Numerical Computation of Asymptotic Variance of IS Quantile Estimator and Covariance

April 12, 2023

This file describes some of the methods applied in developing the source code in [2] used for creating some of the numerical results in the paper [3].

Suppose that importance sampling (IS) is applied to estimate the economic capital (EC), which is the difference between a quantile and mean. The asymptotic variance in the central limit theorem (CLT) of the IS estimator of the EC involves the asymptotic variances and covariance of the IS estimators of the quantile and mean. These values can be expressed as integrals whose values are not known in closed form, so we use quadrature to numerically compute them. But applying quadrature on standard representations of the asymptotic variance and covariance can lead to numerical issues. For example, round-off errors can result in the computed asymptotic variance being negative. To address this issue, this note provides some algebraically equivalent representations for the asymptotic variance and covariance that seem to be more amenable to numerical computation.

1 Model and Notation

Let $Y = c(\mathbf{X}) = \sum_{j=1}^{m} X_j$, where $\mathbf{X} = (X_1, X_2, \ldots, X_m) \sim G$, and $X_1, X_2, \ldots, X_m$ are independent and identically distributed (i.i.d.) with marginal cumulative distribution function (CDF) G_0. Let F be the CDF of Y, and let f be the density function. For $p \in (0,1)$, the p-quantile of Y (or F) is $\xi = F^{-1}(p) = \inf\{y : F(y) \geq p\}$. The EC is defined as $\eta = \xi - \mu$, where $\mu = E_G[Y]$ is the mean of Y under F, and E_G denotes the expectation operator when $\mathbf{X} \sim G$. As in [1], we assume that the quantile level p satisfies $p \equiv p_m = 1 - e^{-\beta m}$, for some fixed constant $\beta > 0$, and want to study analytically and numerically the behavior of EC estimators as $m \to \infty$.

Note that F satisfies

$$F(y) = P(Y \leq y) = E_G[I(c(\mathbf{X}) \leq y)], \tag{1}$$

where $I(\cdot)$ denotes the indicator function. Also, the mean and variance of Y are $\mu = \sum_{j=1}^{m} E_G[X_j] = m\mu_0$ and $\sigma^2 = m\sigma_0^2$, respectively, where $\mu_0 = E_0[X_j]$ and $\sigma_0^2 = \mathrm{Var}_0[X_j]$, with E_0 and Var_0 denoting the expectation and variance operators when $X_j \sim G_0$.

When applying simple random sampling (SRS), we generate n i.i.d. copies $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ of $\mathbf{X}$, and compute $Y_i = c(\mathbf{X}_i)$ for each $i = 1, 2, \ldots, n$. We estimate the CDF F via the empirical distribution $\widehat{F}_{\mathrm{SRS},n}$, where $\widehat{F}_{\mathrm{SRS},n}(y) = (1/n)\sum_{i=1}^{n} I(Y_i \le y)$. The SRS estimator $\widehat{\xi}_{\mathrm{SRS},n} = \widehat{F}^{-1}_{\mathrm{SRS},n}(p)$ has asymptotic variance

$$\kappa^2_{\mathrm{SRS}} = \frac{\chi^2_{\mathrm{SRS}}}{f^2(\xi)}, \qquad \text{with} \quad \chi^2_{\mathrm{SRS}} = p(1-p). \tag{2}$$

Specifically, if $f(\xi) > 0$, then the SRS quantile estimator obeys the following CLT: $\sqrt{n}[\widehat{\xi}_{\mathrm{SRS},n} - \xi] \Rightarrow N(0, \kappa^2_{\mathrm{SRS}})$ as $n \to \infty$ for fixed $p \in (0,1)$, where $\Rightarrow$ denotes convergence in distribution, and $N(a, b^2)$ is a normal random variable with mean a and variance b^2.

Let IS(θ) denote importance sampling using exponential twisting with parameter θ. Specifically, let $M_0(\theta) = E_0[e^{\theta X_j}] = \int e^{\theta x}\, \mathrm{d}G_0(x)$, $\theta \in \Re$, be the moment generating function of G_0, and $Q_0(\theta) = \ln M_0(\theta)$ be the cumulant generating function (CGF) of CDF G_0. For θ satisfying $M_0(\theta) < \infty$, IS(θ) samples $\mathbf{X} = (X_1, \ldots, X_m) \sim \widetilde{G}_\theta$, where under $\widetilde{G}_\theta$, the summands $X_1, \ldots, X_m$ are i.i.d., with each X_j having marginal CDF $\widetilde{G}_{0,\theta}(\cdot)$ defined by

$$\mathrm{d}\widetilde{G}_{0,\theta}(x) = e^{\theta x}\mathrm{d}G_0(x)/M_0(\theta) = e^{\theta_\star x - Q_0(\theta)}\mathrm{d}G_0(x).$$

The resulting likelihood ratio is

$$L_\theta(\mathbf{X}) = \prod_{j=1}^{m} \frac{\mathrm{d}G_0(X_j)}{\mathrm{d}\widetilde{G}_{0,\theta}(X_j)} = \exp\left(mQ_0(\theta) - \theta\sum_{j=1}^{m} X_j\right) = e^{mQ_0(\theta) - \theta Y}. \tag{3}$$

When applying IS(θ) with sampling size n, we generate n i.i.d. copies $\mathbf{X}_i$, $i = 1, 2, \ldots, n$, of $\mathbf{X} \sim \widetilde{G}_\theta$, and the IS($\theta$) estimator of F is $\widehat{F}_{\mathrm{IS}(\theta),n}$, with

$$\widehat{F}_{\mathrm{IS}(\theta),n}(y) = 1 - \frac{1}{n}\sum_{i=1}^{n} I(c(\mathbf{X}_i) > y) L_\theta(\mathbf{X}_i),$$

where as before $c(\mathbf{X}) = \sum_{j=1}^{m} X_j$. The IS($\theta$) estimator of ξ is $\widehat{\xi}_{\mathrm{IS}(\theta),n} = \widehat{F}^{-1}_{\mathrm{IS}(\theta),n}(p)$.

As suggested by [1], we choose $\theta = \theta_\star > 0$ as the root of

$$-\theta_\star Q_0'(\theta_\star) + Q_0(\theta_\star) = -\beta, \tag{4}$$

where $Q_0'(\cdot)$ and $Q_0''(\cdot)$ are the first two derivatives of $Q_0(\cdot)$. If $f(\xi) > 0$, the IS($\theta_\star$) p-quantile estimator $\widehat{\xi}_{\mathrm{IS}(\theta_\star),n}$ obeys a CLT $\sqrt{n}[\widehat{\xi}_{\mathrm{IS}(\theta_\star),n} - \xi] \Rightarrow N(0, \kappa^2_{\mathrm{IS}(\theta_\star)})$ as $n \to \infty$ for fixed $p \in (0,1)$, which has asymptotic variance

$$\kappa^2_{\mathrm{IS}(\theta_\star)} = \frac{\chi^2_{\mathrm{IS}(\theta_\star)}}{f^2(\xi)}, \quad \text{with} \tag{5}$$

$$\chi^2_{\mathrm{IS}(\theta_\star)} \equiv \mathrm{Var}_{\widetilde{G}_{\theta_\star}}[L_{\theta_\star}(\mathbf{X}) I(c(\mathbf{X}) > \xi)] = E_{\widetilde{G}_{\theta_\star}}[L^2_{\theta_\star}(\mathbf{X}) I(c(\mathbf{X}) > \xi)] - \overline{F}^2(\xi), \tag{6}$$

where $E_{\widetilde{G}_{\theta_\star}}$ and $\mathrm{Var}_{\widetilde{G}_{\theta_\star}}$ denote the expectation and variance operators when $\mathbf{X} \sim \widetilde{G}_{\theta_\star}$, and $\overline{F}(y) = 1 - F(y)$ so $\overline{F}^2(\xi) = [1 - F(\xi)]^2 = (1-p)^2$. Recall that (1) expresses the CDF under

SRS (i.e., $\mathbf{X} \sim G$) as an expectation of an indicator function. For IS($\theta_\star$), we can similarly express the CDF $\widetilde{F}_{\theta_\star}$ of $Y = c(\mathbf{X})$ for $\mathbf{X} \sim \widetilde{G}_{\theta_\star}$ as

$$\widetilde{F}_{\theta_\star}(y) = E_{\widetilde{G}_{\theta_\star}}[I(c(\mathbf{X}) \le y)]. \tag{7}$$

When the CDF F of the sum Y under SRS comes from a so-called *exponential family* of distributions (including normal and Erlang), exponential twisting leads to $\widetilde{F}_{\theta_\star}$ being from the same parametric family as F but with different parameters. For example, when F is a normal, exponential twisting results in a normal with a different mean, but the variance remains the same. Also, when F is an Erlang, exponential twisting results again in an Erlang with a different scale parameter, but the shape parameter remains the same. Section 4 will provide further details. The CDF $\widetilde{F}_{\theta_\star}$ of Y under IS($\theta_\star$) will have mean $\breve{\xi} = mQ_0'(\theta_\star)$.

2 Variance of IS Quantile Estimator

A Bahadur representation shows that $\chi^2_{\mathrm{IS}(\theta_\star)}$ is the variance of the IS($\theta_\star$) estimator $\widehat{F}_{\mathrm{IS}(\theta_\star),n}(\xi)$ of $F(\xi)$, where we recall that

$$\widehat{F}_{\mathrm{IS}(\theta_\star),n}(y) = 1 - \frac{1}{n}\sum_{i=1}^{n} I(c(\mathbf{X}_i) > y) L_{\theta_\star}(\mathbf{X}_i), \tag{8}$$

with i.i.d. $\mathbf{X}_i \sim \widetilde{G}_{\theta_\star}$, $i = 1, 2, \ldots, n$, The representation in (6) expresses $\chi^2_{\mathrm{IS}(\theta_\star)}$ as the difference of the second moment, $E_{\widetilde{G}_{\theta_\star}}[L^2_{\theta_\star}(\mathbf{X}) I(c(\mathbf{X}) > \xi)]$, and the squared first moment, $\overline{F}^2(\xi)$, and Jensen's inequality ensures that the difference is nonnegative. Thus, one way of numerically computing $\chi^2_{\mathrm{IS}(\theta_\star)}$ is to separately compute the second moment and the square of the first moment, and then take their difference. If we could do infinite-precision computation, then this would work fine. But when I use Excel and Matlab with this approach, I get reasonable values for $\chi^2_{\mathrm{IS}(\theta_\star)}$ only when the dimension m is small. When m is not small, I sometimes end up with $\chi^2_{\mathrm{IS}(\theta_\star)} < 0$, e.g., for $m \ge 16$ when G_0 is the standard normal CDF Φ. The problem seems to arise from numerical issues with finite-precision calculations caused by having the second moment being just barely larger than the squared first moment, and taking their difference results in loss of precision from round-off error.

To try to address this issue, we can instead rewrite the variance in (6) using a different equivalent representation:

$$\chi^2_{\mathrm{IS}(\theta_\star)} = E_{\widetilde{G}_{\theta_\star}}\left[\left(L_{\theta_\star}(\mathbf{X}) I(c(\mathbf{X}) > \xi) - \overline{F}(\xi)\right)^2\right], \tag{9}$$

which may allow numerical calculations with only nonnegative quantities. To do this, we first define the notation $E[A; B] = E[A\, I(B)]$ for a random variable A and an event B, where $I(\cdot)$ is the indicator function, so for B^c as the complement of B, the law of total probability ensures that $E[A] = E[A; B] + E[A; B^c]$ because $P(B \cup B^c) = 1$ and $P(B \cap B^c) = 0$. Now take A as the quantity in (9) that we are taking the expectation of, $B = \{c(\mathbf{X}) \le \xi\}$, and

$B^c = \{c(\mathbf{X}) > \xi\}$. Thus, (9) becomes

$$
\begin{aligned}
\chi^2_{\mathrm{IS}(\theta_\star)} &= E_{\widetilde{G}_{\theta_\star}}\left[\left(L_{\theta_\star}(\mathbf{X}) I(c(\mathbf{X}) > \xi) - \overline{F}(\xi)\right)^2; c(\mathbf{X}) \le \xi\right], \\
&\quad + E_{\widetilde{G}_{\theta_\star}}\left[\left(L_{\theta_\star}(\mathbf{X}) I(c(\mathbf{X}) > \xi) - \overline{F}(\xi)\right)^2; c(\mathbf{X}) > \xi\right] \\
&= a_1(\xi) + a_2(\xi),
\end{aligned} \tag{10}
$$

with for $z \in \Re$,

$$
a_1(z) = E_{\widetilde{G}_{\theta_\star}}\left[\left(L_{\theta_\star}(\mathbf{X}) I(c(\mathbf{X}) > z) - \overline{F}(z)\right)^2; c(\mathbf{X}) \le z\right], \tag{11}
$$

$$
a_2(z) = E_{\widetilde{G}_{\theta_\star}}\left[\left(L_{\theta_\star}(\mathbf{X}) I(c(\mathbf{X}) > z) - \overline{F}(z)\right)^2; c(\mathbf{X}) > z\right], \tag{12}
$$

where both $a_1(z) \ge 0$ and $a_2(z) \ge 0$ for all $z \in \Re$ as they are expectations of nonnegative quantities. (We work with $a_1(z)$ and $a_2(z)$ for arbitrary z rather than specifically $z = \xi$ so that we can also analyze an approximate variance, which replaces the true quantile $\xi \equiv \xi_m$ with an approximation $\breve{\xi} \equiv mQ_0(\theta_\star)$, which [1] shows satisfies $(\breve{\xi} - \xi_m)/m \to 0$ as $m \to \infty$, and we can further strengthen this to $(\breve{\xi} - \xi_m)/\sqrt{m} \to 0$.) While (6) expresses $\chi^2_{\mathrm{IS}(\theta_\star)}$ as a *difference* of two nonnegative terms, (10) instead gives $\chi^2_{\mathrm{IS}(\theta_\star)}$ as a *sum* of two nonnegative terms, which may be better for numerical computations to result in $\chi^2_{\mathrm{IS}(\theta_\star)}$ being nonnegative. We next analyze separately the two terms $a_1(z)$ and $a_2(z)$.

For $a_1(z)$ in (11), note that $I(c(\mathbf{X}) > z) = 0$ when $c(\mathbf{X}) \le z$, so

$$
a_1(z) = E_{\widetilde{G}_{\theta_\star}}\left[\left(-\overline{F}(z)\right)^2; c(\mathbf{X}) \le z\right] = \overline{F}^2(z) E_{\widetilde{G}_{\theta_\star}}[I(c(\mathbf{X}) \le z)] = \overline{F}^2(z)\widetilde{F}_{\theta_\star}(z) \tag{13}
$$

by (7). Again, it is clear that $a_1(z) \ge 0$ for all $z \in \Re$ as it is the product of probabilities.

For $a_2(z)$ in (12), we have $I(c(\mathbf{X}) > z) = 1$ when $c(\mathbf{X}) > z$, and (3) shows that the likelihood ratio $L_{\theta_\star}(\mathbf{X})$ depends on $\mathbf{X}$ only through the sum $Y = c(\mathbf{X})$. Thus, abusing notation by writing $L_{\theta_\star}(Y)$ instead of $L_{\theta_\star}(\mathbf{X})$ for $Y = c(\mathbf{X})$, we get

$$
\begin{aligned}
a_2(z) &= E_{\widetilde{G}_{\theta_\star}}\left[\left(L_{\theta_\star}(\mathbf{X}) - \overline{F}(z)\right)^2; c(\mathbf{X}) > z\right] \\
&= E_{\widetilde{G}_{\theta_\star}}\left[\left(L_{\theta_\star}(Y) - \overline{F}(z)\right)^2; Y > z\right] \\
&= \int_z^\infty \left(L_{\theta_\star}(y) - \overline{F}(z)\right)^2 \mathrm{d}\widetilde{F}_{\theta_\star}(y).
\end{aligned} \tag{14}
$$

Now we can use numerical quadrature to compute $a_2(z)$ by (14). The key to the representation in (14) is that the integrand is always nonnegative, so numerical quadrature might avoid numerical problems. (There still may be numerical issues because $L_{\theta_\star}(y)$ and $\overline{F}(z)$ can be of vastly different orders of magnitude, so we may lose precision in computing their difference. But at least we square the difference, making the integrand still always nonnegative, so numerical computations should result in $a_2(z)$ always being nonnegative for all $z \in \Re$.) When computing (14), we should **not** expand $(L_{\theta_\star}(y) - \overline{F}(z))^2 = L^2_{\theta_\star}(y) + \overline{F}^2(z) - 2L_{\theta_\star}(y)\overline{F}(z)$ because that would lead to integrands with different signs, destroying the property that may avoid numerical issues. We can further write (14) as

$$a_2(z) = \int_z^\infty \left(e^{mQ_0(\theta_\star)-\theta_\star y} - \overline{F}(z)\right)^2 \mathrm{d}\widetilde{F}_{\theta_\star}(y) \tag{15}$$

by (3) with $\theta = \theta_\star$. Again, note that $a_2(z) \geq 0$ for all $z \in \Re$ because the integrand in (15) is always nonnegative.

Putting all of this together, we can numerically compute $\chi^2_{\mathrm{IS}(\theta_\star)} = a_1(\xi) + a_2(\xi)$ using (10), (13), and (15). Because $a_1(\xi) \geq 0$ and $a_2(\xi) \geq 0$, the numerical computation should produce $\chi^2_{\mathrm{IS}(\theta_\star)} \geq 0$, as desired.

We also approximate the $\mathrm{IS}(\theta_\star)$ quantile estimator's asymptotic variance $\kappa^2_{\mathrm{IS}(\theta_\star)}$ in (5) by

$$\breve{\kappa}^2_{\mathrm{IS}(\theta_\star)} \equiv \frac{\breve{\chi}^2_{\mathrm{IS}(\theta_\star)}}{\breve{f}^2(\breve{\xi})}, \quad \text{with} \tag{16}$$

$$\breve{\chi}^2_{\mathrm{IS}(\theta_\star)} \equiv \mathrm{Var}_{\widetilde{G}_{\theta_\star}}[L_{\theta_\star}(\mathbf{X}) I(c(\mathbf{X}) > \breve{\xi})] = E_{\widetilde{G}_{\theta_\star}}[L^2_{\theta_\star}(\mathbf{X}) I(c(\mathbf{X}) > \breve{\xi})] - \overline{F}^2(\breve{\xi}), \tag{17}$$

where $\breve{f}(\cdot)$ is the saddlepoint approximation to the true density $f(\cdot)$, and the last term in (17) is $\overline{F}^2(\breve{\xi})$ (rather than $(1-p)^2$) to ensure that $\breve{\chi}^2_{\mathrm{IS}(\theta_\star)}$ is nonnegative and a variance, which is needed to handle the covariance for the EC estimator by the Cauchy-Schwarz inequality. We then can numerically compute $\breve{\chi}^2_{\mathrm{IS}(\theta_\star)} = a_1(\breve{\xi}) + a_2(\breve{\xi})$ using (13) and (15) with $z = \breve{\xi}$. Because $a_1(\breve{\xi}) \geq 0$ and $a_2(\breve{\xi}) \geq 0$, numerical computation should produce $\breve{\chi}^2_{\mathrm{IS}(\theta_\star)} \geq 0$.

3 Covariance

For the $\mathrm{IS}(\theta_\star)$ estimator of the EC η, we need to also compute the (asymptotic) covariance $\gamma_{\mathrm{IS}(\theta_\star)}$ and its approximation $\breve{\gamma}_{\mathrm{IS}(\theta_\star)}$ arising from simultaneously estimating the quantile ξ and mean μ with $\mathrm{IS}(\theta_\star)$, where

$$\begin{aligned}
\gamma_{\mathrm{IS}(\theta_\star)} &= \mathrm{Cov}_{\widetilde{G}_{\theta_\star}}[I(c(\mathbf{X}) > \xi) L_{\theta_\star}(\mathbf{X}), c(\mathbf{X}) L_{\theta_\star}(\mathbf{X})], \\
\breve{\gamma}_{\mathrm{IS}(\theta_\star)} &= \mathrm{Cov}_{\widetilde{G}_{\theta_\star}}[I(c(\mathbf{X}) > \breve{\xi}) L_{\theta_\star}(\mathbf{X}), c(\mathbf{X}) L_{\theta_\star}(\mathbf{X})].
\end{aligned}$$

To handle both simultaneously, define $\gamma_{\mathrm{IS}(\theta_\star)}(z) = \mathrm{Cov}_{\widetilde{G}_{\theta_\star}}[I(c(\mathbf{X}) > z) L_{\theta_\star}(\mathbf{X}), c(\mathbf{X}) L_{\theta_\star}(\mathbf{X})]$, so we want to compute $\gamma_{\mathrm{IS}(\theta_\star)}(z)$ for $z = \xi$ and $z = \breve{\xi}$. We now use the representation $\mathrm{Cov}[S, T] = E[(S - E[S])(T - E[T])]$ to obtain an expression for $\gamma_{\mathrm{IS}(\theta_\star)}(z)$ that may be amenable to numerical computation via quadrature. In our setting, we have that $E_{\widetilde{G}_{\theta_\star}}[I(c(\mathbf{X}) > z) L_{\theta_\star}(\mathbf{X})] = E_G[I(c(\mathbf{X}) > z)] = \overline{F}(z)$ and $E_{\widetilde{G}_{\theta_\star}}[c(\mathbf{X}) L_{\theta_\star}(\mathbf{X})] = E_G[c(\mathbf{X})] = \mu$, so

$$\begin{aligned}
\gamma_{\mathrm{IS}(\theta_\star)}(z) &= E_{\widetilde{G}_{\theta_\star}}\left[\big(I(c(\mathbf{X}) > z) L_{\theta_\star}(\mathbf{X}) - \overline{F}(z)\big)\big(c(\mathbf{X}) L_{\theta_\star}(\mathbf{X}) - \mu\big)\right] \\
&= E_{\widetilde{G}_{\theta_\star}}\left[\big(I(Y > z) L_{\theta_\star}(Y) - \overline{F}(z)\big)\big(Y L_{\theta_\star}(Y) - \mu\big)\right] = b_1(z) + b_2(z),
\end{aligned} \tag{18}$$

where

$$b_1(z) = E_{\widetilde{G}_{\theta_\star}}\left[\big(I(Y>z)L_{\theta_\star}(Y) - \overline{F}(z)\big)\big(YL_{\theta_\star}(Y) - \mu\big);\, Y \le z\right], \tag{19}$$

$$b_2(z) = E_{\widetilde{G}_{\theta_\star}}\left[\big(I(Y>z)L_{\theta_\star}(Y) - \overline{F}(z)\big)\big(YL_{\theta_\star}(Y) - \mu\big);\, Y > z\right]. \tag{20}$$

We first work out an expression for $b_1(z)$ in (19):

$$\begin{aligned}
b_1(z) &= -\overline{F}(z)E_{\widetilde{G}_{\theta_\star}}\left[YL_{\theta_\star}(Y) - \mu;\, Y \le z\right] \\
&= -\overline{F}(z)E_{\widetilde{G}_{\theta_\star}}\left[YL_{\theta_\star}(Y)I(Y \le z)\right] + \overline{F}(z)E_{\widetilde{G}_{\theta_\star}}\left[\mu I(Y \le z)\right] && (21) \\
&= -\overline{F}(z)E_G\left[YI(Y \le z)\right] + \overline{F}(z)\mu E_{\widetilde{G}_{\theta_\star}}\left[I(Y \le z)\right] && (22) \\
&= -\overline{F}(z)\int_{-\infty}^{z} y\,\mathrm{d}F(y) + \overline{F}(z)\mu\widetilde{F}_{\theta_\star}(z) && (23)
\end{aligned}$$

by (7), and we can evaluate (23) numerically. For $b_2(z)$ in (20), we have that

$$\begin{aligned}
b_2(z) &= E_{\widetilde{G}_{\theta_\star}}\left[\big(L_{\theta_\star}(Y) - \overline{F}(z)\big)\big(YL_{\theta_\star}(Y) - \mu\big);\, Y > z\right] \\
&= \int_{z}^{\infty}\big(L_{\theta_\star}(y) - \overline{F}(z)\big)\big(yL_{\theta_\star}(y) - \mu\big)\,\mathrm{d}\widetilde{F}_{\theta_\star}(y), && (24)
\end{aligned}$$

which we can evaluate numerically. Finally, we evaluate $\gamma_{\mathrm{IS}(\theta_\star)}(z)$ by puttting (23) and (24) into (18).

3.1 Numerical Problems

Note that in (22), we applied a change of measure to the first term (but not to the second term). The reason for this is that if we don't apply a change of measure to the first term, then Matlab has numerical problems. Specifically, by (21), we could have instead expressed the first term in (23) as

$$-\overline{F}(z)E_{\widetilde{G}_{\theta_\star}}\left[YL_{\theta_\star}(Y)I(Y \le z)\right] = -\overline{F}(z)\int_{-\infty}^{z} yL_{\theta_\star}(y)\,\mathrm{d}\widetilde{F}_{\theta_\star}(y).$$

But in my numerical computations with this representation, Matlab issues a warning: `Infinite or Not-a-Number value encountered`, and gives no results. It seems that the problem arises in trying to evaluate $yL_{\theta_\star}(y)$ over the range $y \in (-\infty, z]$. I do not get the Matlab warning when using the representation for the first term in (23). Also, I do not get the Matlab warning for (24), so it seems that evaluating $yL_{\theta_\star}(y)$ over the range $y \in [z, \infty)$ is not a problem. Thus, in all numerical computations, we avoid computing integrals over the range $y \in (-\infty, z]$ with $yL_{\theta_\star}(y)$ appearing in the integrand.

We can also express the covariance $\gamma_{\mathrm{IS}(\theta_\star)}(z)$ as

$$\begin{aligned}
\gamma_{\mathrm{IS}(\theta_\star)}(z) &= E_{\widetilde{G}_{\theta_\star}}\left[I(Y>z)L_{\theta_\star}(Y)YL_{\theta_\star}(Y)\right] - E_{\widetilde{G}_{\theta_\star}}\left[I(Y>z)L_{\theta_\star}(Y)\right]E_{\widetilde{G}_{\theta_\star}}\left[YL_{\theta_\star}(Y)\right] \\
&= E_G\left[I(Y>z)YL_{\theta_\star}(Y)\right] - E_G\left[I(Y>z)\right]E_G\left[Y\right] \\
&= \int_{z}^{\infty} yL_{\theta_\star}(y)\,\mathrm{d}F(y) - \overline{F}(z)\mu. && (25)
\end{aligned}$$

In my numerical computations for $G_0 = N(0,1)$, Matlab gives consistent results when $m \le 16$ for the representations in (25) and by puttting (23) and (24) into (18). But for $m = 32$, I get some discrepancy: about 60% difference for $z = \xi$, and about 0.8% difference for $z = \breve{\xi}$. I'm not sure if the numerical results from (25) or (18) is more accurate. But because of the numerical problems of using (6) when computing the variance, it is probably better to compute the covariance $\gamma_{\mathrm{IS}(\theta_\star)}(z)$ by using (23) and (24) in (18) rather than using (25). (When I numerically compute variance instead of covariance, it is obvious that (6) is problematic numerically because I get negative values for the variance when $m \ge 16$. But a covariance can be positive or negative, so there is no obvious evidence that one numerical computation is more accurate.) In any case, this may not be that much of an issue because when computing the asymptotic variance of the EC estimator, the Cauchy-Schwarz inequality implies that the covariance term $\gamma_{\mathrm{IS}(\theta_\star)}(\xi)$ will be nondominant compared to the larger of asymptotic variances of the $\mathrm{IS}(\theta_\star)$ estimators of ξ and μ, the larger of which will be for the estimator of μ.

4 Specific Choices for G_0

Example 1. Suppose that each summand $X_j \sim G_0$, where G_0 is the $N(\mu_0, \sigma_0^2)$ CDF, so $G_0(x) = \Phi\left(\frac{x-\mu_0}{\sigma_0}\right)$, with Φ as the $N(0,1)$ CDF. Let ϕ be the $N(0,1)$ density function. The CGF of G_0 is $Q_0(\theta) = \mu_0\theta + (\sigma_0^2\theta^2/2)$, so $Q_0'(\theta) = \mu_0 + \sigma_0^2\theta$ and $Q_0''(\theta) = \sigma_0^2$. Therefore, (4) becomes $-\theta_\star(\mu_0 + \sigma_0^2\theta_\star) + \mu_0\theta_\star + (\sigma_0^2\theta_\star^2/2) = -\beta$, which is a quadratic equation in $\theta_\star$, leading to $\theta_\star = \pm\sqrt{2\beta}/\sigma_0$, and we take the positive root: $\theta_\star = \sqrt{2\beta}/\sigma_0$. [1]When $\mathbf{X} \sim G$, the sum $Y = \sum_{j=1}^m X_j \sim N(m\mu_0, m\sigma_0^2)$ has CDF $F(y) = \Phi\left(\frac{y-m\mu_0}{\sigma_0\sqrt{m}}\right)$ under SRS, and the $\mathrm{IS}(\theta_\star)$ CDF of Y is $\widetilde{F}_{\theta_\star}(y) = \Phi\left(\frac{y-\breve{\xi}}{\sigma_0\sqrt{m}}\right)$, where $\breve{\xi} = mQ_0'(\theta_\star) = m(\mu_0 + \sigma_0^2\theta_\star)$. Thus, (13) becomes

$$a_1(z) = \overline{F}^2(z)\widetilde{F}_{\theta_\star}(z) = \left[1 - \Phi\left(\frac{z-\mu_0 m}{\sigma_0\sqrt{m}}\right)\right]^2 \Phi\left(\frac{z-\breve{\xi}}{\sigma_0\sqrt{m}}\right),$$

[1]For the IS mean estimator, its variance $\sigma^2_{\mathrm{IS}(\theta)} = m[\alpha(\theta)]^m\Big(m[Q_0'(-\theta)]^2 + Q_0''(-\theta)\Big) - (m\mu_0)^2$ when $m \ge 2$. We always have that $Q_0''(-\theta) > 0$ and $\alpha(\theta) \equiv M_0(\theta)M_0(-\theta) \ge 1$, with $\alpha(\theta) > 1$ for $\theta \ne 0$. Thus, $\sigma^2_{\mathrm{IS}(\theta)}$ grows at least at rate $m[\alpha(\theta)]^m$, and it grows at the faster rate $m^2[\alpha(\theta)]^m$ if $Q_0'(-\theta) \ne 0$. It is possible to have $Q_0'(-\theta) = 0$. For example, if G_0 is $N(\mu_0, \sigma_0^2)$, then $Q_0'(-\theta) = 0$ for $-\theta = \mu_0/\sigma_0^2$.

which can be used to get $a_1(\xi)$ and $a_1(\breve{\xi})$, where for $z = \breve{\xi}$, we have that $\Phi\left(\frac{z-\breve{\xi}}{\sigma_0\sqrt{m}}\right) = \Phi(0) = 1/2$. For $a_2(z)$ as in (15), let $\overline{\Phi}(y) = 1 - \Phi(y)$, so $\overline{F}(y) = \overline{\Phi}\left(\frac{y-\mu_0 m}{\sigma_0\sqrt{m}}\right)$, and we get

$$\begin{aligned}
a_2(z) &= \int_z^\infty \left[e^{mQ_0(\theta_\star)-\theta_\star y} - \overline{\Phi}\left(\frac{z-\mu_0 m}{\sigma_0\sqrt{m}}\right)\right]^2 \mathrm{d}\Phi\left(\frac{y-\breve{\xi}}{\sigma_0\sqrt{m}}\right) \\
&= \int_z^\infty \left[e^{mQ_0(\theta_\star)-\theta_\star y} - \overline{\Phi}\left(\frac{z-\mu_0 m}{\sigma_0\sqrt{m}}\right)\right]^2 \phi\left(\frac{y-\breve{\xi}}{\sigma_0\sqrt{m}}\right)\frac{\mathrm{d}y}{\sigma_0\sqrt{m}} \\
&= \int_z^\infty \left[e^{mQ_0(\theta_\star)-\theta_\star y} - \overline{\Phi}\left(\frac{z-\mu_0 m}{\sigma_0\sqrt{m}}\right)\right]^2 \frac{1}{\sigma_0\sqrt{2\pi m}}\exp\left[-\frac{1}{2}\left(\frac{y-\breve{\xi}}{\sigma_0\sqrt{m}}\right)^2\right]\mathrm{d}y \qquad (26)
\end{aligned}$$

to obtain $a_2(\xi)$ and $a_2(\breve{\xi})$. In numerically computing (26), we should use numerical quadrature, e.g., with `scipy.integrate` in Python, possibly replacing the upper limit ∞ of the integral by some value $y_0 \gg \breve{\xi}$ or $y_0 \gg \xi$. All of the numerical computations should be carried out with as high precision as possible (double or preferably quad precision). We then use $a_1(\xi)$ and $a_2(\xi)$ in (10) to obtain the numerical value of $\chi^2_{\mathrm{IS}(\theta_\star)}$ in (6), which is the numerator in (5) of the asymptotic variance of the IS$(\theta_\star)$ estimator of ξ. We can similarly use (23) and (24) in (18) to numerically compute the IS$(\theta_\star)$ covariance $\gamma_{\mathrm{IS}(\theta_\star)}(z)$.

Example 2. Now suppose that each summand $X_j \sim G_0$, where G_0 is the Erlang CDF with $s \geq 1$ stages, each having mean $1/\lambda$, so each summand has CDF $G_0(x) = \int_0^x g_0(t)\,\mathrm{d}t$, with density $g_0(t) = \frac{\lambda^s t^{s-1} e^{-\lambda t}}{(s-1)!}$ for $t \geq 0$. The MGF of G_0 is $M_0(\theta) = [\lambda/(\lambda-\theta)]^s$ for $\theta < \lambda$, so its CGF and its derivatives are

$$Q_0(\theta) = s[\ln(\lambda) - \ln(\lambda-\theta)], \quad Q_0'(\theta) = \frac{s}{\lambda-\theta}, \quad Q_0''(\theta) = \frac{s}{(\lambda-\theta)^2}. \qquad (27)$$

Therefore, (4) becomes

$$-\theta\frac{s}{\lambda-\theta} + s[\ln(\lambda) - \ln(\lambda-\theta)] = -\beta,$$

whose root $\theta = \theta_\star$ in general needs to be solved for numerically using a root-finding approach, e.g., Newton's method or the false-position method. When $\mathbf{X} \sim G$, the sum $Y = \sum_{j=1}^m X_j$ is Erlang with ms stages, each with mean $1/\lambda$, which has

$$\text{CDF} \quad F(y) = \int_0^y f(t)\,\mathrm{d}t \quad \text{and density} \quad f(t) = \frac{\lambda^{ms} t^{ms} e^{-\lambda t}}{(ms-1)!} I(t \geq 0). \qquad (28)$$

Under IS$(\theta_\star)$, the sum Y is Erlang with ms stages, each with mean $1/(\lambda-\theta_\star)$, which has

$$\text{CDF} \quad \widetilde{F}_{\theta_\star}(y) = \int_0^y \widetilde{f}_{\theta_\star}(t)\,\mathrm{d}t \quad \text{and density} \quad \widetilde{f}_{\theta_\star}(t) = \frac{(\lambda-\theta_\star)^s t^{s-1} e^{-(\lambda-\theta_\star)t}}{(s-1)!} I(t \geq 0). \qquad (29)$$

Thus, we can use (27), (28), and (29) in (13) and (15) to compute $a_1(z)$ and $a_2(z)$. We then use $a_1(\xi)$ and $a_2(\xi)$ in (10) to obtain the numerical value of $\chi^2_{\mathrm{IS}(\theta_\star)}$ in (6), which is the numerator in (5) of the asymptotic variance of the IS$(\theta_\star)$ estimator of ξ. We can similarly use (23) and (24) in (18) to numerically compute the IS$(\theta_\star)$ covariance $\gamma_{\mathrm{IS}(\theta_\star)}(z)$.

References

[1] P. W. Glynn. Importance sampling for Monte Carlo estimation of quantiles. In S. M. Ermakov and V. B. Melas, editors, *Mathematical Methods in Stochastic Simulation and Experimental Design: Proceedings of the 2nd St. Petersburg Workshop on Simulation*, pages 180–185, St. Petersburg, Russia, 1996. Publishing House of St. Petersburg Univ.

[2] Y. Li, Z. T. Kaplan, and M. K. Nakayama. Github repository: Monte Carlo methods for economic capital, 2023.

[3] Y. Li, Z. T. Kaplan, and M. K. Nakayama. Monte Carlo methods for economic capital. *INFORMS Journal on Computing*, 2023. Accepted.