Better Confidence Intervals for Importance Sampling

Halis Sak, Wolfgang Hörmann, Josef Leydold
Better Confidence Intervals for Importance Sampling*

Halis Sak†
Department of Statistics and Mathematics, WU (Vienna University of Economics and Business), Augasse 2-6, A-1090 Wien, Austria
halis.sak@wu.ac.at

Wolfgang Hörmann
Department of Industrial Engineering, Boğaziçi University, 34342 Bebek-Istanbul, Turkey
hormannw@boun.edu.tr

Josef Leydold
Department of Statistics and Mathematics, WU (Vienna University of Economics and Business), Augasse 2-6, A-1090 Wien, Austria
josef.leydold@wu.ac.at

It is well known that for highly skewed distributions the standard method of using the $t$ statistic for the confidence interval of the mean does not give robust results. This is an important problem for importance sampling (IS) as its final distribution is often skewed due to a heavy tailed weight distribution. In this paper, we first explain Hall’s transformation and its variants to correct the confidence interval of the mean and then evaluate the performance of these methods for two numerical examples from finance which have closed-form solutions. Finally, we assess the performance of these methods for credit risk examples. Our numerical results suggest that Hall’s transformation or one of its variants can be safely used in correcting the two-sided confidence intervals of financial simulations.

Keywords: confidence intervals; skewness removal; Hall’s transformation; importance sampling; quantitative risk management.

1. Introduction

In quantitative risk management Monte Carlo simulation is the only alternative when we are interested in the tail of the loss distribution in complex models; see the monographs [10] and [3] for a reference on quantitative risk management and Monte Carlo methods in finance. The problems we face in quantitative finance often include the evaluation of an integral, most prominently we have to compute

---


†Corresponding Author: Department of Statistics and Mathematics, WU (Vienna University of Economics and Business), Augasse 2-6, A-1090 Wien, Austria. Phone: +43 1 31336-4180, Fax: +43 1 31336-774, Email: halis.sak@wu.ac.at
the expectation
\[ E_P[g(x)] = \int g(x) \, dP = \int g(x) \varphi(x) \, dx \]
where \( P \) is the absolute continuous probability measure with density function \( \varphi(x) \).
Then naive Monte Carlo evaluates this expectation as
\[ E_P[g(x)] \approx \frac{1}{n} \sum_{k=1}^{n} g(X_i), \quad X_i \sim \varphi, \]
where the \( X_i \) are random numbers generated from density \( \varphi \) and \( n \) is the number of replications. The confidence intervals produced by naive Monte Carlo simulations are too wide to give precise results for rare event simulations. Thus we are in need of variance reduction techniques such as importance sampling, control variates, etc. (see Chapter 4 of [3] for a good description of the methods and examples).

The fundamental idea of importance sampling (IS) is to introduce a probability measure \( Q \) that has high probability in regions where the integrand \( g(x) \) has large values, i.e., for the “important domain” of \( g \). Thus we get
\[ E_P[g(x)] = \int g(x) \varphi(x) \, dx = \int g(x) \frac{\varphi(x)}{\psi(x)} \psi(x) \, dx = E_Q[g(x)w(x)] \]
where \( \psi \) denotes the probability density function for \( Q \) and \( w(x) = \varphi(x)/\psi(x) \) is called the likelihood ratio or weight of IS. Our aim is to find a new density \( \psi \) such that \( g(x)w(x) \) under measure \( Q \) has a lower variance than \( g(x) \) under measure \( P \). However, we must be careful about the distribution of the likelihood ratio [see 2, 8].

It is well known that the standard method of using the \( t \) distribution for computing the confidence interval of the mean does not give reliable results whenever we have a large skewness in a Monte Carlo simulation, since then the Gaussian approximation does not work very well, see [3]. This is in particular the case when only moderate sample sizes are possible as described in the example below.

We are in particular interested in the tail probability
\[ P \left( L = \frac{1}{\sqrt{d}} \sum_{i=1}^{d} Z_i > x \right) = \int_{\mathbb{R}^d} 1_{\{\sum_{i=1}^{d} z_i > x\}} \phi(z_i) \, dz \]
where \( Z_i \)'s are i.i.d. standard normal variates. We apply an IS strategy where we shift the mean of the variates \( Z \) to obtain the importance distribution [see, e.g., 4]. Hence we get a joint importance distribution of i.i.d. normal distributions with mean \( \mu_i = x/\sqrt{d} \) and where we choose a standard deviation \( \sigma_i = \sigma \) for some \( \sigma \) close to 1. The density of sampling and sample mean distribution for this IS strategy are visualized for \( \sigma = 1.2 \), \( d = 50 \), and \( n = 100 \) in Figure 1. It is clearly seen that the central limit theorem does not apply due to high skewness present in the likelihood function of the IS.

[9] and [7] have proposed transformations of the \( t \) variable to correct these deficiencies. In particular, Hall’s transformation [7] has the desirable characteristics of monotonicity and invertibility to correct for both bias and skewness from the
distribution of the $t$ statistic. [13] compared the performance of Johnson’s [9] and Hall’s transformations as well as their respective bootstrap versions by looking at the coverage accuracy of one-sided confidence intervals, that is, they compared expected and observed (true) confidence levels for the respective methods in several sampling problems. Their findings confirm superiority of Hall’s transformation for the lower endpoint confidence interval compared to Johnson’s transformation. Moreover, [12] and [11] propose new transformations, similar to Hall’s transformation, and survey existing ones to improve coverage accuracy of one- and two-sample problems.

To our knowledge using the $t$ statistic for confidence interval constructions for the mean is the standard in financial simulations. In this study we evaluate the performance of Hall’s transformation and its variants given in [12] and [11] in correcting the estimated confidence intervals for financial simulations that include IS. In this respect, our work is a contribution to better confidence intervals in this area.

The paper is organized as follows: Section 2 describes Hall’s transformation and its variants for the $t$ statistic. In Section 3, we consider two numerical examples from finance that have closed-form solutions where we apply an IS strategy to decrease the resultant variance of the simulations and assess the coverage accuracy of one-sided and two-sided confidence intervals using both ordinary $t$ statistic and the transformations. In Section 4 we look at the industry standard model of credit risk and evaluate the performance of the transformations in correcting the confidence intervals of simulations that use the two-step IS of [5]. Finally, in Section 5 we present our conclusions.

2. Transformations on the $t$ Statistic

Assume we have a random sample of size $n$, $X_1, ..., X_n$, from a population with mean $\theta$ and variance $\tau^2$. If we ignore skewness then we may use the $t$ statistic

$$T = \frac{\hat{\theta} - \theta}{\hat{\tau}/\sqrt{n}}$$
where $\hat{\theta} = \frac{1}{n} \sum_{i=1}^{n} X_i$ and $\hat{\tau}^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \hat{\theta})^2$ to construct the respective $(1-\alpha) \times 100\%$ confidence intervals on $\theta$ as $(\infty, \hat{\theta} - n^{-1/2}\tau z_{\alpha})$ and $(\hat{\theta} - n^{-1/2}\tau z_{1-\alpha}, \infty)$ where $z_{\alpha}$ is the $\alpha$-quantile of the standard normal distribution. If the distribution of $X$ is absolutely continuous and $E[X^4] < \infty$, then following [13], $T$ admits a first-order Edgeworth expansion

$$P(T \leq x) = \Phi(x) + n^{-1/2}\gamma(ax^2 + b)\phi(x) + O(n^{-1})$$

where $a = 1/3, b = 1/6$, and $\gamma$ is the skewness, defined as $E[(X - \theta)^3/\tau^3]$. $\Phi(\cdot)$ and $\phi(\cdot)$ denote cumulative distribution function and density of the standard normal distribution, respectively. Thus a one-sided standard $t$ interval has a coverage error of size $O(n^{-1/2})$, i.e.,

$$P(\theta \in (\infty, \hat{\theta} - n^{-1/2}\tau z_{\alpha})) = 1 - \alpha + O(n^{-1/2}).$$

To eliminate the effect of the skewness $\gamma$, [7] proposes the following monotone and invertible transformation on $T$:

$$g_1(T) = T + n^{-1/2}\hat{\gamma}(aT^2 + b) + n^{-1}(1/3) (a\hat{\gamma})^2 T^3$$

where $\hat{\gamma} = \frac{1}{n} \sum_{i=1}^{n} (X_i - \hat{\theta})^3/\hat{\tau}^3$ is an estimator for the skewness of $X$. Its unique inverse is then given by

$$T = g_1^{-1}(x) = n^{1/2}(a\hat{\gamma})^{-1} \left( \left(1 + 3a\hat{\gamma}(n^{-1/2}x - n^{-1/2}b\hat{\gamma})\right)^{1/3} - 1 \right). \quad (2.1)$$

Thus we obtain corrected $(1-\alpha) \times 100\%$ upper and lower endpoint confidence intervals

$$(\infty, \hat{\theta} - n^{-1/2}\tau g_1^{-1}(z_{\alpha})) \quad \text{and} \quad (\hat{\theta} - n^{-1/2}\tau g_1^{-1}(z_{1-\alpha}), \infty) \quad (2.2)$$

which have a coverage error of size $O(n^{-1})$ only. In the same way, the $(1-\alpha) \times 100\%$ two-sided equal-tailed corrected confidence interval for the mean is

$$\left(\hat{\theta} - n^{-1/2}\tau g_1^{-1}(z_{1-\alpha/2}), \hat{\theta} - n^{-1/2}\tau g_1^{-1}(z_{\alpha/2})\right). \quad (2.3)$$

[12] proposes a simpler transformation than Hall’s transformation which leads to the two-sided equal-tailed confidence interval for the mean

$$\left(\hat{\theta} - \tau g_2^{-1}(n^{-1/2}z_{1-\alpha/2}), \hat{\theta} - \tau g_2^{-1}(n^{-1/2}z_{\alpha/2})\right),$$

where $g_2^{-1}(x) = \left(1 + 3(x - n^{-1/2}\hat{\gamma})\right)^{1/3} - 1$.

Moreover, using the technique of Hall’s transformation, [11] proposes the two-sided equal-tailed confidence interval for the mean

$$\left(\hat{\theta} - n^{-1/2}\tau g_3^{-1}(t_{\nu,1-\alpha/2}), \hat{\theta} - n^{-1/2}\tau g_3^{-1}(t_{\nu,\alpha/2})\right)$$

where $g_3^{-1}(x) = (2r)^{-1} \left(1 + 6r(x - r))^{1/3} - 1\right), r = \frac{\hat{\mu}_3/\hat{\tau}^3}{\sqrt{\nu}}, \nu = n - 1, t_{\nu,\beta}$ is the $\beta$ quantile of the $t$-distribution with $\nu$ degrees of freedom and $\hat{\mu}_3 = n(n-1)^{-1}(n-
2)^{-1} \sum_{i=1}^{n} \left(X_i - \hat{\theta}\right)^3 \text{ (note that this estimate of the third central moment is different to the estimate of the skewness used above). For large samples } (n), \ t_{\nu, \beta} \text{ is of course very close to } z_\beta. \text{ We omit the formulas for the one-sided confidence intervals as they can be easily developed from the two-sided intervals using (2.2).}

3. Two Examples with Closed-Form Solutions

We want to assess the performance of Hall’s transformation compared to the standard \( t \) statistic for constructing confidence intervals. For this task we first perform simple financial simulations that include IS. We choose examples for which analytical solutions are available such that we are able to compute coverage frequencies. Following [7], [13], [12], and [11] we look at observed (true) confidence levels of one-sided and two-sided intervals which we call coverage level in the remaining part of the paper. The term \( 1 - \alpha \) confidence interval always refers to the interval that we have constructed with expected confidence level \( 1 - \alpha \) (and which should have coverage level \( 1 - \alpha \)).

The experiments are performed as following: Draw a random sample of size \( n \), compute the 95\% confidence interval (i.e., \( \alpha = 0.05 \)) and test whether it covers the parameter \( \theta \). Repeat this Bernoulli trial \( 10^4 \) times and use the observed frequency as estimator for the coverage level. Thus the formula that we use for the achieved coverage level of two-sided confidence interval is

\[
\frac{1}{10^4} \sum_{k=1}^{10^4} 1\{L^{(k)} \leq \theta \& U^{(k)} \geq \theta\}
\]

where \( L^{(k)} \) and \( U^{(k)} \) are calculated lower and upper bounds of the two-sided confidence interval for replication \( k \). Notice that this estimator is binomially distributed with approximate standard error \( \sqrt{\alpha(1-\alpha)/10^4} \approx 0.0022 \) (for \( \alpha = 0.05 \)).

Our first numerical example has already been described in Section 1. In this example we simulate the tail probability of a sum of standard normal variates. Table 1 shows the achieved coverage levels for one-sided (upper and lower endpoints, respectively) and two-sided 95\% intervals for a series of \( \sigma \) values, dimensions \( d \), and sample sizes \( n \). With Ord. \( t \), Hall, Z&D and Willink we denote the coverage levels using the ordinary \( t \) statistic, and transformations suggested in [7], [12], and [11], respectively. The coverage levels are expected to be 0.95 with an approximate standard error of 0.0022 as mentioned before. In [13] it was observed that upper endpoint confidence intervals for the transformations have poor coverage accuracy when compared to lower endpoint confidence intervals in small samples. We see the same behavior, especially for the case \( \sigma = 0.9, d = 50, n = 100 \). However, for all confidence intervals considered in Table 1 using one of the transformations always lead to more reliable results than the standard \( t \) statistic when \( n \) is small. This is in particular the case for increasing dimensions and \( \sigma \neq 1 \). Hall’s and Willink’s transformations produce very similar coverage levels that are slightly better than
those of Z&D’s transformation especially for the highly skewed cases \((n = 100)\).

Our second example focuses on geometric average Asian call option pricing. The price of a geometric average Asian options is quite close to an arithmetic average Asian option but has a closed form solution when using a Black-Scholes model. So let us assume that the maturity of this Asian option is 0.2 (years) with control points at 0.12, 0.14, 0.16, 0.18, and 0.2. We apply a similar IS strategy as in the previous example. The parameters \(\mu\) and \(\sigma\) for the marginal normal distributions of the importance distribution are computed to get minimal variance for the Monte Carlo estimate. For this task we used the Nelder-Mead simplex method as implemented in GSL [1]. The achieved coverage levels for 95% confidence intervals are presented in Table 2 for \(S_0 = 100\) and \(r = 0.09\). The standard \(t\) statistic leads to close to correct results for small values of the strike price \(K\); Hall’s and Willink’s transformations produce practically the same results but one sided intervals of [12] perform worse than the \(t\) statistic for \(n = 100\) and \(K = 90\) or 100. For increasing strike price \(K\) the \(t\) statistic leads to wrong confidence intervals and the intervals produced by the transformations are clearly better. As in the case of Table 1, Hall’s and Willink’s transformations produce very similar coverage levels but different to the situation of Table 1 the coverage levels obtained by Z&D’s transformation are clearly worse.
Table 1. Achieved coverage levels for one-sided and two-sided 95% confidence intervals for IS in computing tail probabilities of a sum of standard normal variate ($\sigma$).

<table>
<thead>
<tr>
<th>n</th>
<th>d</th>
<th>Upper Endpoint</th>
<th>Lower Endpoint</th>
<th>Two-Sided</th>
<th>Hall</th>
<th>Z&amp;D Willink</th>
<th>Ord. t Hall</th>
<th>Z&amp;D Willink</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0.912</td>
<td>0.940</td>
<td>0.952</td>
<td>0.946</td>
<td>0.951</td>
<td>0.931</td>
<td>0.945</td>
<td>0.946</td>
</tr>
<tr>
<td>20</td>
<td>0.900</td>
<td>0.887</td>
<td>0.932</td>
<td>0.931</td>
<td>0.931</td>
<td>0.949</td>
<td>0.948</td>
<td>0.949</td>
</tr>
<tr>
<td>50</td>
<td>1.000</td>
<td>0.896</td>
<td>0.925</td>
<td>0.925</td>
<td>0.925</td>
<td>0.951</td>
<td>0.951</td>
<td>0.951</td>
</tr>
<tr>
<td>100</td>
<td>1.000</td>
<td>0.896</td>
<td>0.925</td>
<td>0.925</td>
<td>0.925</td>
<td>0.951</td>
<td>0.951</td>
<td>0.951</td>
</tr>
<tr>
<td>500</td>
<td>1.000</td>
<td>0.896</td>
<td>0.925</td>
<td>0.925</td>
<td>0.925</td>
<td>0.951</td>
<td>0.951</td>
<td>0.951</td>
</tr>
</tbody>
</table>

Better Confidence Intervals for Importance Sampling 7
Table 2. Achieved coverage levels for one-sided and two-sided 95% confidence intervals for the IS in geometric average Asian call option pricing ($S_0 = 100$, $r = 0.09$).

<table>
<thead>
<tr>
<th>volatility</th>
<th>$K$</th>
<th>optimal IS</th>
<th>$n$</th>
<th>Upper Endpoint</th>
<th>Lower Endpoint</th>
<th>Two-Sided</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>90</td>
<td>105</td>
<td>1.002</td>
<td>100</td>
<td>0.948</td>
<td>0.948</td>
<td>0.977</td>
</tr>
<tr>
<td>100</td>
<td>0.279</td>
<td>1.033</td>
<td>100</td>
<td>0.940</td>
<td>0.949</td>
<td>0.973</td>
</tr>
<tr>
<td>105</td>
<td>0.481</td>
<td>1.116</td>
<td>100</td>
<td>0.920</td>
<td>0.949</td>
<td>0.958</td>
</tr>
<tr>
<td>110</td>
<td>0.758</td>
<td>1.281</td>
<td>100</td>
<td>0.869</td>
<td>0.947</td>
<td>0.916</td>
</tr>
<tr>
<td>115</td>
<td>1.168</td>
<td>1.505</td>
<td>100</td>
<td>0.933</td>
<td>0.952</td>
<td>0.943</td>
</tr>
<tr>
<td>120</td>
<td>1.470</td>
<td>1.687</td>
<td>100</td>
<td>0.586</td>
<td>0.787</td>
<td>0.631</td>
</tr>
<tr>
<td>125</td>
<td>1.750</td>
<td>1.750</td>
<td>100</td>
<td>0.586</td>
<td>0.787</td>
<td>0.631</td>
</tr>
</tbody>
</table>
4. Credit Risk Application

4.1. The Normal Copula Model

We are interested in the normal copula model of CreditMetrics (see [6]) for the
dependence structure across obligors. Following [5] we use the following notation:

- \( m \): number of obligors in portfolio
- \( Y_j \): default indicator for \( j \)th obligor (1 if default occurs, 0 otherwise)
- \( c_j \): loss resulting from the default of \( j \)th obligor
- \( p_j \): marginal default probability of \( j \)th obligor
- \( L = \sum_{j=1}^{m} c_j Y_j \): total loss of portfolio
- \( n \): number of replications in a simulation.

The loss values \( c_j \) and the default probabilities \( p_j \) are known and constant over
the fixed horizon.

The normal copula model introduces a multivariate normal vector \((X_1, \ldots, X_m)\)
of latent variables to obtain dependence across obligors. Relationships between the
default indicators and the latent variables are represented by

\[
Y_j = 1_{\{X_j > x_j\}}, \quad j = 1, \ldots, m,
\]

where \( X_j \) has standard normal distribution and \( x_j = \Phi^{-1}(1 - p_j) \). Obviously, the
threshold value of \( x_j \) is chosen such that \( P(Y_j = 1) = p_j \). The correlations among
\( X_j \) are modeled as

\[
X_j = b_j \epsilon_j + a_{j1} Z_1 + \cdots + a_{jd} Z_d, \quad j = 1, \ldots, m, \tag{4.1}
\]

where \( \epsilon_j \) and \( Z_1, \ldots, Z_d \) are independent standard normal random variables with
\( b_j^2 + a_{j1}^2 + \cdots + a_{jd}^2 = 1 \). While \( Z_1, \ldots, Z_d \) are systematic risk factors affecting all of
the obligors, \( \epsilon_j \) is the idiosyncratic risk factor affecting only obligor \( j \). Furthermore,
\( a_{j1}, \ldots, a_{jd} \) are constant, nonnegative factor loadings, assumed to be known. Thus,
given the vector \( Z = (Z_1, \ldots, Z_d) \), we have the conditionally independent default probabilities

\[
p_j(Z) = P(Y_j = 1|Z) = \Phi\left(\frac{a_j Z + \Phi^{-1}(p_j)}{b_j}\right), \quad j = 1, \ldots, m, \tag{4.2}
\]

where \( a_j = (a_{j1}, \ldots, a_{jd}) \).

[5] proposes a two-step IS strategy for computing \( P(L > x) \) in the normal copula
model. First they employ exponential twisting to increase the conditional default
probabilities and get

\[
p_{j, \theta_x}(Z) = \frac{p_j(Z) e^{\theta_x(Z) c_j}}{p_j(Z) e^{\theta_x(Z) c_j} + 1 - p_j(Z)}, \quad j = 1, \ldots, m, \tag{4.3}
\]

where \( \theta_x(Z) \) is chosen such that it minimizes the upper bound of the second moment
of the IS estimator. For this, if \( E[L|Z] = \sum_{j=1}^{m} p_j(Z) c_j < x \) then \( \theta_x(Z) \) is the unique
solution to

\[
\psi'_{L|Z}(\theta_x(Z)) = x \tag{4.4}
\]
where
\[ \psi_{L|Z}(\theta_x(Z)) = \sum_{j=1}^{m} \log \left( 1 + p_j(Z) \left( e^{c_j \theta_x(Z)} - 1 \right) \right). \]

Otherwise we set \( \theta_x(Z) = 0 \). The likelihood ratio of this step is
\[ \exp \left( -\theta_x(Z)L + \psi_{L|Z}(\theta_x(Z)) \right). \]

For further reduction of the variance (for highly dependent obligors exponential twisting is not enough) [5] shift the mean of \( Z \) from the origin to some point \( \mu \) to get the importance distribution of the second step. For finding \( \mu \) they solve the optimization problem
\[ \mu = \max_x P(L > x|Z = z) e^{-\frac{1}{2}z^T z} \tag{4.5} \]
which is the mode of the zero-variance IS function [see 4]. A tail bound approximation in which
\[ P(L > x|Z = z) \approx e^{-\theta_x(z) x + \psi_{L|Z}(\theta_x(z))} \]
is used in (4.5) to solve for an approximate \( \mu \). The likelihood ratio for the second step is \( \exp(-\mu^T Z + \frac{1}{2} \mu^T \mu) \). The details of the combined IS algorithm and the application of the two-sided confidence interval (using Hall’s transformation) is presented as Algorithm 1.

There are two numerical examples (10- and 21-factor models) given in [5]. We define a third numerical example ourselves. It is a 5-factor model with 4800 obligors. Obligors are separated into 6 segments of size 800 each as seen in Table 3. Default probabilities, exposure levels and factor loadings are the same in each segment for the obligors. We evaluate the performance of confidence intervals using the transformations compared to standard \( t \) statistic on these three numerical examples.

Table 3. Portfolio composition for the 5-factor model; default probabilities, exposure levels and factor loadings for six segments.

<table>
<thead>
<tr>
<th>Segment</th>
<th>Obligor j</th>
<th>( p_j )</th>
<th>( c_j )</th>
<th>( a_{j,1} )</th>
<th>( a_{j,2} )</th>
<th>( a_{j,3} )</th>
<th>( a_{j,4} )</th>
<th>( a_{j,5} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1A</td>
<td>1 – 800</td>
<td>0.01</td>
<td>20</td>
<td>0.7</td>
<td>0.5</td>
<td>0.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1B</td>
<td>801 – 1600</td>
<td>0.02</td>
<td>10</td>
<td>0.7</td>
<td>0.5</td>
<td>0.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2A</td>
<td>1601 – 2400</td>
<td>0.02</td>
<td>10</td>
<td>0.7</td>
<td>0.2</td>
<td>0.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2B</td>
<td>2401 – 3200</td>
<td>0.04</td>
<td>5</td>
<td>0.7</td>
<td>0.2</td>
<td>0.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3A</td>
<td>3201 – 4000</td>
<td>0.03</td>
<td>5</td>
<td>0.7</td>
<td>0.4</td>
<td>0.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3B</td>
<td>4001 – 4800</td>
<td>0.05</td>
<td>1</td>
<td>0.7</td>
<td>0.4</td>
<td>0.5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4 shows the achieved coverage levels for 95% one-sided and two-sided confidence intervals for the tail loss probabilities of all three models. The sample size is 100. As suggested in [5] for the case of several \( x \)-values, we compute the mean shift \( \langle \mu \rangle \) for the smallest \( x \) value and use it for all \( x \) values at step 1 of Algorithm 1. (To
Algorithm 1 Computation of tail loss probability and its corrected two-sided confidence interval (using Hall’s transformation) for dependent obligors

**Output:** Tail loss probability and its two-sided \((1 - \alpha) \times 100\%\) confidence interval for dependent obligors using two step IS of [5] and Hall’s transformation.

1: Compute \(\mu\) using (4.5).
2: for \(k = 1, \ldots, n\) do
3: Generate i.i.d. \(z_l \sim N(\mu_l, 1)\) for \(l = 1, \ldots, d\).
4: Calculate \(p_j(Z)\) for \(j = 1, \ldots, m\) using (4.2).
5: if \(E[L|Z] = \sum_{j=1}^{m} p_j(Z) c_j < x\) then
6: Compute \(\theta_x(Z)\) using (4.4).
7: else
8: Set \(\theta_x(Z) = 0\).
9: Calculate \(p_j,\theta_x(Z)\) for \(j = 1, \ldots, m\) using (4.3).
10: Calculate total loss \(L^{(k)} = \sum_{j=1}^{m} c_j Y_j\) for for replication \(k\).
11: Calculate likelihood ratio for replication \(k\):
\[
    w^{(k)} = \exp(-\mu^T Z + \frac{1}{2} \mu^T \mu - \theta_x(Z)L + \psi_L|Z(\theta_x(Z)))
\]
12: Calculate loss probability for replication \(k\):
\[
    X^{(k)} = 1_{\{L^{(k)} > x\}} w^{(k)}
\]
13: Calculate \(\hat{\theta} = \frac{1}{n} \sum_{k=1}^{n} X^{(k)}\).
14: Calculate standard deviation and skewness estimates using \(\hat{\tau}^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(X^{(k)} - \hat{\theta}\right)^2\) and \(\hat{\gamma} = \frac{1}{n} \sum_{i=1}^{n} \left(X^{(k)} - \hat{\theta}\right)^3 / \hat{\tau}^3\).
15: Calculate two-sided \((1 - \alpha) \times 100\%\) confidence interval using (2.3)

compute the nearly exact tail loss probabilities required to calculate the coverage levels, we made simulations with \(10^6\) repetitions.) Our simulation results show that the two-sided confidence intervals using the transformations are better than those using the standard \(t\) statistic for our credit risk examples. Hall and Willink transformations again produce very similar coverage levels slightly better than those of Z&D transformation.
Table 4. Achieved coverage levels for one-sided and two-sided 95% confidence intervals for computing tail loss probabilities in credit risk examples.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ord. $t$</td>
<td>Hall Z&amp;D</td>
<td>Willink</td>
<td>Ord. $t$</td>
<td>Hall Z&amp;D</td>
</tr>
<tr>
<td>500</td>
<td>3.862 × 10^{-2}</td>
<td>0.917 0.945 0.956</td>
<td>0.947 0.969 0.950</td>
<td>0.941 0.951 0.933</td>
<td>0.948 0.949 0.949</td>
</tr>
<tr>
<td>1000</td>
<td>8.581 × 10^{-3}</td>
<td>0.921 0.954 0.962</td>
<td>0.958 0.966 0.951</td>
<td>0.940 0.952 0.938</td>
<td>0.955 0.949 0.955</td>
</tr>
<tr>
<td>10-factor</td>
<td>2000</td>
<td>8.434 × 10^{-4}</td>
<td>0.903 0.968 0.947</td>
<td>0.972 0.974 0.956</td>
<td>0.951 0.956 0.926</td>
</tr>
<tr>
<td></td>
<td>3000</td>
<td>1.059 × 10^{-4}</td>
<td>0.860 0.977 0.905</td>
<td>0.978 0.986 0.959</td>
<td>0.964 0.960 0.884</td>
</tr>
<tr>
<td></td>
<td>4000</td>
<td>1.302 × 10^{-5}</td>
<td>0.715 0.801 0.746</td>
<td>0.801 0.995 0.966</td>
<td>0.978 0.966 0.732</td>
</tr>
<tr>
<td>21-factor</td>
<td>2500</td>
<td>5.016 × 10^{-2}</td>
<td>0.905 0.939 0.949</td>
<td>0.943 0.973 0.950</td>
<td>0.945 0.951 0.928</td>
</tr>
<tr>
<td></td>
<td>10000</td>
<td>1.122 × 10^{-2}</td>
<td>0.918 0.949 0.958</td>
<td>0.952 0.968 0.953</td>
<td>0.942 0.954 0.938</td>
</tr>
<tr>
<td></td>
<td>20000</td>
<td>2.717 × 10^{-3}</td>
<td>0.915 0.953 0.957</td>
<td>0.957 0.972 0.955</td>
<td>0.947 0.956 0.936</td>
</tr>
<tr>
<td></td>
<td>30000</td>
<td>6.156 × 10^{-4}</td>
<td>0.896 0.960 0.941</td>
<td>0.964 0.978 0.956</td>
<td>0.952 0.957 0.922</td>
</tr>
<tr>
<td></td>
<td>40000</td>
<td>7.348 × 10^{-5}</td>
<td>0.858 0.976 0.905</td>
<td>0.979 0.987 0.957</td>
<td>0.962 0.958 0.882</td>
</tr>
<tr>
<td>5-factor</td>
<td>5000</td>
<td>4.640 × 10^{-2}</td>
<td>0.903 0.933 0.946</td>
<td>0.936 0.977 0.952</td>
<td>0.951 0.953 0.928</td>
</tr>
<tr>
<td></td>
<td>15000</td>
<td>8.343 × 10^{-3}</td>
<td>0.886 0.935 0.936</td>
<td>0.938 0.983 0.957</td>
<td>0.959 0.957 0.917</td>
</tr>
<tr>
<td></td>
<td>20000</td>
<td>3.966 × 10^{-3}</td>
<td>0.870 0.937 0.926</td>
<td>0.942 0.980 0.952</td>
<td>0.955 0.953 0.906</td>
</tr>
<tr>
<td></td>
<td>25000</td>
<td>1.851 × 10^{-3}</td>
<td>0.884 0.959 0.932</td>
<td>0.963 0.979 0.953</td>
<td>0.956 0.954 0.910</td>
</tr>
<tr>
<td></td>
<td>30000</td>
<td>7.753 × 10^{-4}</td>
<td>0.874 0.971 0.921</td>
<td>0.975 0.985 0.955</td>
<td>0.959 0.956 0.901</td>
</tr>
</tbody>
</table>
5. Conclusions

We demonstrated with simulation examples taken from the field of quantitative finance that Hall’s transformation [7] or variants of it should be used for correcting the confidence intervals of financial simulations that employ IS. The difference between the standard $t$ statistic and the transformations becomes large when we have a small number of simulations or very large skewness. Notice that in practice we may have to use a small number of simulations for real world credit risk problems as portfolios with $10^5$ or more obligors are considered in practice. Furthermore, the use of the transformations does not lead to any disadvantage as they are easily implemented and require a negligible amount of additional computations.

Considering the variants, we observed that the results for Willink’s transformation [11] were almost identical to the results of Hall’s transformation in all our experiments. Z&D’s transformation [12] performed slightly worse in two of our experiments and clearly worse in the third one. We therefore recommend to use Hall’s or Willink’s transformation instead of the standard $t$ statistic to construct two-sided confidence intervals for simulations that use importance sampling.

Acknowledgements

The authors thank the anonymous referee for the valuable suggestions that improved the presentation of the paper. This work was partially supported by Bogazici Research Fund Project 5020.

References


