$$

**Theorem 3.2:** If $\{Y_n, n \geq 1\}$ are negatively dependent, then

$$P\left[\left|\widetilde{\xi}(n) - \xi_r\right| \geq \varepsilon\right] \leq e^{-n\Delta_+(\varepsilon,n)} + e^{-n\Delta_-(\varepsilon,n)}, \forall \varepsilon > 0, \tag{23}$$

where

$$\begin{aligned}
\Delta_+(\varepsilon, n) &= \exp\left[-n \sup_{-\infty < \lambda \leq \varepsilon}\left(\lambda\frac{L(n)}{n} - \frac{\log E\{\exp[\lambda \sum_{i=1}^n I(Y_i \leq \xi_r + \varepsilon)]\}}{n}\right)\right], \\
\Delta_-(\varepsilon, n) &= \exp\left[-n \sup_{0 \leq \lambda < \infty}\left(\lambda\frac{L(n)}{n} - \frac{\log E\{\exp[\lambda \sum_{i=1}^n I(Y_i \leq \xi_r - \varepsilon)]\}}{n}\right)\right].
\end{aligned}$$

And, moreover, the rate is enhanced by negatively dependence in the sense that

$$\Delta_+(\varepsilon, n) \geq \Delta_+^0(\varepsilon, n) = \sup_{-\infty < \lambda \leq 0}\left(\lambda\frac{L(n)}{n} - \log E\{\exp[\lambda I(Y \leq \xi_r + \varepsilon)]\}\right), \tag{24}$$

$$\Delta_+(\varepsilon, n) \geq \Delta_-^0(\varepsilon, n) = \sup_{0 \leq \lambda < \infty}\left(\lambda\frac{L(n)}{n} - \log E\{\exp[\lambda I(Y \leq \xi_r - \varepsilon)]\}\right), \tag{25}$$

and

$$\Delta_+^0(\varepsilon, n) > 0, \tag{26}$$

$$\Lambda_-^0(\varepsilon, n) > 0, \tag{27}$$

for $n$ satisfying

$$r - F(\xi_r - \varepsilon) - \frac{(1+\delta)(2n \log\log n/2)^{\frac{1}{2}} + 1}{n} > 0. \tag{28}$$

**Proof**. As in Theorem 3.1, (23), (24) and (25) can be proved in an analogous manner. It remains to show (26) and (27). Like (19) and (20), it suffices to show

$$\frac{L(n)}{n} - F(\xi_r + \varepsilon) < 0$$

and

$$\frac{L(n)}{n} - F(\xi_r - \varepsilon) > 0.$$

From the definition of $L(n)$,

$$\begin{aligned}\frac{L(n)}{n} - F(\xi_r + \varepsilon) &\le \frac{nr - (1+\delta)(2n\log\log n/2)^{\frac{1}{2}}}{n} - F(\xi_r + \varepsilon) \\ &= r - F(\xi_r + \varepsilon) - (1+\delta)(2n\log\log n/2)^{\frac{1}{2}} < 0,\end{aligned}$$

where the last inequality follows from that $r \le F(\xi_r + \varepsilon)$, proving (26).

We now turn to showing (27). Noticing that $r - F(\xi_r - \varepsilon) > 0$,

$$\begin{aligned}\frac{L(n)}{n} - F(\xi_r - \varepsilon) &> \frac{nr - (1+\delta)(2n\log\log n/2)^{\frac{1}{2}} - 1}{n} - F(\xi_r - \varepsilon) \\ &= r - F(\xi_r - \varepsilon) - \frac{(1+\delta)(2n\log\log n/2)^{\frac{1}{2}} + 1}{n} > 0\end{aligned}$$

for $n$ satisfying (28). □

# 4 New Quantiles Estimate

Now we consider the output random variable as a function of input random variables. Specifically, $Y = h(X_1, ..., X_d)$, where $X_1, ..., X_d$ are $d$ independent random variables with respective c.d.f.'s $G_1, ..., G_d$. In the last section, it was shown that the convergence rate of single-sample direct-simulation estimator $\widehat{\xi}(n)$ is exponential, and this convergence can be accelerated by Latin Hypercube Sampling. But it is still not clear how much improvement can be achieved over independent sampling. In this section, we will investigate this important issue by defining a new quantile estimator, which is defined by modifying the empirical c.d.f. of (2) to the following:

$$F_n(x) = \frac{1}{n^d} \sum_{i_d=1}^{n} \cdots \sum_{i_1=1}^{n} I(h(G_1^{-1}(U_1^{(i_1)}), ..., G_d^{-1}(U_d^{(i_d)})) \le x) \tag{29}$$

where $G_k^{-1}(\cdot)$ is the inverse function of $G_k(\cdot)$.

Note that the output sample size in the empirical distribution function (29) is $n^d$ although only $nd$ input samples are generated. The simulation for estimating the quantile $\xi_r$ by using the emprical distribution function (29) proceeds as follows:

**Step 1**: Generate $d$ independent sequences $\{U_1^{(i_1)}, i_1 = 1, ..., n\}, ..., \{U_d^{(i_d)}, i_d = 1, ..., n\}$ by Latin Hypercube Sampling;

**Step 2**: Generate $d$-dimensional random vectors $\{(G_1^{-1}(U_1^{(i_1)}), ..., G_d^{-1}(U_d^{(i_d)})), i_1 = 1, ..., n, ..., i_d = 1, ..., n\}$ and then get $n^d$ random variables $\{h(G_1^{-1}(U_1^{(i_1)}), ..., G_d^{-1}(U_d^{(i_d)}), i_1 = 1, ..., n, ..., i_d = 1, ..., n\}$;

**Step 3**: Sort out the $\lfloor n^d r \rfloor$th order statistics of sequence $\{h(G_1^{-1}(U_1^{(i_1)}), ..., G_d^{-1}(U_d^{(i_d)}), i_1 = 1, ..., n, ..., i_d = 1, ..., n\}$, giving the estimator of $\xi_r$.

Rather than (29), the estimator of Avramidis and Wilson (1998) uses the empirical distribution function given by (2), which can also be written as

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} I(h(G_1^{-1}(U_1^{(i)}), ..., G_d^{-1}(U_d^{(i)})) \le x), \tag{30}$$

where $\{U_1^{(i)}, i = 1, ..., n\}, ..., \{U_d^{(i)}, i = 1, ..., n\}$ are $d$ independent sequences generated by Latin Hypercube Sampling. Note that to generate an output sample size of $n^d$ requires the generation of $dn$ random numbers using (29), versus $dn^d$ random numbers required for (30). This difference can be enormous for high dimensions and/or large sample sizes.

A real-valued function $f(x_1, ..., x_d)$ of $d$ arguments will be called *monotonic* if, considered as a function of each of the individual coordinates (with all other coordinates held fixed), the function is monotone in the ordinary sense, i.e., as a function of a single variable.

**Theorem 4.1:** Suppose the function $h(x_1, x_2, ..., x_d)$ is monotone and

$$\min(F(\xi_r + \varepsilon) - r, r - F(\xi_r - \varepsilon)) = c > 0.$$

Then, for $\varepsilon > 0$,

$$P[F_n^{-1}(r) \in (\xi_r - \varepsilon, \xi_r + \varepsilon)] = 1$$

for $n > n_0 = \frac{2d}{c}$.

We give a queueing example to illustrate the monotonic condition.

**Example:** $G/G/1$ Queue

Consider the standard first-come, first-served, single-server queue. We verify that the waiting time $W_n$ of the $n$th customer and average waiting time $\overline{W}_n$ over the first $n$ customers are monotone functions of the service and interarrival times.

Let $a_1$ be the interval time of the first customer, $a_n$ $(n \geq 2)$ be the interarrival time between the $(n-1)$th customer and the $n$th customer, and $s_n$ $(n \geq 1)$ be the service time of $n$th customer. And, moreover, let $D_n$ denote the delay time of $n$th customer in queue. Assume $\{a_n, n \geq 1\}$ and $\{s_n, n \geq 1\}$ are two independent sequences of independent and identically distrbuted random variables with exponenetial distribution. Then we have the following relations:

$$W_n = D_n + s_n, \; n = 1, 2, ...,$$

$$D_n = (D_{n-1} + s_{n-1} - a_n)^+, \; n = 2, 3, ...,$$

where $(x)^+ = \max\{0, x\}$. Particularly, $W_1 = s_1$ and $D_1 = 0$.

By induction, it is not hard to show that $D_n$ is non-decreasing function of $s_i$, $i = 1, 2, ..., n-1$ and non-increasing function of $a_i$, $i = 1, 2, ..., n$ and so are $W_n$ and $\overline{W}_n$, arriving at our conclusions. □

Before proving Theorem 4.1, we give a lemma. To this end, define

$$\widetilde{h}_1(y : x_1, x_2, ..., x_d) = I(h(x_1, x_2, ..., x_d) \leq y),$$

$$\widetilde{h}_j(y : x_j, x_{j+1}, ..., x_d) = P[h(X_1, ..., X_{j-1}, x_j, x_{j+1}, ..., x_d) \leq y], \text{ for } j = 2, ..., d,$$

i.e., for $j = 2, ..., d$, $\widetilde{h}_j(y : x_j, x_{j+1}, ..., x_d)$ is the conditional distribution function of $h(X_1, ..., X_{j-1}, x_j, x_{j+1}, ..., x_d)$ given $X_l = x_l, l = j, ..., d$ and

$$\widetilde{h}_{j+1}(y : x_{j+1}, ..., x_d) = \int \widetilde{h}_j(y : x_j, x_{j+1}, ..., x_d) dP[X_j \leq x_j]$$

since $X_1, ..., X_d$ are mutually independent.

**Lemma 4.1:** Suppose the assumption of Theorem 4.1 holds. Then, when $n > n_0 = \frac{2d}{c}$,

$$P\left[\sum_{i_j=1}^{n} \widetilde{h}_j\left(\xi_r - \varepsilon : G_j^{-1}(U_j^{(i_j)}), x_{j+1}, ..., x_d\right) \geq \widetilde{h}_{j+1}(\xi_r - \varepsilon, x_{j+1}, ..., x_d) + \frac{c}{2d}\right] = 0,$$

$$P\left[\sum_{i_j=1}^{n} \widetilde{h}_j\left(\xi_r + \varepsilon : G_j^{-1}(U_j^{(i_j)}), x_{j+1}, ..., x_d\right) \leq \widetilde{h}_{j+1}(\xi_r + \varepsilon : x_{j+1}, ..., x_d) - \frac{c}{2d}\right] = 0$$

for $j = 1, ..., d$,and $x_{j+1}, ..., x_d$.

**Proof:** See Appendix.

**Proof of Theorem 4.1**: From Lemma 4.1, given $U_k^{(i_k)} = u_k^{(i_k)}, i_k = 1, ..., n, k = 2, ..., d$, when $n > n_0$,

$$\begin{aligned}
&\frac{1}{n}\sum_{i_1=1}^{n} I(h(G_1^{-1}(U_1^{(i_1)}), G_2^{-1}(u_2^{(i_2)}), ..., G_d^{-1}(u_d^{(i_d)})) \leq \xi_r - \varepsilon) \\
&\qquad \leq \widetilde{h}_2\left(\xi_r - \varepsilon, G_2^{-1}(u_2^{(i_2)}), ..., G_d^{-1}(u_d^{(i_d)})\right) + \frac{c}{2d}
\end{aligned}$$

with probability one and therefore, when $n > n_0$,

$$\begin{aligned}
P[F_n^{-1}(r) \leq \xi_r - \varepsilon] &= P\left[\frac{1}{n^d}\sum_{i_d=1}^{n}\cdots\sum_{i_1=1}^{n} I(h(G_1^{-1}(U_1^{(i_1)}), ..., G_d^{-1}(U_d^{(i_d)})) \leq \xi_r - \varepsilon) \geq r\right] \\
&= \int ... \int P\left[\frac{1}{n^d}\sum_{i_d=1}^{n}\cdots\sum_{i_1=1}^{n} I(h(G_1^{-1}(U_1^{(i_1)}), G_2^{-1}(u_2^{(i_2)}), ..., G_d^{-1}(u_d^{(i_d)})) \leq \xi_r - \varepsilon) \geq r\right. \\
&\qquad \left| U_k^{(i_k)} = u_k^{(i_k)}, i_k = 1, ..., n, k = 2, ..., d\right] d\prod_{k=2}^{d} P\left[U_k^{(i_k)} \leq u_k^{(i_k)}, i_k = 1, ..., n\right] \\
&= \int ... \int P\left[\frac{1}{n^d}\sum_{i_d=1}^{n}\cdots\sum_{i_1=1}^{n} I(h(G_1^{-1}(U_1^{(i_1)}), G_2^{-1}(u_2^{(i_2)}), ..., G_d^{-1}(u_d^{(i_d)})) \leq \xi_r - \varepsilon) \geq r\right] \\
&\qquad d\prod_{k=2}^{d} P\left[G_k^{-1}(U_k^{(i_k)}) \leq u_k^{(i_k)}, i_k = 1, ..., n\right] \\
&\leq P\left[\frac{1}{n^{d-1}}\sum_{i_d=1}^{n}\cdots\sum_{i_2=1}^{n} \widetilde{h}_2\left(\xi_r - \varepsilon, G_2^{-1}(u_2^{(i_2)}), ..., G_d^{-1}(u_d^{(i_d)})\right) + \frac{c}{2d} \geq r\right] \\
&\qquad d\prod_{k=2}^{d} P\left[G_k^{-1}(U_k^{(i_k)}) \leq u_k^{(i_k)}, i_k = 1, ..., n\right] \\
&\leq P\left[\frac{1}{n^{d-1}}\sum_{i_d=1}^{n}\cdots\sum_{i_2=1}^{n} \widetilde{h}_2\left(\xi_r - \varepsilon, G_2^{-1}(U_2^{(i_2)}), ..., G_d^{-1}(U_d^{(i_d)})\right) + \frac{c}{2d} \geq r\right].
\end{aligned}$$

By exactly the same technique, we have the following result:

$$P\left[\frac{1}{n^{d-1}}\sum_{i_d=1}^{n}\cdots\sum_{i_2=1}^{n} \widetilde{h}_2\left(\xi_r - \varepsilon, G_2^{-1}(U_2^{(i_2)}), ..., G_d^{-1}(U_d^{(i_d)})\right) + \frac{c}{2d} \geq r\right]$$

$$
\begin{aligned}
&\le P\left[\frac{1}{n^{d-2}}\sum_{i_d=1}^{n}\cdots\sum_{i_3=1}^{n}\widetilde{h}_3\left(\xi_r-\varepsilon, G_3^{-1}(U_3^{(i_3)}),...,G_d^{-1}(U_d^{(i_d)})\right)+\frac{2c}{2d}\ge r\right] \\
&\cdots \\
&= P\left[\frac{1}{n}\sum_{i_d=1}^{n}\widetilde{h}_d\left(\xi_r-\varepsilon : G_d^{-1}(U_d^{(i_d)})\right)+\frac{(d-1)c}{2d}\ge r\right] \\
&= P\left[F\left(\xi_r-\varepsilon\right)\ge r-\frac{c}{2}\right]=0.
\end{aligned}
$$

Likewise, we can show that when $n > \frac{2d}{c}$,

$$P[F_n^{-1}(r) \ge \xi_r + \varepsilon] = 0,$$

completing the proof of Theorem 4.1. □

In particular, when $d = 1$, Theorem 4.1 shows that by using Latin Hypercube Sampling, the convergence rate of $\widehat{\xi}(n)$ is infinite when the sample size is large enough, i.e.,

$$\sup_{0\le\lambda<\infty}\left(\lambda r-\frac{\log E\{\exp[\lambda\sum_{i=1}^{n}I(Y_i\le\xi_r-\varepsilon)]\}}{n}\right)=\infty,$$

and

$$\sup_{-\infty<\lambda\le 0}\left(\lambda r-\frac{\log E\{\exp[\lambda\sum_{i=1}^{n}I(Y_i\le\xi_r+\varepsilon)]\}}{n}\right)=\infty,$$

for $n > (\min(r - F(\xi_r - \varepsilon), F(\xi_r + \varepsilon) - r))^{-1}$. In other words, the probability that the estimate of a quantile is in a neighborhood of the true quantile (the *coverage probability*) is one for sufficiently large sample size. On the other hand, the next example demonstrates that the infinite convergence rate cannot be obtained for the case of independent sampling, i.e., an infinite sample size is required for 100% coverage.

**Example:** Assume $\{Y_n, n \ge 1\}$ are i.i.d. with distribution $F(x) = 1 - e^{-x}$ if $x \ge 0$; and $F(x) = 0$ if $x \le 0$. We will show that convergence rates

$$\sup_{0\le\lambda<\infty}\left(\lambda r-\log E\{\exp[\lambda I(Y\le\xi_r-\varepsilon)]\}\right)<\infty$$

and

$$\sup_{-\infty<\lambda\le 0}\left(\lambda r-\log E\{\exp[\lambda I(Y\ge\xi_r+\varepsilon)]\}\right)<\infty.$$

Simple algebra gives

$$\sup_{0\le\lambda<\infty}\left(\lambda r-\log E\{\exp[\lambda I(Y\le\xi_r-\varepsilon)]\}\right)=r\ln\left[\frac{r\exp[-(\xi_r-\varepsilon)]}{(1-r)(1-\exp[-(\xi_r-\varepsilon)])}\right]-\ln\left[\frac{\exp[-(\xi_r-\varepsilon)]}{1-r}\right],$$

which is a positive and finite number.

Likewise,

$$\sup_{-\infty<\lambda\le 0}\left(\lambda r-\log E\{\exp[\lambda I(Y\ge\xi_r+\varepsilon)]\}\right)=r\ln\left[\frac{r\exp[-(\xi_r+\varepsilon)]}{(1-r)(1-\exp[-(\xi_r+\varepsilon)])}\right]-\ln\left[\frac{\exp[-(\xi_r-\varepsilon)]}{1-r}\right],$$

which is a positive and finite number, arriving at our conclusion. □

So far, we have not analyzed estimator variance. It is well known that the variance of an estimate of an average provided by Monte Carlo sampling decreases in proportion to the inverse of the square root of the number of trials. That is, to improve the accuracy of the estimate by a factor of 10, the number of trials must be increased by a factor 100. Avramidis and Wilson (1998) established the same result for Latin Hypercube Sampling. When $d = 1$, the following result represents a huge improvement, as a $1/n$ (vs. $1/\sqrt{n}$) convergence rate is established. Furthermore, the bound is not asymptotic but holds for all $n$.

**Corollary 4.1:** If $F(\cdot)$ is differentiable in a neighborhood $[\xi_r - \delta, \xi_r + \delta]$ of $\xi_r$ with $\inf_{x\in[\xi_r-\delta,\xi_r+\delta]} F'(x) = c_0 > 0$, then

$$\left|E\left[\widehat{\xi}(n)\right] - \xi_r\right| \leq \frac{3}{nc_0}$$

and

$$Var\left[\widehat{\xi}(n)\right] \leq \left(\frac{3}{nc_0}\right)^2.$$

**Proof.** Note that, by taking $\varepsilon = \frac{3}{nc_0} \leq \delta$,

$$\frac{2}{c} = \frac{2}{\min(F(\xi_r + \varepsilon) - r, r - F(\xi_r - \varepsilon))} \leq \frac{2}{c_0\varepsilon} = \frac{2n}{3} < n.$$

Thus, $n$ and $\varepsilon$ satisfy Theorem 4.1 and therefore, with probability one,

$$\left|\widehat{\xi}(n) - \xi_r\right| \leq \varepsilon = \frac{3}{nc_0}$$

implying the two conclusions of this theorem. □

# 5 Application: Value-at-Risk Estimation

In this section, we illustrate the application of our quantile estimate techniques to the simulation of Value-at-Risk (VaR) in a classic financial model. Risk exposures are typically quantified in terms of VaR. Formally, *VaR* measures the worst expected loss over a given time interval under normal market conditions at a given confidence level. *VaR* provides users with a summary measure of market risk. For instance, a bank might say that the daily *VaR* of its trading portfolio is \$35 million at the 99 percent confidence level. In other words, there is only 1 chance in a 100, under normal market conditions, for a loss greater than \$35 million to occur. This single number summarizes the bank's exposure to market risk as well as the probability of an adverse move. In the terminology of statistics, *VaR* is nothing more than a quantile of a portfolio's potential profit and loss process over a given time period.

Suppose that an institution has an exposure to an asset, $S_t$, whose process is governed by the following stochastic differential equation:

$$\frac{dS_t}{S_t} = \mu dt + \sigma dz_t,$$

where $\mu$ and $\sigma$ are the drift and the diffusion, respectively, of the asset value, and $z_t$ is a standard Brownian motion. One can regard this asset either as a single asset, or as a portfolio of assets like, for example, the S&P index, or a portfolio of the institution's currency exposures. As such, the analysis is better suited to an institution concerned with their exposure to commodity prices, equities, or exchange rates.

The institution is concerned about its exposure to the asset over the next $\tau$ periods, that is, the institution is concerned about the loss at the $r\%$ level of the distribution of the institution's exposure $S_{t+\tau}$, which is called as $\text{VaR}_{t+\tau}$. Given the lognormality of $S_t$, we

$$P\left[S_t - S_t \exp[(\mu - \frac{1}{2}\sigma^2)\tau + \sigma\sqrt{\tau}Z] \geq VaR_{t+\tau}\right] = r$$

that is

$$VaR_{t+\tau} = S_t - S_t \exp[(\mu - \frac{1}{2}\sigma^2)\tau + c(r)\sigma\sqrt{\tau}],$$

where $Z \sim N(0,1)$ is a standard normal random variable and $c(r)$ is the cut-off point of the cumulative distribution of a standard normal at the $r\%$ level. That is, the VaR at the $r\%$ level is the quantile at the $(1-r)\%$ level of process:

$$Y = S_t - S_t \exp\left[(\mu - \frac{1}{2}\sigma^2)\tau + \sigma\sqrt{\tau}Z\right]. \tag{31}$$

In our implementation, we use the Box-Muller method (see Law and Kelton 2000) to generate standard normal random variables. This method uses two independent random numbers $U_1$ and $U_2$. Consequently, (31) can be expressed as

$$\begin{aligned} Y &= S_t - S_t \exp[(\mu - \frac{1}{2}\sigma^2)\tau + \sigma\sqrt{\tau}\sqrt{-2\ln U_1}\cos(2\pi U_2)] \\ &= S_t - S_t \exp[(\mu - \frac{1}{2}\sigma^2)\tau + \sigma\sqrt{\tau}X_1 X_2] \\ &= h(X_1, X_2), \end{aligned}$$

where

$$X_1 = \sqrt{-2\ln U_1},\ X_2 = \cos(2\pi U_2).$$

Obviously, the function $h(X_1, X_2)$ satisfies the assumptions of Theorem 4.1, where the dimension $d$ in this problem is 2.

In our numerical experiments, we take $t = 0, \tau = 1$ day$= \frac{1}{250}$ year (since there are 250 business days within each year), $\mu = 0.2, \sigma = 0.2$, $\varepsilon = 1\%\xi_r$, and $S_0 = 1000$. Because our estimator "reuses" sample, it is not straightforward to compare with other estimators in literature. When we generate $nd$ input samples, we obtain $n^d$ output samples. Since $d = 2$ here, we will generate $n^2$ output samples for the other estimators, whether using Latin Hypercube Sampling or independent sampling. In the tables, LHS*, LHS and IND denote the quantile estimate generated by (29), the quantile estimate generated by (30), and the quantile estimate generated by using independent sampling, respectively. From the tables and graphs, it can be seen that LHS* significantly outperforms IND in all cases. For rare events, e.g., $r = 0.01$ and $r = 0.99$, LHS* also beats LHS by a significant margin, but the reverse is true for the median $r = 0.5$. The variance of LHS and IND are very close for the two extreme cases, although LHS clearly dominates IND in terms of coverage probability. However, the latter is a function of the choice of $\varepsilon$, so a smaller value of $\varepsilon$ would reduce the difference.

| $n$ | 50 | 100 | 200 |
|---|---|---|---|
| LHS* | 0.5649 | 0.1997 | 0.0693 |
| LHS | 0.8841 | 0.3955 | 0.2020 |
| IND | 0.8969 | 0.4066 | 0.2169 |

Table 1: Standard error estimates for $r$=0.01.

| $n$ | 50 | 100 | 200 |
|---|---|---|---|
| LHS* | 0.2796 | 0.0320 | 0.0116 |
| LHS | 0.0149 | 0.0029 | 0.00082 |
| IND | 0.2871 | 0.1601 | 0.0869 |

Table 2: Standard error estimates for $r$=0.5.

| $n$ | 50 | 100 | 200 |
|---|---|---|---|
| LHS* | 0.4800 | 0.1869 | 0.0698 |
| LHS | 0.8781 | 0.4876 | 0.2136 |
| IND | 0.9296 | 0.4950 | 0.2683 |

Table 3: Standard error estimates for $r$=0.99.

| $n$ | 50 | 100 | 200 |
|---|---|---|---|
| LHS* | 0.49 | 0.94 | 1 |
| LHS | 0.25 | 0.55 | 0.82 |
| IND | 0.08 | 0.16 | 0.16 |

Table 4: Coverage probability for $r$=0.01.

| $n$ | 50 | 100 | 200 |
|---|---|---|---|
| LHS* | 0.05 | 0.15 | 0.49 |
| LHS | 0.51 | 0.99 | 1 |
| IND | 0.01 | 0.02 | 0.01 |

Table 5: Coverage probability for $r$=0.5.

| $n$ | 50 | 100 | 200 |
|---|---|---|---|
| LHS* | 0.33 | 0.93 | 1 |
| LHS | 0.21 | 0.54 | 0.83 |
| IND | 0.07 | 0.09 | 0.22 |

Table 6: Coverage probability for $r$=0.99.

# 6 Conclusions

Traditionally, simulation efficiency is improved by reducing estimator variance, which is taken to be the variance of limiting normal distribution. This approach, however, suffers some drawbacks: first the variance is asymptotic, and second, it may be cumbersome to validate some of the assumptions underlying the central limit theorem. In this paper we analyze simulation efficiency from another point of view—the coverage probability obtained by applying the large deviations principle. We established large deviations results for quantile estimation, and the convergence rate is considered as a measure of estimator efficiency. It is shown that variance reduction techniques can be used to enhance the convergence rate.

Particularly and perhaps more importantly, we proposed a new sampling plan for estimating quantiles using the Latin Hypercube Sampling. Both theoretical and experimental results provide substantial evidence that our new sampling plan is orders of magnitude better than independent sampling in all cases and improves upon a previous estimator in simulating rare events.

Immediate work in progress includes investigating whether or not Corollary 4.1 holds for $d > 1$, and obtaining satisfying explanations for the empirical results involving the median estimator. Avenues for further research include applying some commonly used variance reduction techniques, e.g., control variates and antithetic sampling, to increase convergence rate, and to combine important sampling with our new sampling plan to improve the estimator accuracy. We would also like to investigate the use of quasi-Monte Carlo methods (e.g., Niederreiter (1992)) in place of, or in combination with, LHS.

# 7 Appendix

**Proof of Lemma 2.2**: We prove this result by using Lemma 2.1. To this end, let $g_i(\cdot) : R \to R$ be a non-negative monotone function, $i = 1, ..., n$ and they are concordant. It is easy to verify from the definition of negative dependence that $g_i(Y_i), i = 1, ..., n$, are negative dependent if $Y_i, i = 1, ..., n$, are negative dependent. Note that $X^{(j)} = (X_1^{(j)}, ..., X_n^{(j)}), j = 1, ..., d$, are mutually independent. Then by Lemma 2.1,

$$
\begin{aligned}
& E\left[\prod_{i=1}^{n} g_i(f_i(X_i^{(1)}, ..., X_i^{(d)}))\right] \\
= \; & E\left[E\left[\prod_{i=1}^{n} g_i(f_i(X_i^{(1)}, ..., X_i^{(d)}))\middle| X_i^{(2)}, ..., X_i^{(d)}, i = 1, ..., n\right]\right] \\
\leq \; & E\left[\prod_{i=1}^{n} E\left[g_i(f_i(X_i^{(1)}, ..., X_i^{(d)}))\middle| X_i^{(2)}, ..., X_i^{(d)}, i = 1, ..., n\right]\right] \\
= \; & E\left[E\left[\prod_{i=1}^{n} E\left[g_i(f_i(X_i^{(1)}, ..., X_i^{(d)}))\middle| X_i^{(2)}, ..., X_i^{(d)}, i = 1, ..., n\right]\middle| X_i^{(3)}, ..., X_i^{(d)}, i = 1, ..., n\right]\right] \\
\leq \; & E\left[\prod_{i=1}^{n} E\left[g_i(f_i(X_i^{(1)}, ..., X_i^{(d)}))\middle| X_i^{(3)}, ..., X_i^{(d)}, i = 1, ..., n\right]\right] \\
\leq \; & ... \\
\leq \; & \prod_{i=1}^{n} E\left[g_i(f_i(X_i^{(1)}, ..., X_i^{(d)}))\right].
\end{aligned}
$$

By taking $g_i(x) = I(x \leq x_i), i = 1, ..., n$ and $g_i(x) = I(x \leq x_i), i = 1, ..., n$ respectively, we arrive at the conclusion. $\square$

**Proof of Lemma 4.1:** Without loss of the generality, suppose the function $h(x_1, x_2, ..., x_d)$ is increasing with respect to each coordinate since, for example, the function $h(-x_1', x_2, ..., x_d)$ (where $x_1' = -x_1$) will be decreasing function of $x_1'$(with all other coordinates held fixed) if $h(x_1, x_2, ..., x_d)$ is increasing with respect to $x_1$. And, moreover, let $h_j^{-1}(x_1, ..., x_{j-1}, x_{j+1}, ..., x_d)$ denote the inverse function of $h(x_1, x_2, ..., x_d)$ which is considered as the function of $x_j$ (with all other coordinates held fixed).

Note that by definition, $\widetilde{h}_j(\xi_r - \varepsilon : x_j, x_{j+1}, ..., x_d)$, considered as a function of $x_j$ (with all other coordinates held fixed), is a non-increasing function.

Hence,

$$\begin{aligned} & P\left[\sum_{i_j=1}^{n} \widetilde{h}_j\left(\xi_r - \varepsilon : G_j^{-1}(U_j^{(i_j)}), x_{j+1}, ..., x_d\right) \geq \widetilde{h}_{j+1}(\xi_r - \varepsilon : x_{j+1}, ..., x_d) + \frac{c}{2d}\right] \\ = \; & P\left[\sum_{i_j=1}^{n} \widetilde{h}_j\left(\xi_r - \varepsilon : G_j^{-1}\left(\frac{\pi_j(i_j) - U^*_{i_j j}}{n}\right), x_{j+1}, ..., x_d\right) \geq \widetilde{h}_{j+1}(\xi_r - \varepsilon : x_{j+1}, ..., x_d) + \frac{c}{2d}\right] \\ \leq \; & P\left[\sum_{i_j=1}^{n} \widetilde{h}_j\left(\xi_r - \varepsilon : G_j^{-1}\left(\frac{\pi_j(i_j) - 1}{n}\right), x_{j+1}, ..., x_d\right) \geq \widetilde{h}_{j+1}(\xi_r - \varepsilon : x_{j+1}, ..., x_d) + \frac{c}{2d}\right] \end{aligned}$$

And moreover, by Lemma 2.3,

$$\begin{aligned} & P\left[\sum_{i_j=1}^{n} \widetilde{h}_j\left(\xi_r - \varepsilon : G_j^{-1}\left(\frac{\pi_j(i_j) - 1}{n}\right), x_{j+1}, ..., x_d\right) \geq \widetilde{h}_{j+1}(\xi_r - \varepsilon : x_{j+1}, ..., x_d) + \frac{c}{2d}\right] \\ \leq \; & \exp\left[-n \sup_{0 \leq \lambda < \infty}\left\{\lambda\left(\widetilde{h}_{j+1}(\xi_r - \varepsilon : x_{j+1}, ..., x_d) + \frac{c}{2d}\right) - \frac{\log M(\lambda)}{n}\right\}\right] \end{aligned}$$

where

$$M(\lambda) = E\left[\exp\left[\lambda \sum_{i_j=1}^{n} \widetilde{h}_j\left(\xi_r - \varepsilon : G_j^{-1}\left(\frac{\pi_j(i_j) - 1}{n}\right), x_{j+1}, ..., x_d\right)\right]\right].$$

Next we want to prove

$$\sup_{0 \leq \lambda < \infty}\left\{\lambda\left(\widetilde{h}_{j+1}(\xi_r - \varepsilon : x_{j+1}, ..., x_d) + \frac{c}{2d}\right) - \frac{\log M(\lambda)}{n}\right\} = \infty$$

when $n > \frac{2d}{c}$.

Since $\pi_j(i_j), i_j = 1, ..., n$, is the permutation of $1, ..., n$,

$$\begin{aligned} & \sum_{i_j=1}^{n} \widetilde{h}_j\left(\xi_r - \varepsilon : G_j^{-1}\left(\frac{\pi_j(i_j) - 1}{n}\right), x_{j+1}, ..., x_d\right) \\ = \; & \sum_{i_j=1}^{n} \widetilde{h}_j\left(\xi_r - \varepsilon : G_j^{-1}\left(\frac{j - 1}{n}\right), x_{j+1}, ..., x_d\right) \end{aligned}$$

implying

$$\begin{aligned} \frac{\log M(\lambda)}{n} &= \frac{\lambda \sum_{i_j=1}^{n} \widetilde{h}_j\left(\xi_r - \varepsilon : G_j^{-1}\left(\frac{j-1}{n}\right), x_{j+1}, ..., x_d\right)}{n} \\ &= \frac{\lambda \sum_{i_j=1}^{n} \widetilde{h}_j\left(\xi_r - \varepsilon : G_j^{-1}\left(\frac{\pi_j(i_j)-1}{n}\right), x_{j+1}, ..., x_d\right)}{n} \\ &= \lambda E\left[\widetilde{h}_j\left(\xi_r - \varepsilon : G_j^{-1}\left(\frac{\pi_j - 1}{n}\right), x_{j+1}, ..., x_d\right)\right]. \end{aligned}$$

Note that

$$
\begin{aligned}
&= \lambda E\left[\widetilde{h}_j\left(\xi_r-\varepsilon : G_j^{-1}\left(\frac{\pi_j-1}{n}\right), x_{j+1}, ..., x_d\right)\right] \\
&= \lambda P\left[h_j\left(X_1, ..., X_{j-1}, G_j^{-1}\left(\frac{\pi_j-1}{n}\right), x_{j+1}, ..., x_d\right) \leq \xi_r-\varepsilon\right] \\
&= \lambda P\left[G_j^{-1}\left(\frac{\pi_j-1}{n}\right) \leq h_j^{-1}\left(\xi_r-\varepsilon : X_1, ..., X_{j-1}, x_{j+1}, ..., x_d\right)\right] \\
&= \lambda P\left[\frac{\pi_j-1}{n} \leq G_j\left(h_j^{-1}\left(\xi_r-\varepsilon : X_1, ..., X_{j-1}, x_{j+1}, ..., x_d\right)\right)\right] \\
&\leq \lambda P\left[\frac{\pi_j-U_j^*}{n} \leq G_j\left(h_j^{-1}\left(\xi_r-\varepsilon : X_1, ..., X_{j-1}, x_{j+1}, ..., x_d\right)\right)+\frac{1}{n}\right] \\
&\leq \lambda P\left[\frac{\pi_j-U_j^*}{n} \leq G_j\left(h_j^{-1}\left(\xi_r-\varepsilon : X_1, ..., X_{j-1}, x_{j+1}, ..., x_d\right)\right)\right]+\frac{\lambda}{n} \\
&= \lambda P\left[h_j\left(X_1, ..., X_{j-1}, G_j^{-1}\left(\frac{\pi_j-U_j^*}{n}\right), x_{j+1}, ..., x_d\right) \leq \xi_r-\varepsilon\right]+\frac{\lambda}{n} \\
&= \lambda \widetilde{h}_{j+1}\left(\xi_r-\varepsilon : x_{j+1}, ..., x_d\right)+\frac{\lambda}{n}
\end{aligned}
$$

Hence,

$$
\begin{aligned}
& \sup_{0 \leq \lambda<\infty}\left\{\lambda\left(\widetilde{h}_{j+1}\left(\xi_r-\varepsilon : x_{j+1}, ..., x_d\right)+\frac{c}{2d}\right)-\frac{\log M(\lambda)}{n}\right\} \\
\geq\ & \sup_{0 \leq \lambda<\infty}\left[\lambda\left(\widetilde{h}_{j+1}\left(\xi_r-\varepsilon : x_{j+1}, ..., x_d\right)+\frac{c}{2d}\right)-\lambda \widetilde{h}_{j+1}\left(\xi_r-\varepsilon : x_{j+1}, ..., x_d\right)-\frac{\lambda}{n}\right] \\
\geq\ & \sup_{0 \leq \lambda<\infty}\left[\lambda\left(\frac{c}{2d}-\frac{1}{n}\right)\right]=\infty
\end{aligned}
$$

for $n > \frac{2d}{c}$. $\square$

## References


[1] Avramidis, A.N. and J.R. Wilson, "Correlation-induction techniques for estimating quantiles in simulation experiments," *Operations Research* 46 (1998), 574-591.

[2] Bucklew, J.A., *Large Deviations Techniques in Decision, Simulation, and Estimation*, Wiley, 1990.

[3] Dembo, A., and O. Zeitouni, *Large Deviations Techniques and Applications,* 2nd edition, Springer-Verlag, 1998.

[4] Deuschel, J.D., and D. Strook, *Large Deviations,* Academic Press, 1989.

[5] Feldman, D., and H.G. Tucker, "Estimation of non-unique quantiles," *Annals of Mathematical Statistics*, 37 (1966), 451-457.

[6] Fu, M.C., and X. Jin, "On the convergence rate of ordinal comparison of random variables," *IEEE Transactions on Automatic Control*, forthcoming, Dec. 2001.

[7] Hsu, J.C., and B.L. Nelson, "Control variates for quantile estimation," *Management Science*, 36 (1990), 835-851.

[8] Hesterberg, T.C., and B.L. Nelson, "Control variates for probability and quantile estimation," *Management Science*, 44 (1998), 1295-1312.

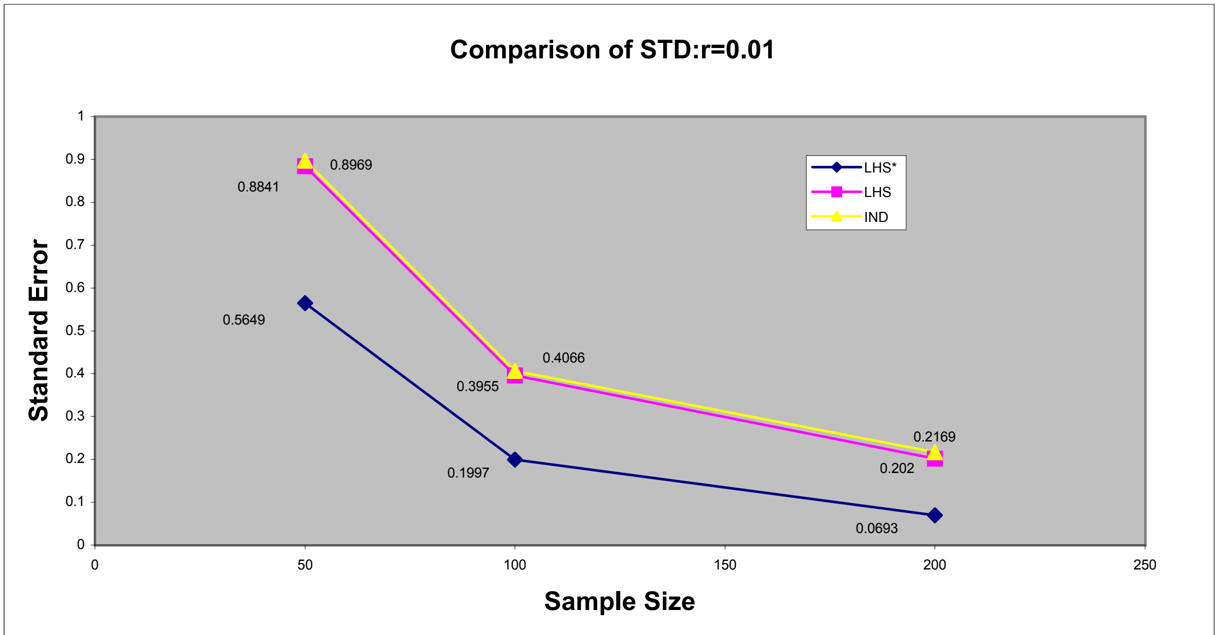
[9] Law A.M. and W.D. Kelton, 2000. *Simulation Modeling and Analysis*, 3rd edition, McGraw-Hill, New York.

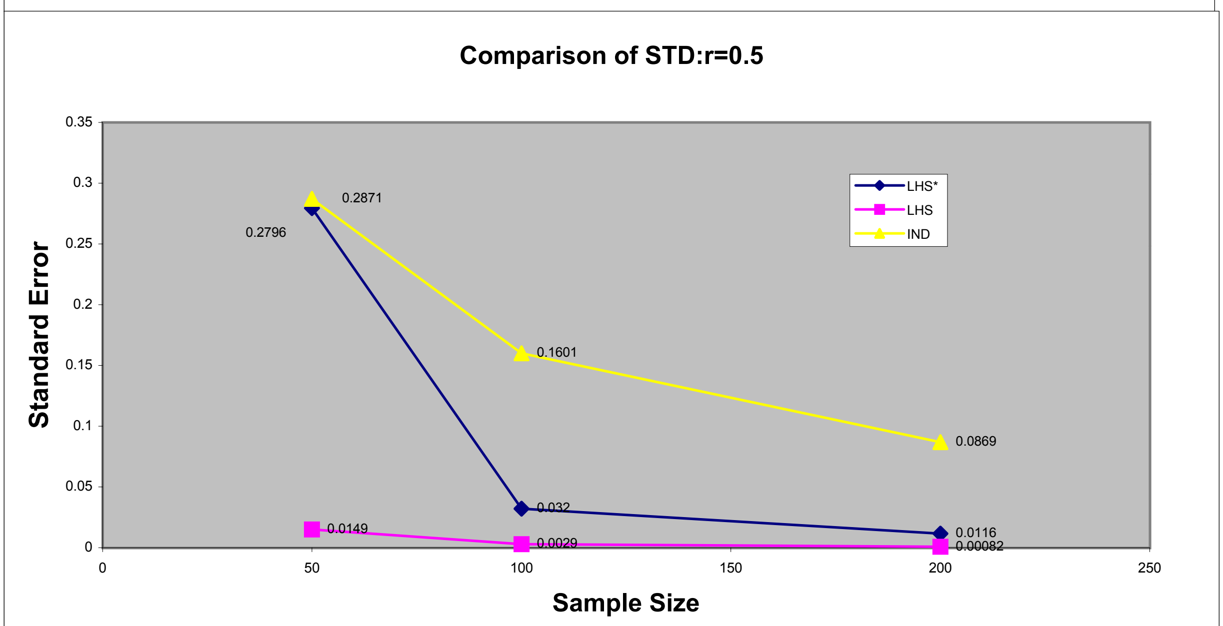
[10] Lehmann, E.L., "Some concepts of dependence," *Annals of Mathematical Statistics*, 37(1966), 1137-1153.

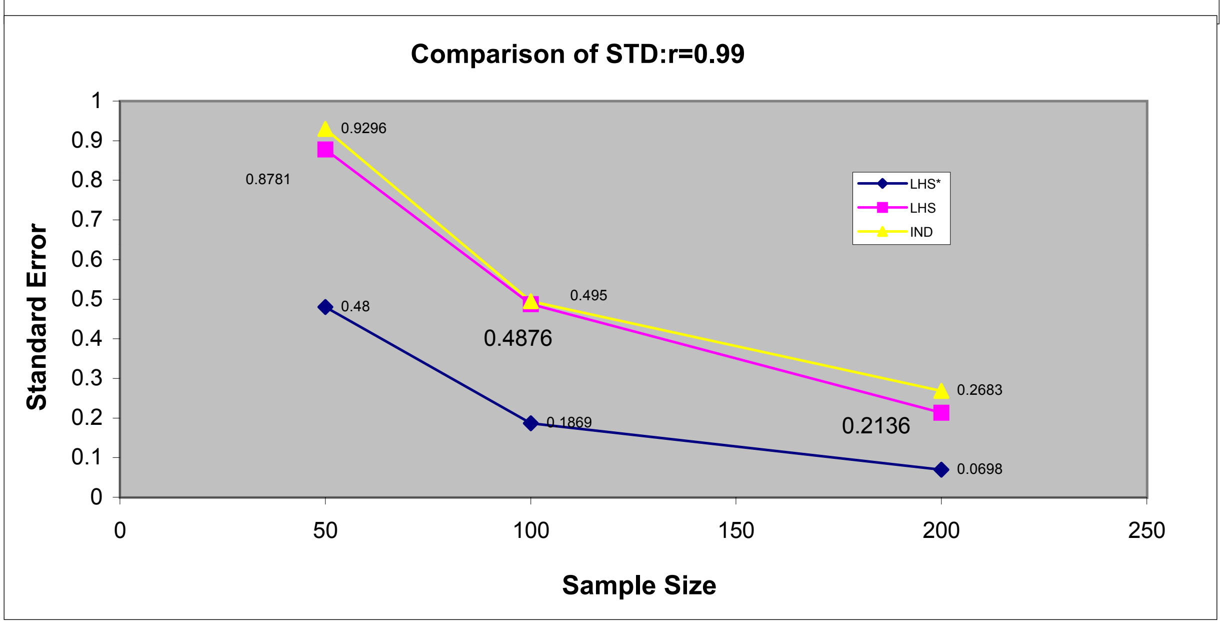
[11] McKay, M.D., R.J. Beckman, and W.J. Conover, "A comparison of three methods for selecting values of input variables in the analysis of output from a computer code," *Technometrics*, 21(1979), 239-245.

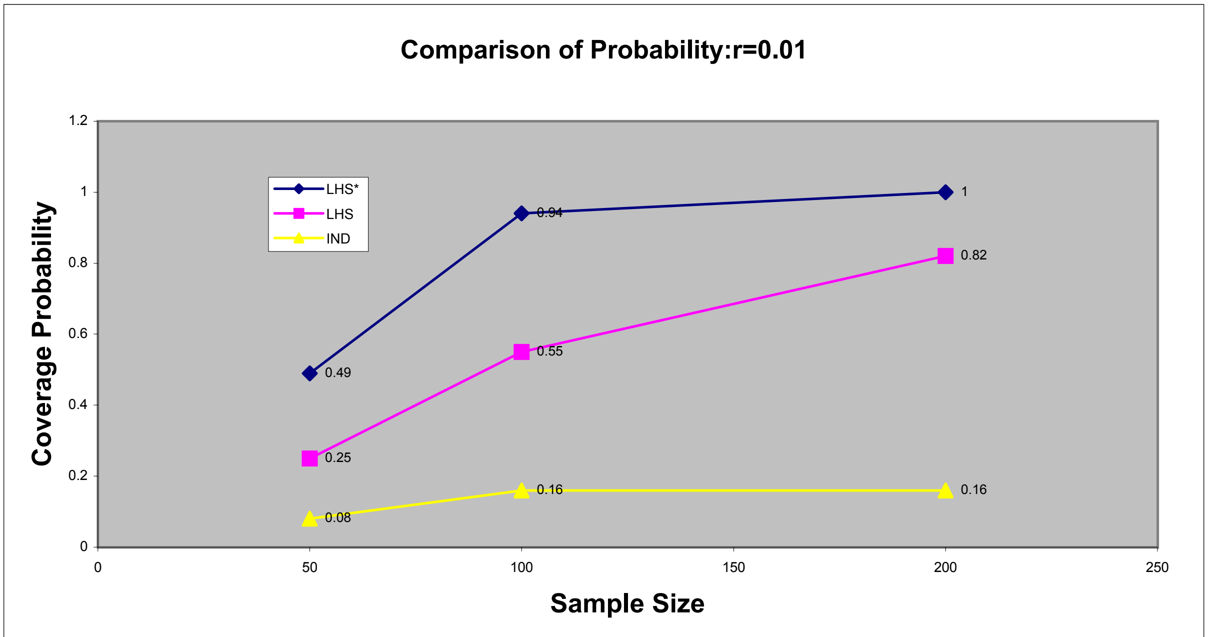
[12] Nelsen, R.B., *An Introduction to Copulas*, Springer, New York, 1999.

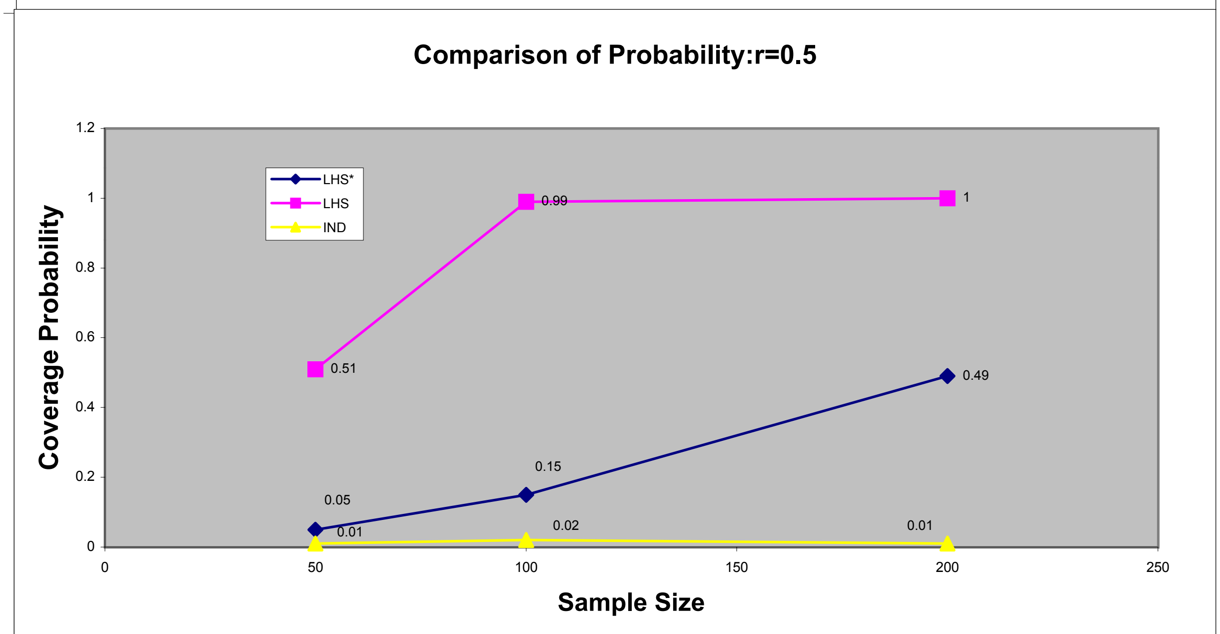
[13] Niederreiter, H., *Random Number Generation and Quasi-Monte Carlo Methods*, SIAM, Philadelphia, 1992.

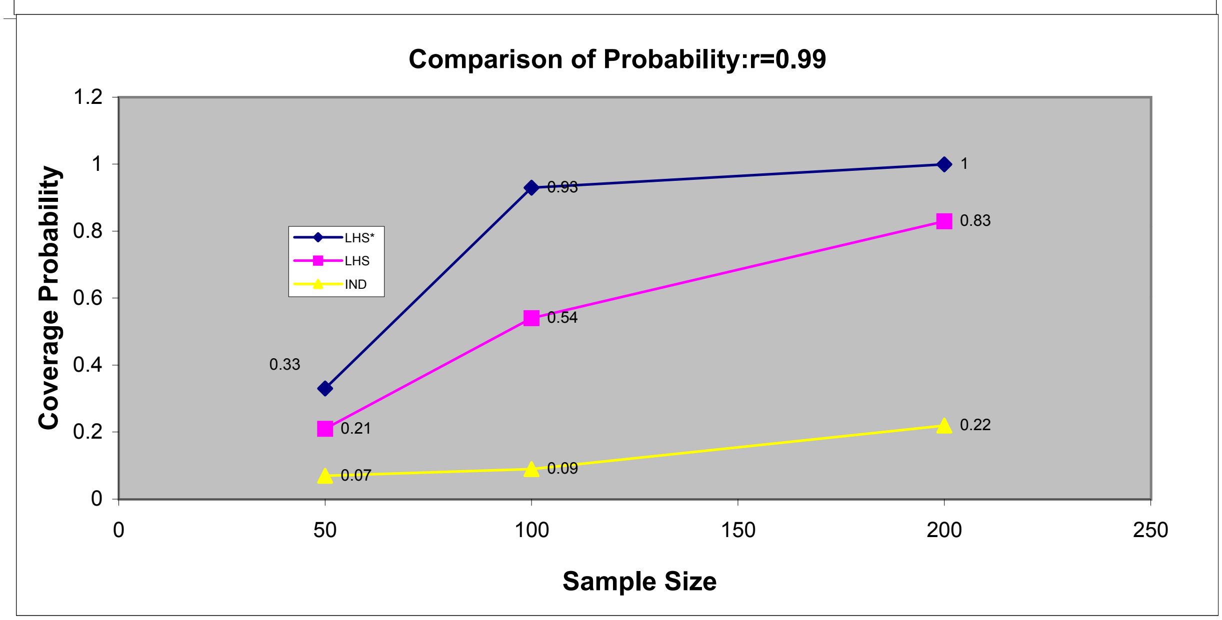

Comparison of STD:r=0.01
Standard Error
Sample Size
LHS*
LHS
IND
0.8969
0.8841
0.5649
0.4066
0.3955
0.2169
0.202
0.1997
0.0693
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
50
100
150
200
250


Comparison of STD:r=0.5
Standard Error
Sample Size
LHS*
LHS
IND
0.2871
0.2796
0.1601
0.0869
0.032
0.0116
0.0149
0.0029
0.00082
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0
50
100
150
200
250


Comparison of STD:r=0.99
Standard Error
Sample Size
LHS*
LHS
IND
0.9296
0.8781
0.48
0.495
0.4876
0.2683
0.2136
0.1869
0.0698
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
50
100
150
200
250


Comparison of Probability:r=0.01
Coverage Probability
Sample Size
LHS*
LHS
IND
0.49
0.94
1
0.25
0.55
0.82
0.08
0.16
0.16
0
0.2
0.4
0.6
0.8
1
1.2
0
50
100
150
200
250


Comparison of Probability:r=0.5
Coverage Probability
Sample Size
LHS*
LHS
IND
0.51
0.99
1
0.05
0.15
0.49
0.01
0.02
0.01
0
0.2
0.4
0.6
0.8
1
1.2
0
50
100
150
200
250


Comparison of Probability:r=0.99
Coverage Probability
Sample Size
LHS*
LHS
IND
0.33
0.93
1
0.21
0.54
0.83
0.07
0.09
0.22
0
0.2
0.4
0.6
0.8
1
1.2
0
50
100
150
200
250