Available online at www.sciencedirect.com
SCIENCE DIRECT®
European Journal of Operational Research 155 (2004) 154–169
EUROPEAN JOURNAL OF OPERATIONAL RESEARCH
www.elsevier.com/locate/dsw

Stochastics and Statistics

Estimation of means and covariances of inverse-Gaussian order statistics

Huifen Chen [a,*], Kuo-Hwa Chang [a], Liuying Cheng [b]

[a] *Department of Industrial Engineering, Chung-Yuan Christian University, Chung Li 320, Taiwan*
[b] *Department of Industrial Engineering, Dayeh University, Chang Hwa 515, Taiwan*

Received 30 November 2001; accepted 3 September 2002

Abstract

We propose a simulation algorithm to estimate means, variances, and covariances for a set of order statistics from inverse-Gaussian (IG) distributions. Given a set of Monte Carlo data, the algorithm estimates these values simultaneously. Two types of control variates are used: internal uniform and external exponential. Simulation results show that exponential control variates work better, best when the IG skewness is near the exponential skewness value 2. Either type of control variate provides substantial variance reduction for IG distributions that have low skewness.
© 2003 Elsevier B.V. All rights reserved.

Keywords: Simulation; BLUE; Control variates; Order statistics; Variance reduction

1. Introduction

We consider the problem of computing the mean vector and covariance matrix of a set of order statistics from the inverse-Gaussian (IG) population. Let $(X_1, X_2, \ldots, X_n)$ denote a random sample of size n from the IG distribution with mean μ and standard deviation σ. The n observations ranked in increasing order, denoted $\{X_{1:n} \leqslant X_{2:n} \leqslant \cdots \leqslant X_{n:n}\}$, form a set of IG order statistics. When the IG population mean μ and standard deviation σ are unknown but the skewness k is known, we can use the order statistics to compute the best linear unbiased estimates (BLUEs) of μ and σ. The BLUEs of μ and σ are linear combinations of $X_{1:n}, \ldots, X_{n:n}$, where the coefficients depend on the mean vector and covariance matrix of the standardized IG order statistics $Z_{1:n}, Z_{2:n}, \ldots, Z_{n:n}$, where $Z_{i:n} = (X_{i:n} - \mu)/\sigma$ for $i = 1, \ldots, n$.

For this application, we propose algorithms to compute the mean vector and covariance matrix of a set of order statistics from a standardized IG population with known skewness. Given the IG mean μ, variance σ, and skewness k, it is easy to implement these algorithms to compute the mean vector and covariance matrix for non-standardized IG order statistics in applications such as estimation of trimmed means or

* Corresponding author.
E-mail address: huifen@cycu.edu.tw (H. Chen).

0377-2217/$ - see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/S0377-2217(02)00821-4

quantiles. Notice that $Z_{i:n} = (X_{i:n} - \mu)/\sigma$ and hence, $E(X_{i:n}) = \mu + \sigma E(Z_{i:n})$ and $Cov(X_{i:n}, X_{j:n}) = \sigma^2 Cov(Z_{i:n}, Z_{j:n})$ for $1 \leqslant i, j \leqslant n$. Therefore, algorithms for standardized IG are sufficient for non-standardized IG.

A traditional approach for computing moments of order statistics is numerical integration. Balakrishnan and Chen (1997) use Gaussian quadrature methods (Press et al., 1997) to compute the mean vector and covariance matrix of a set of standardized IG order statistics with a properly determined upper limit of integration. For means and variances, they use 512 abscissae; for covariances, they use 96 abscissae. For $n = 1(1)25$ and $k = 0(0.1)2.5$, they provide tables of means, variances, and covariances ($n \geqslant 2$). Another example is Parrish (1992a,b), who also uses Gaussian quadrature methods to compute the mean vector and covariance matrix for normal order statistics.

A Monte Carlo approach has two advantages over numerical integration. The first is that they can estimate all constants—means, variances, and covariances—simultaneously using only one set of samples. (We assume that computations of all constants are necessary for purposes such as constructing BLUEs.) Moreover, the control-variate variance-reduction technique can be used to increase the computational efficiency. Numerical integration methods are therefore less efficient because they compute each constant separately. Since there is a total of $2n$ one-dimensional integrations (for means and variances) and $n(n-1)/2$ two-dimensional integrations (for covariances), the Monte Carlo approach may be more efficient when n is large. The second advantage of the Monte Carlo approach is the robustness of Monte Carlo convergence with respect to the IG skewness k. When k is large, the IG distribution shape is steep and hence the integrands for computing means and covariances (Eqs. (7)–(9) in Section 2) are not smooth. Gaussian quadrature methods may have large numerical error in this case (e.g., $k = 50$). Other numerical integration methods such as the equally-spaced-abscissae methods (e.g., Simpson's rule; see Section 2) converge as the number of abscissae increases to infinity. However, when k is large, the equally-spaced-abscissae methods are very inefficient. On the other hand, when n and k are small, numerical approaches can be quite accurate and efficient.

The rest of this paper is organized as follows. In Section 2, we introduce IG distributions, IG order statistics, and BLUEs of the population mean and variance. In Section 3, we propose a Monte Carlo method for estimating the mean vector and covariance matrix of a set of standardized IG order statistics. Two kinds of control variates are used to reduce the variances of crude estimates. In Section 4, we investigate the accuracy and variance-reduction effects of control-variate estimates.

2. Inverse-Gaussian distributions

Schrödinger (1915) and Smoluchowski (1915) derived, by a different method, the distribution function of the first passage time of Brownian motion with positive drift (denoted as TBMP distribution by Wasan, 1968; see Seshadri, 1993). Tweedie (1945) established the inverse relationship between the cumulant generating functions of the TBMP and Gaussian distributions and established the alternate name the inverse-Gaussian distribution. Independent of Tweedie's work, Wald (1947, p. 193) derived the IG distribution (as shown in Eq. (1) with $\delta = 1$) as the limiting distribution of the sample size in a sequential probability ratio test. Therefore, the IG distribution is also called the standard Wald distribution. In two important works, Tweedie (1957a,b) profiled the statistical properties of IG distributions.

IG distributions have wide uses in reliability. Since the first passage time of Brownian motion has an IG distribution, IG distributions have been used as probability models of product/device lifetime, for example, the lifetime of deep-groove ball bearings (Lieblein and Zelen, 1956). Folks and Chhikara (1978) fitted an IG distribution to shelf-life data of food products and found that the IG is as reasonable a model as the lognormal or Weibull (Gacula and Kubala, 1975). Furthermore, the failure function for IG is non-monotonic, where it first increases and then decreases, approaching a constant as the lifetime goes to

infinity (Chhikara and Folks, 1977). This property makes the IG useful to model the repair time when the lifetime is dominated by the early occurrence of an event, e.g., the repair time for an airborne communication transceiver (Von Alven, 1964). The IG has a wide range of distribution shapes, from highly right skewed to almost normal. Sichel (1975) showed that the IG provides a good fit to the highly skewed observed distribution of word frequencies. Other applications of probability modeling include interpurchase times within individual households (Banerjee and Bhattacharyya, 1976), durations of labor strikes (Lancaster, 1972), employee service times (Whitmore, 1979), and cycle times for particles in the bloodstream (Wise et al., 1968; Wise, 1971; Wise, 1975). Gunes et al. (1997) discuss several goodness-of-fit tests for the IG distribution. Statistical properties and detailed discussions of the IG can be found in Chhikara and Folks (1989) and Johnson et al. (1994).

There are various forms of IG density functions. The two-parameter type given by Tweedie (1956) is

$$f(x|\delta,\lambda)=\sqrt{\frac{\lambda}{2\pi x^3}}\exp\left\{-\frac{\lambda(x-\delta)^2}{2\delta^2 x}\right\},\quad x>0,\ \delta>0,\ \lambda>0, \tag{1}$$

where δ is the scale parameter and $\phi=\lambda/\delta$ is the shape parameter (Chhikara and Folks, 1989, p. 9; Johnson et al., 1994, p. 268). The density function is unimodal with mode $x=\delta\left[\sqrt{1+9/(4\phi^2)}-3/(2\phi)\right]$ (Tweedie, 1957a). The cumulative distribution function (cdf) is

$$F(x|\delta,\lambda)=\Phi[(\lambda/x)^{1/2}(x/\delta-1)]+\mathrm{e}^{2\lambda/\delta}\Phi[-(\lambda/x)^{1/2}(x/\delta+1)], \tag{2}$$

where $\Phi(\ )$ is the standard normal cdf. The first four moments are

mean $\mu=\delta$,

variance $\sigma^2=\delta^3/\lambda$,

skewness $\sqrt{\beta_1}=3\sqrt{\delta/\lambda}$,

kurtosis $\beta_2=3+15\delta/\lambda$.

Since $\beta_2=3+5\beta_1/3$, this inverse Gaussian line lies between the gamma and lognormal lines in the Pearson(β_1,β_2) plane. As λ increases to infinity (with δ fixed), the standardized IG distribution approaches the standard normal distribution.

By adding a threshold value η, we obtain a three-parameter variant of the IG distribution with the mean shifted to $\delta+\eta$ and the other three moments unchanged. Following the notations used in Balakrishnan and Chen (1997), μ denotes the shifted IG mean $(\delta+\eta)$; σ^2, the variance δ^3/λ, and $k>0$, the skewness $3\sqrt{\delta/\lambda}$. Then the density function and cdf of the three-parameter IG, denoted IG(μ,σ,k), are

$$f(x|\mu,\sigma,k)=\frac{1}{\sqrt{2\pi}\sigma}\left[\frac{3\sigma}{3\sigma+k(x-\mu)}\right]^{3/2}\exp\left\{-\frac{3(x-\mu)^2}{2\sigma[3\sigma+k(x-\mu)]}\right\},\quad x>\mu-\frac{3\sigma}{k}, \tag{3}$$

and

$$F(x|\mu,\sigma,k)=\Phi\left[\frac{x-\mu}{\sigma}\left(1+\frac{k(x-\mu)}{3\sigma}\right)^{-1/2}\right]+\mathrm{e}^{18/k^2}\Phi\left[-\left(\frac{x-\mu}{\sigma}+\frac{6}{k}\right)\left(1+\frac{k(x-\mu)}{3\sigma}\right)^{-1/2}\right]. \tag{4}$$

The mode of IG(μ,σ,k) is

$$x_M=\mu-\frac{3\sigma}{k}+\frac{\sigma(\sqrt{36+k^4}-k^2)}{2k}. \tag{5}$$

In Section 3, we use the mode as an initial guess to generate the first order statistic by the inverse-transformation method. The remainder of this paper is devoted to three-parameter IGs. Application of the results to the two-parameter IGs is straightforward.

Order statistics are useful in parameter estimations. The BLUE is the best (in the sense of minimal variance) unbiased L-estimator; the L-estimator is defined as the linear combination of the order statistics. In the case of $\mathrm{IG}(\mu,\sigma,k)$, constructing BLUEs for μ, σ, and k is difficult (see Cohen and Whitten, 1988). However, when the skewness k is fixed, the BLUEs of the location parameter μ and scale parameter σ can be easily derived (Arnold et al., 1992, p. 172). Let $\mathbf{X}=(X_{1:n},\ldots,X_{n:n})'$ be a set of order statistics from $\mathrm{IG}(\mu,\sigma,k)$. The standardized order statistic $\mathbf{Z}=(Z_{1:n},\ldots,Z_{n:n})'$ is defined by $Z_{i:n}=(X_{i:n}-\mu)/\sigma$. Let $\boldsymbol{\mu}=(\mu_{i:n})_{i=1}^{n}$ and $\boldsymbol{\Sigma}=(\sigma_{i,j:n})_{i,j=1}^{n}$ denote the mean vector and covariance matrix of $\mathbf{Z}$, respectively. Given an observed $\mathbf{X}$, the BLUEs of μ and σ are

$$\hat{\mu}=-\boldsymbol{\mu}'\boldsymbol{\Gamma}\mathbf{X} \quad \text{and} \quad \hat{\sigma}=\mathbf{1}'\boldsymbol{\Gamma}\mathbf{X}, \tag{6}$$

where matrix $\boldsymbol{\Gamma}=\boldsymbol{\Sigma}^{-1}(\mathbf{1}\boldsymbol{\mu}'-\boldsymbol{\mu}\mathbf{1}')\boldsymbol{\Sigma}^{-1}/\Delta$, $\mathbf{1}$ is an $n\times 1$ vector with all elements being 1, and constant $\Delta=(\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu})(\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1})-(\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\mathbf{1})^2$. These BLUEs are linear combinations of order statistics where the coefficients are functions of $\mu_{i:n}$ and $\sigma_{i,j:n}$.

To construct BLUEs for μ and σ, the values of $\mu_{i:n}$ and $\sigma_{i,j:n}$ are needed. Let $f(z)$ denote the standardized IG density function; $F(z)$, the standardized IG cdf; $f_{i:n}(z)$, the density function of $Z_{i:n}$; and $f_{i,j:n}(z_i,z_j)$, the joint density of $Z_{i:n}$ and $Z_{j:n}$. Then for $i=1,\ldots,n$,

$$\mu_{i:n}=\mathrm{E}(Z_{i:n})=\int_{-3/k}^{\infty} z f_{i:n}(z)\,\mathrm{d}z=\int_{-3/k}^{\infty}\frac{n!}{(i-1)!(n-i)!}z[F(z)]^{i-1}[1-F(z)]^{n-i}f(z)\,\mathrm{d}z \tag{7}$$

and

$$\sigma_{i,i:n}=\mathrm{V}(Z_{i:n})=\int_{-3/k}^{\infty} z^2 f_{i:n}(z)\,\mathrm{d}z-\mu_{i:n}^2=\int_{-3/k}^{\infty}\frac{n!}{(i-1)!(n-i)!}z^2[F(z)]^{i-1}[1-F(z)]^{n-i}f(z)\,\mathrm{d}z-\mu_{i:n}^2, \tag{8}$$

and for $1\leqslant i<j\leqslant n$,

$$\begin{aligned}\sigma_{i,j:n}&=\mathrm{Cov}(Z_{i:n},Z_{j:n})=\int_{-3/k}^{\infty}\int_{z_i}^{\infty} z_i z_j f_{i,j:n}(z_i,z_j)\,\mathrm{d}z_j\,\mathrm{d}z_i-\mu_{i:n}\mu_{j:n},\\ &=\int_{-3/k}^{\infty}\int_{z_i}^{\infty}\frac{n!}{(i-1)!(j-i-1)!(n-j)!}z_i z_j[F(z_i)]^{i-1}[1-F(z_j)]^{n-j}[F(z_j)-F(z_i)]^{j-i-1}f(z_i)f(z_j)\,\mathrm{d}z_j\,\mathrm{d}z_i\\ &\quad-\mu_{i:n}\mu_{j:n}.\end{aligned} \tag{9}$$

Functions $f(z)$ and $F(z)$ can be obtained from Eqs. (3) and (4), respectively, by setting $\mu=0$ and $\sigma=1$.

There are two numerical integration approaches for calculating $\mu_{i:n}$ and $\sigma_{i,i:n}$ (both one-dimensional integrals) and $\sigma_{i,j:n}$ (a two-dimensional integral). Numerical integration methods in each approach are well known (Conte and de Boor, 1980, p. 303) and their published computer code is available (e.g., Press et al., 1997). The first approach, the equally-spaced-abscissae approach, evaluates the integrand at a set of N equally spaced abscissae. Numerical integration methods such as midpoint, trapezoid rule, and Simpson's rule belong to the first approach. Though these methods converge as N goes to infinity, the convergence is slow when the magnitude of the integrand is high only in a small region but almost zero everywhere. By comparison, the second approach, the Gaussian quadrature method, is based on N unequally spaced abscissae chosen so that the approximation is exact (assuming no numerical error) for polynomial integrands of order $\leqslant(2N-1)$. The efficiency is twice that of the equally-spaced-abscissae approach. However, bigger N does not guarantee more accuracy; when the integrand is not smooth, Gaussian quadrature may have large numerical errors even for large N. In our application, when the IG skewness k is high (e.g., $k=50$)

and the order i is small, the integrands in Eqs. (7)–(9) are steeply shaped, causing significant Gaussian-quadrature-integration errors. For example, when $n = 10$, $k = 50$ and $i = 1$, the first IG order statistic mean is about -0.059923 (computed by simulation with standard error $<10^{-7}$) but the mean value computed by Gaussian quadrature (with Laguerre abscissae) is 0. As the order i increases, the integrand becomes smoother and the numerical error decreases. For example, when i changes from 1 to 10, the mean value (which changes to 0.504) is accurately computed by Gaussian quadrature. The computational complexity of both numerical integration approaches increases quadratically in n. This is because there are $2n$ one-dimensional integrals and $n(n-1)/2$ two-dimensional integrals, and usually the latter dominates the computation time.

The Monte Carlo estimation approach used in this research has approximately linear complexity. This approach generates m realizations of the order statistics $\{Z_{1:n}, Z_{2:n}, \ldots, Z_{n:n}\}$ and uses them to estimate all values in the mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. Two computation steps are required: generation of observations and computation of estimates (e.g., Eq. (10)). The time required is linear in n for the first step; quadratic in n for the second. Since the first step usually dominates (unless n is very large), the computational complexity of the simulation approach increases approximately linearly in n. To increase the simulation efficiency, the control-variate technique is used to reduce the variance of the crude estimates. We describe our implementation of the Monte Carlo approach in Section 3.

3. Methods

We propose a Monte Carlo approach to estimate the mean vector and covariance matrix for a set of standardized IG order statistics $\{Z_{1:n}, Z_{2:n}, \ldots, Z_{n:n}\}$. The crude estimates are adjusted by control variates for variance reduction. Two kinds of control variates are considered: sample moments of uniform order statistics and sample moments of exponential order statistics. In this section, we discuss the crude estimates, the generation of IG observations, and the two kinds of control variates.

Let $\{z_{1:n}^{(b)}, \ldots, z_{n:n}^{(b)} : b = 1, \ldots, m\}$ denote m realizations of the standardized IG order statistics $Z_{1:n}, \ldots, Z_{n:n}$ shown in Fig. 1. Given these observations, we can calculate crude, unbiased estimates of the means $\mu_{i:n} = \mathrm{E}(Z_{i:n})$, variances $\sigma_{i,i:n} = \mathrm{V}(Z_{i:n})$, and covariances $\sigma_{i,j:n} = \mathrm{Cov}(Z_{i:n}, Z_{j:n})$ using

$$\begin{aligned}
\hat{\mu}_{i:n} &= \bar{z}_{i:n} = \sum_{b=1}^{m} z_{i:n}^{(b)}/m, \\
\hat{\sigma}_{i,i:n} &= s_{i,i:n} = \sum_{b=1}^{m} (z_{i:n}^{(b)} - \bar{z}_{i:n})^2/(m-1), \\
\hat{\sigma}_{i,j:n} &= s_{i,j:n} = \sum_{b=1}^{m} (z_{i:n}^{(b)} - \bar{z}_{i:n})(z_{j:n}^{(b)} - \bar{z}_{j:n})/(m-1).
\end{aligned} \tag{10}$$

The Monte Carlo estimation approach requires generation of standardized IG order statistics. Michael et al. (1976) propose a chi-square method to generate IG observations. Suppose that a random variable X has a two-parameter IG distribution as in Eq. (2). Then the transformed variable $Y^2 = \lambda(X-\delta)^2/(\delta^2 X)$ has

	Order 1	⋯	Order i	⋯	Order j	⋯	Order n
1st sample:	$z_{1:n}^{(1)}$	⋯	$z_{i:n}^{(1)}$	⋯	$z_{j:n}^{(1)}$	⋯	$z_{n:n}^{(1)}$
⋮	⋮		⋮		⋮		⋮
mth sample:	$z_{1:n}^{(m)}$	⋯	$z_{i:n}^{(m)}$	⋯	$z_{j:n}^{(m)}$	⋯	$z_{n:n}^{(m)}$

Fig. 1. Monte Carlo observations: m realizations of a set of standardized IG order statistics with sample size n.

a chi-square distribution with 1 degree of freedom (denoted χ_1^2). To generate an observation x of X, we can: (i) generate an observation y^2 of Y^2 from χ_1^2; (ii) solve the equation $y^2 = \lambda(x-\delta)^2/(\delta^2 x)$ for x, yielding two roots, $x_1 = \delta[2\lambda + \delta y^2 - \sqrt{4\lambda\delta y^2 + \delta^2 y^4}]/(2\lambda)$ and $x_2 = \delta^2/x_1$; (iii) randomly choose x_1 with probability $\delta/(\delta + x_1)$ and x_2 with probability $x_1/(\delta + x_1)$. The procedure to generate observations from a three-parameter IG(μ, σ, k) distribution of Eq. (4) is: (i) compute $\delta = 3\sigma/k$ and $\lambda = \delta^3/\sigma^2$; (ii) generate an observation x from the two-parameter IG based on the values for δ and λ computed in step (i); (iii) move x by $\mu - 3\sigma/k$ units. For the standardized IG, we merely set $\mu = 0$ and $\sigma = 1$. Given a sample of n standardized IG random variates, the order statistics from the sample are the n random variates ranked in increasing order.

For the purpose of variance reduction, we use the inverse-transformation method to generate standardized IG order statistics. Specifically, the inverse-transformation procedure is: (i) generate a set of n independent $U(0,1)$ random numbers $U_1, \ldots, U_n$; (ii) rank them in increasing order to obtain $U(0,1)$ order statistics $U_{1:n}, \ldots, U_{n:n}$; and (iii) compute $Z_{i:n} = F^{-1}(U_{i:n})$ for $i = 1, \ldots, n$, where F^{-1} is the inverse of the standardized IG cdf. A combination of the Newton and regula-falsi methods is used to solve the rootfinding problem in (iii). The combination approach starts with Newton's method until the stopping rule is satisfied or a bounding interval $[x_l, x_u]$ is found. After the bounding interval is found, the Newton's iterate is used only when it falls within the bounding interval. Otherwise, the regula-falsi iterate is computed. Given a bounding interval $[x_l, x_u]$, the regula-falsi method approximates the rootfinding function $F(x)$ by the linear function that goes through the points $(x_l, F(x_l))$ and $(x_u, F(x_u))$. The regula-falsi iterate is defined as the root of the linear function. To accelerate the search, we use the mode $z_M = -3/k + (\sqrt{36 + k^4} - k^2)/(2k)$ (Eq. (5)) as the initial guess for $Z_{1:n}$ and the computed value of $Z_{i-1:n}$ as the initial guess for all subsequent roots $Z_{i:n}, i = 2, \ldots, n$. Cheng (2000) shows that the inverse-transformation method is faster than the chi-square method proposed by Michael et al. (1976).

Inverse transformation has another advantage in that it is easy to apply control variates to reduce the variance of crude estimators (Eq. (10)). Two kinds of control variates are used here: (a) $U(0,1)$ control variates, which are mean, variance, and covariance estimators of $U(0,1)$ order statistics, and (b) exp(1) control variates, which are mean, variance, and covariance estimators of exp(1) order statistics, where exp(1) denotes the exponential distribution with mean 1. Since the true values of the $U(0,1)$ and exp(1) order-statistic means, variances, and covariances are known (e.g., Bickel and Doksum, 2001, p. 528), their estimators can be used as control variates. We discuss the two kinds of control variates in turn.

Because the ith order statistic $U_{i:n}$ of $U(0,1)$ has a beta$(i,\ n-i+1)$ distribution and the difference $(U_{j:n} - U_{i:n})$ has a beta$(j-i,\ n+i-j+1)$ distribution, $1 \leqslant i < j \leqslant n$,

$$\begin{aligned}
&\mathrm{E}(U_{i:n}) = i/(n+1),\\
&\mathrm{V}(U_{i:n}) = i(n-i+1)/[(n+1)^2(n+2)],\\
&\mathrm{Cov}(U_{i:n}, U_{j:n}) = i(n-j+1)/[(n+1)^2(n+2)].
\end{aligned}$$

Given m random samples $\{u_{1:n}^{(b)}, \ldots, u_{n:n}^{(b)} : b = 1, \ldots, m\}$ of $U(0,1)$ order statistics $U_{1:n}, \ldots, U_{n:n}$, the values of $\mathrm{E}(U_{i:n})$, $\mathrm{V}(U_{i:n})$, and $\mathrm{Cov}(U_{i:n}, U_{j:n})$, $1 \leqslant i < j \leqslant n$, can be estimated by

$$\widehat{\mathrm{E}}(U_{i:n}) = \bar{u}_{i:n} = \sum_{b=1}^{m} u_{i:n}^{(b)}/m,$$

$$\widehat{\mathrm{V}}(U_{i:n}) = \sum_{b=1}^{m} (u_{i:n}^{(b)} - \bar{u}_{i:n})^2/(m-1),$$

$$\widehat{\mathrm{Cov}}(U_{i:n}, U_{j:n}) = \sum_{b=1}^{m} (u_{i:n}^{(b)} - \bar{u}_{i:n})(u_{j:n}^{(b)} - \bar{u}_{j:n})/(m-1).$$

Using these estimators as control variates, the control-variate estimators of $\mathrm{E}(Z_{i:n})$, $\mathrm{V}(Z_{i:n})$, and $\mathrm{Cov}(Z_{i:n}, Z_{j:n})$, respectively, are

$$\begin{aligned}
\hat{\mu}^U_{i:n} &= \hat{\mu}_{i:n} - \alpha_i \left[\widehat{\mathrm{E}}(U_{i:n}) - \frac{i}{n+1} \right], \quad i = 1, \ldots, n, \\
\hat{\sigma}^U_{i,i:n} &= \hat{\sigma}_{i,i:n} - \beta_i \left[\widehat{\mathrm{V}}(U_{i:n}) - \frac{i(n-i+1)}{(n+1)^2(n+2)} \right], \quad i = 1, \ldots, n, \\
\hat{\sigma}^U_{i,j:n} &= \hat{\sigma}_{i,j:n} - \gamma_{i,j} \left[\widehat{\mathrm{Cov}}(U_{i:n}, U_{j:n}) - \frac{i(n-j+1)}{(n+1)^2(n+2)} \right], \quad 1 \leqslant i < j \leqslant n,
\end{aligned} \tag{11}$$

where the crude estimators $\hat{\mu}_{i:n}$, $\hat{\sigma}_{i,i:n}$, and $\hat{\sigma}_{i,j:n}$ are defined in Eq. (10). The coefficients α_i, β_i, and $\gamma_{i,j}$ are set as $\widehat{\mathrm{Cov}}[\hat{\mu}_{i:n}, \widehat{\mathrm{E}}(U_{i:n})]/\widehat{\mathrm{V}}[\widehat{\mathrm{E}}(U_{i:n})]$, $\widehat{\mathrm{Cov}}[\hat{\sigma}_{i,i:n}, \widehat{\mathrm{V}}(U_{i:n})]/\widehat{\mathrm{V}}[\widehat{\mathrm{V}}(U_{i:n})]$, and $\widehat{\mathrm{Cov}}[\hat{\sigma}_{i,j:n}, \widehat{\mathrm{Cov}}(U_{i:n}, U_{j:n})]/\widehat{\mathrm{V}}[\widehat{\mathrm{Cov}}(U_{i:n}, U_{j:n})]$, respectively (see Law and Kelton, 2000, p. 605). Using the same observations $\{u^{(b)}_{1:n}, \ldots, u^{(b)}_{n:n} : b = 1, \ldots, m\}$, these coefficients can be computed by the micro/macroreplication method—dividing the m replications into r macroreplications of $l(= m/r)$ microreplications. The details of the $U(0,1)$ control-variate method are listed in Appendix A. Notice that the $U(0,1)$ control variates $\widehat{\mathrm{E}}(U_{i:n})$, $\widehat{\mathrm{V}}(U_{i:n})$, and $\widehat{\mathrm{Cov}}(U_{i:n}, U_{j:n})$ are internal controls because the uniform order statistics have already been obtained from the crude experiment for generating IG order statistics by inverse transformation. Calculation of the coefficients α_i, β_i, and $\gamma_{i,j}$ has a negligible impact on overall computation time.

The exp(1) control-variate method is identical to the $U(0,1)$ method except that the $U(0,1)$ order statistics are replaced by exp(1) order statistics. Unlike the internal $U(0,1)$ control variates, the exp(1) control variates are external because the crude experiment does not generate exp(1) order statistics. Let $\{Y_{1:n}, \ldots, Y_{n:n}\}$ denote a set of exp(1) order statistics of size n. The means, variances, and covariances of the exp(1) order statistics are

$$\mathrm{E}(Y_{i:n}) = \sum_{c=0}^{i-1} (n-c)^{-1}, \quad \mathrm{V}(Y_{i:n}) = \sum_{c=0}^{i-1} (n-c)^{-2} \quad \text{for } i = 1, \ldots, n, \quad \text{and}$$

$$\mathrm{Cov}(Y_{i:n}, Y_{j:n}) = \mathrm{V}(Y_{i:n}) = \sum_{c=0}^{i-1} (n-c)^{-2} \quad \text{for } 1 \leqslant i < j \leqslant n.$$

Given m random samples $\{y^{(b)}_{1:n}, \ldots, y^{(b)}_{n:n} : b = 1, \ldots, m\}$ of exp(1) order statistics $Y_{1:n}, \ldots, Y_{n:n}$, the exp(1) control variates are

$$\widehat{\mathrm{E}}(Y_{i:n}) = \bar{y}_{i:n} = \sum_{b=1}^{m} y^{(b)}_{i:n}/m, \quad i = 1, \ldots, n,$$

$$\widehat{\mathrm{V}}(Y_{i:n}) = \sum_{b=1}^{m} (y^{(b)}_{i:n} - \bar{y}_{i:n})^2/(m-1), \quad i = 1, \ldots, n,$$

$$\widehat{\mathrm{Cov}}(Y_{i:n}, Y_{j:n}) = \sum_{b=1}^{m} (y^{(b)}_{i:n} - \bar{y}_{i:n})(y^{(b)}_{j:n} - \bar{y}_{j:n})/(m-1), \quad 1 \leqslant i < j \leqslant n.$$

Hence, the exp(1) control-variate estimators are

$$\hat{\mu}^E_{i:n} = \hat{\mu}_{i:n} - \tilde{\alpha}_i \left[\widehat{\mathrm{E}}(Y_{i:n}) - \sum_{c=0}^{i-1} (n-c)^{-1} \right], \quad i = 1, \ldots, n,$$
$$\hat{\sigma}^E_{i,i:n} = \hat{\sigma}_{i,i:n} - \tilde{\beta}_i \left[\widehat{\mathbf{V}}(Y_{i:n}) - \sum_{c=0}^{i-1} (n-c)^{-2} \right], \quad i = 1, \ldots, n, \tag{12}$$
$$\hat{\sigma}^E_{i,j:n} = \hat{\sigma}_{i,j:n} - \tilde{\gamma}_{i,j} \left[\widehat{\mathrm{Cov}}(Y_{i:n}, Y_{j:n}) - \sum_{c=0}^{i-1} (n-c)^{-2} \right], \quad 1 \leqslant i < j \leqslant n.$$

The coefficients $\tilde{\alpha}_i$, $\tilde{\beta}_i$, and $\tilde{\gamma}_{i,j}$—the exp(1) counterparts of the $U(0,1)$ values of α_i, β_i, and $\gamma_{i,j}$ in Eq. (11)—can be computed by the micro/macroreplication method. The procedure for computing exp(1) control-variate estimates is listed in Appendix B. The variance-reduction effect of the $U(0,1)$ and exp(1) control variates is discussed in Section 4.

4. Empirical results

Here we use simulation experiments to study the variance-reduction effects of the two kinds of control variates discussed in Section 3. Twenty simulation runs are made, each with $m = 50{,}000$ replications, which are divided into 20 macroreplications and 2500 microreplications (i.e., $r = 20$ and $l = 2500$). Common random numbers are used for all cases. Simulation results show that (i) both control variates provide fairly good variance reduction, (ii) generally the exp(1) control variate performs better than the $U(0,1)$ control variate because the IG has a long right tail, and (iii) the exp(1) control variate has the best variance reduction when the IG skewness k equals 2 (the skewness of exp(1) is also 2). Results (ii) and (iii) indicate that the closer the distribution shape of the control variate to that of the IG, the more the variance is reduced.

Fig. 2 shows the variances of the three mean estimates $\hat{\mu}_{i:n}$, $\hat{\mu}^U_{i:n}$, and $\hat{\mu}^E_{i:n}$ for order $i = 1, \ldots, n$, sample size $n = 10$, and IG skewness $k = 0.1$, 2, 5, and 50. In each subfigure, the y-axis represents a log scale for dispersing the curves. The first subfigure compares the crude estimates. When $k = 0.1$, the variance $\mathrm{V}(\hat{\mu}_{i:n})$ decreases for $i = 1$ to 4 and then increases with the order i. As k increases, $\mathrm{V}(\hat{\mu}_{1:n})$ declines but $\mathrm{V}(\hat{\mu}_{n:n})$ rises. When the IG is more right skewed—i.e., k is larger—the IG left tail gets shorter and the right tail gets lower. Hence, estimation of the mean becomes easier for the smallest order statistics but harder for the largest order statistics. Therefore, when $k \geqslant 5$, $\mathrm{V}(\hat{\mu}_{i:n})$ is an increasing function of i. For most values of k, $\mathrm{V}(\hat{\mu}_{n:n})$ is largest among all orders and hence is the main object for variance reduction.

The second subfigure shows that the $U(0,1)$ control variates work well for small and medium orders i but not so well for high orders, especially for $i = n$. This is because the bounded shape of $U(0,1)$ does not match the long-right-tailed shape of the IG. On the other hand, the exp(1) control variates work well for large values of i as shown in the third subfigure; this is because both the exp(1) and the IG have a long right tail. Since in many situations $\mathrm{V}(\hat{\mu}_{n:n})$ is largest among all the $\mathrm{V}(\hat{\mu}_{i:n})$'s, we conclude that the exp(1) control variate is more effective than the $U(0,1)$ control variate.

Fig. 3 plots the variances of $\hat{\sigma}_{i,i:n}$, $\hat{\sigma}^U_{i,i:n}$, and $\hat{\sigma}^E_{i,i:n}$, $i = 1, \ldots, n$, for sample size $n = 10$ and skewness $k = 0.1$, 2, 5, and 50. As in Fig. 2, the y-axis is rendered in log scale for better readability. The behavior of the variance estimators is similar to that of the mean estimators in Fig. 2. The $U(0,1)$ control variates have good variance reduction for small and medium orders while the exp(1) control variates work better for high orders. Hence, again exp(1) is superior to $U(0,1)$ for variance reduction.

Fig. 4 plots the variances of the covariance estimators $\hat{\sigma}_{i,j:n}$, $\hat{\sigma}^U_{i,j:n}$, and $\hat{\sigma}^E_{i,j:n}$ for $i, j = 1, \ldots, n$. The sample size n is still 10 but, for simplicity, only three values of the skewness—0.1, 2, and 50—are studied. There are in total nine subfigures. Each column corresponds to one value of k; the first row corresponds to the crude

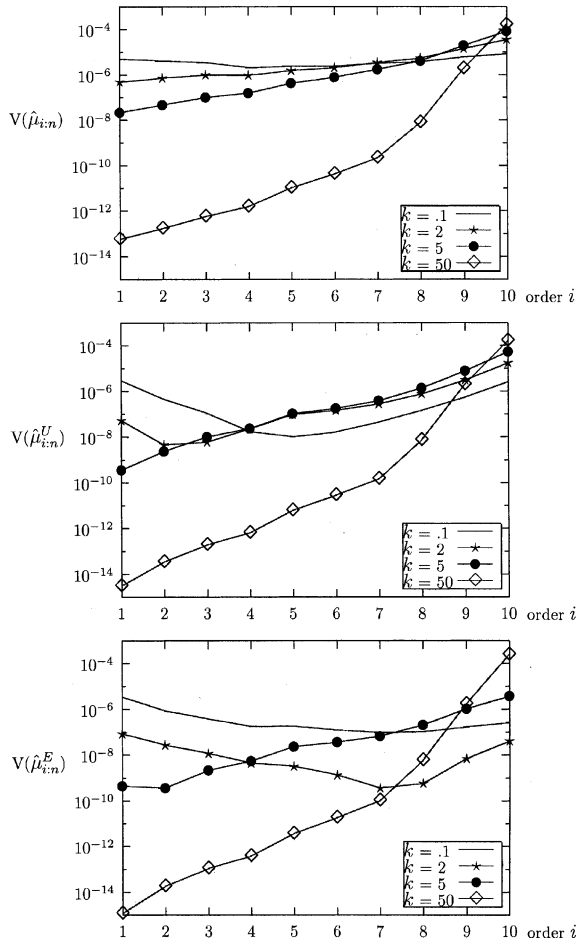

Fig. 2. Plots of $V(\hat{\mu}_{i:n})$, $V(\hat{\mu}^U_{i:n})$, and $V(\hat{\mu}^E_{i:n})$ for $n = 10$ and $k = 0.1$, 2, 5, 50.

estimators, the second to the $U(0,1)$ control-variate estimators, and the third to the exp(1) control-variate estimators. In each subfigure of Fig. 4, the heights above the "$i = j$" diagonal line represent the variances of the variance estimators as shown in Fig. 3. Fig. 4 shows that $V(\hat{\sigma}_{i,j:n})$, $i < j$, is largest when $(i,j) = (1,10)$; that is, the order statistics with lowest and highest orders have the largest correlation. To compare the variance-reduction effects of the $U(0,1)$ and exp(1) control variates, Table 1 lists the percentage of variance reduction, $\{\sum_{1\leqslant i,j\leqslant n} V(\hat{\sigma}_{i,j:n}) - \sum_{1\leqslant i,j\leqslant n} V(\hat{\sigma}^U_{i,j:n})\}/\sum_{1\leqslant i,j\leqslant n} V(\hat{\sigma}_{i,j:n})$ for $U(0,1)$ and $\{\sum_{1\leqslant i,j\leqslant n} V(\hat{\sigma}_{i,j:n}) - \sum_{1\leqslant i,j\leqslant n} V(\hat{\sigma}^E_{i,j:n})\}/\sum_{1\leqslant i,j\leqslant n} V(\hat{\sigma}_{i,j:n})$ for exp(1). When $k = 0.1$, the $U(0,1)$ and exp(1) control variates reduce the variances of all covariance estimators by about the same amount. When $k = 2$, the effect of the $U(0,1)$ control variates is hard to discern. The exp(1) control variates, however, perform well, reducing 99.4% of the variances. Unfortunately, when k increases to 50, both kinds of control variates don't work well, especially the $U(0,1)$ control variates with zero variance reduction.

Sometimes we want to simultaneously compute means, variances, and covariances of all orders—e.g., when constructing BLUEs. Table 2 compares the cumulative effect of variance reduction across all orders for two kinds of control variates. For evaluation purposes, we adopt a new measure the v-ratio, which is the ratio of the sum of the crude estimate variances over all orders to the sum of the control-variate estimate

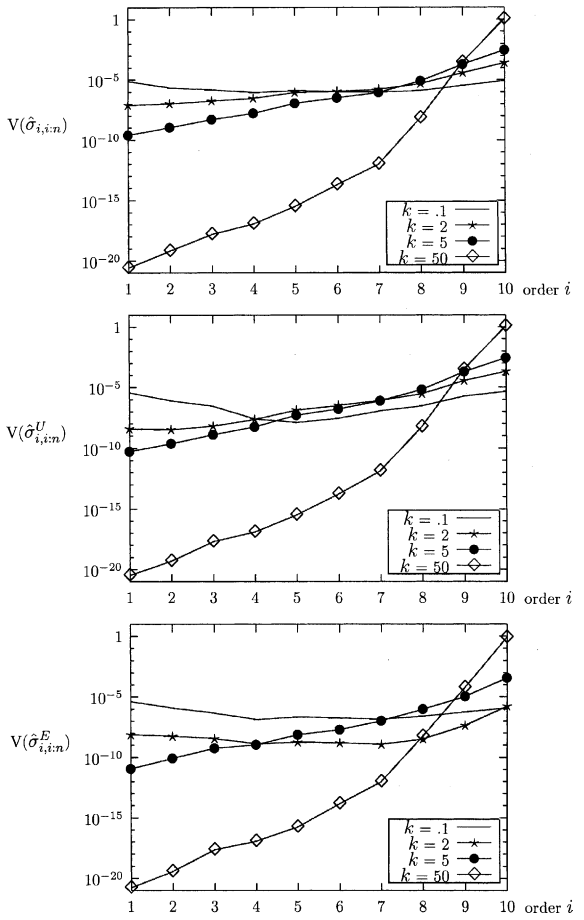

Fig. 3. Plots of $V(\hat{\sigma}_{i,i:n})$, $V(\hat{\sigma}^U_{i,i:n})$, and $V(\hat{\sigma}^E_{i,i:n})$ for $n = 10$, $k = 0.1$, 2, 5, 50.

variances over all orders. There are 42 design points: 21 for each kind of control variate—$U(0,1)$ and exp(1), denoted U and E, respectively—corresponding to $n \in \{10, 20, 30\}$ and $k \in \{0.1, 0.5, 1, 1.6, 2, 5, 50\}$. Table 2 compares the mean v-ratios $\sum_{i=1}^n V(\hat{\mu}_{i:n}) / \sum_{i=1}^n V(\hat{\mu}^U_{i:n})$ and $\sum_{i=1}^n V(\hat{\mu}_{i:n}) / \sum_{i=1}^n V(\hat{\mu}^E_{i:n})$, the variance v-ratios $\sum_{i=1}^n V(\hat{\sigma}_{i,i:n}) / \sum_{i=1}^n V(\hat{\sigma}^U_{i,i:n})$ and $\sum_{i=1}^n V(\hat{\sigma}_{i,i:n}) / \sum_{i=1}^n V(\hat{\sigma}^E_{i,i:n})$, and the covariance v-ratios $\sum_{1 \leqslant i<j \leqslant n} V(\hat{\sigma}_{i,j:n}) / \sum_{1 \leqslant i<j \leqslant n} V(\hat{\sigma}^U_{i,j:n})$ and $\sum_{1 \leqslant i<j \leqslant n} V(\hat{\sigma}_{i,j:n}) / \sum_{1 \leqslant i<j \leqslant n} V(\hat{\sigma}^E_{i,j:n})$. For each design point, the CPU time (hour:minute:second) and the v-ratios for the mean, variance (denoted as var), and covariance (denoted as cov) estimators are shown. A higher value of the ratio implies more variance reduction. The v-ratio is computed from twenty simulation runs, each of $m = 50{,}000$ replications, as in Figs. 2–4. Common random numbers are used for all design points.

Table 2 shows that the v-ratio for $U(0,1)$ is highest when $k = 0.1$, and decreases as k increases. Therefore, the $U(0,1)$ control variates work better for small values of k. As k approaches zero, the $IG(0,1,k)$ distribution converges to the standard normal distribution and, hence, the symmetric uniform distribution works well for small values of k. The v-ratio for exp(1) is highest when k is near 2 and decreases as k moves away from 2. This is because the skewness of exp(1) is also 2. The closer the exp(1) shape to the IG shape, the more the variance is reduced. For most design points, the exp(1) v-ratio is higher than the $U(0,1)$ v-ratio. Although the external exp(1) control variates require more computation, the extra

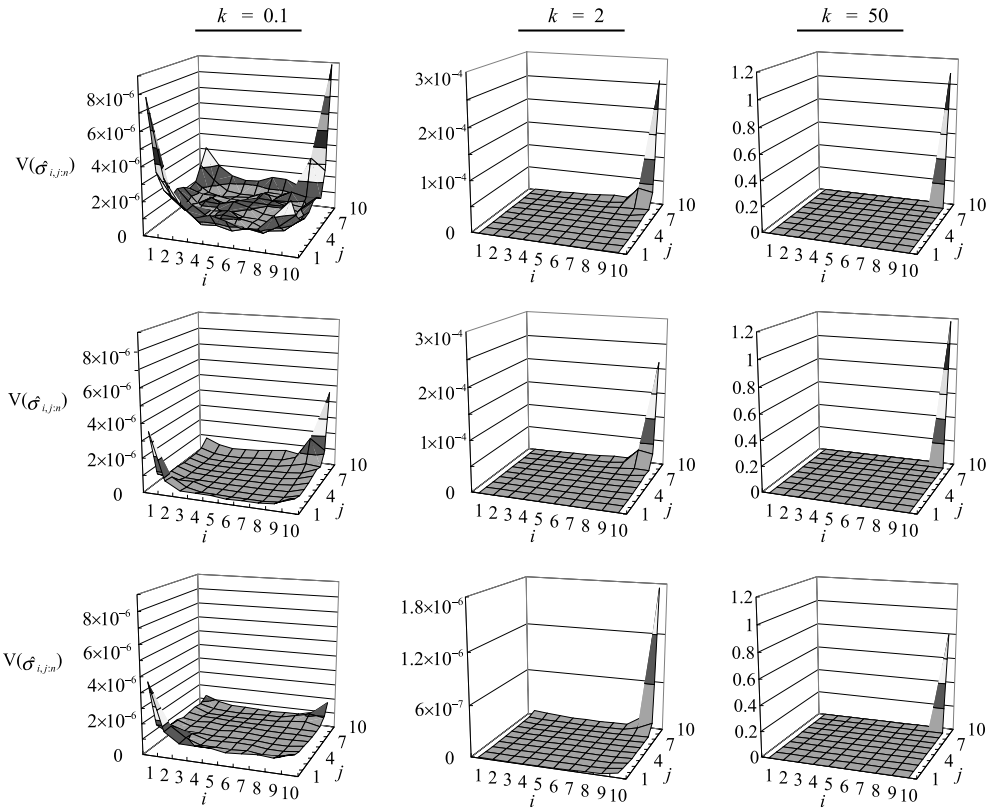

Fig. 4. Plots of $V(\hat{\sigma}_{i,j:n})$, $V(\hat{\sigma}^{U}_{i,j:n})$, and $V(\hat{\sigma}^{E}_{i,j:n})$ for $1 \leqslant i, j \leqslant n$, $n = 10$, and $k = 0.1$, 2, and 50.

Table 1
Percentages of variance-reduction on the total variances of all covariance estimators

k	0.1	2	50
$U(0,1)$	76.4%	22.6%	0%
exp(1)	76.8%	99.4%	26.8%

computation time is minor, at less than 4%. Therefore, exp(1) is preferable to $U(0,1)$ based on the v-ratio comparison.

5. Conclusions

We have proposed a Monte Carlo simulation algorithm for computing means, variances, and covariances for a set of order statistics from the inverse Gaussian distribution. The control-variate variance-reduction technique is used to increase the precision of the estimators. There are two kinds of control variates: the internal-control $U(0,1)$ and the external-control exp(1). Simulation results show that the exponential control variates are more efficient than the uniform control variates, especially when the inverse-Gaussian skewness is near 2, the exponential skewness value. When the inverse-Gaussian population

Table 2
Ratios to evaluate overall variance-reduction effects

k	CV	$n = 10$				$n = 20$				$n = 30$			
		Mean	Var	Cov	CPU	Mean	Var	Cov	CPU	Mean	Var	Cov	CPU
0.1	*U*	5.95	2.56	5.15	8:56	8.91	1.58	3.96	18:11	9.36	1.65	4.19	27:31
0.5	*U*	5.84	1.97	4.01	12:30	7.18	1.24	2.82	25:22	7.07	1.31	3.91	38:29
1	*U*	4.65	1.49	2.59	11:34	5.19	1.09	1.89	23:29	4.63	1.11	3.06	35:39
1.6	*U*	3.49	1.27	1.74	11:05	3.83	1.04	1.43	22:23	3.14	1.01	2.40	33:59
2	*U*	2.99	1.21	1.46	10:59	3.31	1.02	1.29	22:03	2.62	0.98	2.13	33:28
5	*U*	1.73	1.07	0.98	11:39	1.99	1.01	1.08	23:38	1.53	0.90	1.36	35:39
50	*U*	0.95	0.92	0.83	25:20	1.03	1.04	1.34	51:18	1.10	1.00	0.84	1:16:53
0.1	*E*	6.96	3.43	4.63	9:08	17.7	4.19	5.91	18:36	17.8	3.63	5.43	28:10
0.5	*E*	13.6	9.34	7.91	12:49	36.9	17.5	11.5	25:59	34.3	16.3	12.3	39:18
1	*E*	37.7	49.4	19.6	11:52	112	132	35.5	24:06	95.8	138	42.1	36:32
1.6	*E*	167	761	98	11:22	550	2274	221	22:58	419	1367	248	34:43
2	*E*	355	167	208	11:15	878	426	442	22:48	654	260	384	34:33
5	*E*	21.4	8.42	7.01	12:00	45.9	28.5	30.4	24:19	36.82	18.8	14.1	36:54
50	*E*	0.65	1.37	1.78	25:42	2.12	2.16	2.94	51:54	1.43	1.78	1.28	1:17:56

is highly right skewed, neither kind of control variate works well. Finding control variates whose population-distribution shapes match the inverse-Gaussian shape of interest is a topic for future research. Weibull shapes are an option. The theoretical statistical properties of Weibull order statistics, however, are unknown except for special cases, e.g., exponential. To use Weibull control variates, initialization would require computing the means, variances, and covariances of Weibull order statistics, either numerically or stochastically.

Acknowledgements

This research is supported by National Science Council in Taiwan under Grant NSC89-2213-E-003-063. We thank Carol Troy for proofreading this paper and referees for insightful comments.

Appendix A. The $U(0,1)$ control-variate method

Given:

n: sample size.
k: inverse-Gaussian skewness.
r: number of macroreplications.
l: number of microreplications.

Procedure:

0. Let $m = rl$.
1. For $p = 1, \ldots, r$, do steps 1.1 and 1.2:
   1.1 For $q = 1, \ldots, l$:
      (i) Generate a random sample $\{u_1, \ldots, u_n\}$ from the $U(0,1)$ distribution.

(ii) Rank $\{u_1,\ldots,u_n\}$ in increasing order to obtain uniform order statistics $u_{1:n}^{(b)} \leqslant \cdots \leqslant u_{n:n}^{(b)}$, where $b=(p-1)l+q$.

(iii) Compute $z_{i:n}^{(b)} = F^{-1}(u_{i:n}^{(b)})$, $i=1,\ldots,n$, where F^{-1} is the inverse of the standardized inverse-Gaussian cdf.

1.2 Compute the following microsample statistics for $1 \leqslant i < j \leqslant n$:

$$\bar{u}_{i:n}^{(p)} = \sum_{b=(p-1)l+1}^{pl} u_{i:n}^{(b)}/l.$$

$$\bar{z}_{i:n}^{(p)} = \sum_{b=(p-1)l+1}^{pl} z_{i:n}^{(b)}/l.$$

$$\widehat{\mathbf{V}}^{(p)}(U_{i:n}) = \frac{[\sum_{b=(p-1)l+1}^{pl} (u_{i:n}^{(b)})^2] - l[\bar{u}_{i:n}^{(p)}]^2}{l-1}.$$

$$\widehat{\mathbf{V}}^{(p)}(Z_{i:n}) = \frac{[\sum_{b=(p-1)l+1}^{pl} (z_{i:n}^{(b)})^2] - l[\bar{z}_{i:n}^{(p)}]^2}{l-1}.$$

$$\widehat{\mathbf{Cov}}^{(p)}(U_{i:n}, U_{j:n}) = \frac{[\sum_{b=(p-1)l+1}^{pl} u_{i:n}^{(b)} u_{j:n}^{(b)}] - l\bar{u}_{i:n}^{(p)} \bar{u}_{j:n}^{(p)}}{l-1}.$$

$$\widehat{\mathbf{Cov}}^{(p)}(Z_{i:n}, Z_{j:n}) = \frac{[\sum_{b=(p-1)l+1}^{pl} z_{i:n}^{(b)} z_{j:n}^{(b)}] - l\bar{z}_{i:n}^{(p)} \bar{z}_{j:n}^{(p)}}{l-1}.$$

2. Compute the crude estimates:

$$\hat{\mu}_{i:n} = \bar{\bar{z}}_{i:n} = \sum_{p=1}^{r} \bar{z}_{i:n}^{(p)}/r = \sum_{b=1}^{m} z_{i:n}^{(b)}/m, \quad i=1,\ldots,n.$$

$$\hat{\sigma}_{i,i:n} = s_{i,i:n} = \frac{[\sum_{b=1}^{m} (z_{i:n}^{(b)})^2] - m(\bar{\bar{z}}_{i:n})^2}{m-1}, \quad i=1,\ldots,n.$$

$$\hat{\sigma}_{i,j:n} = s_{i,j:n} = \frac{[\sum_{b=1}^{m} z_{i:n}^{(b)} z_{j:n}^{(b)}] - m\bar{\bar{z}}_{i:n}\bar{\bar{z}}_{j:n}}{m-1}, \quad 1 \leqslant i < j \leqslant n.$$

3. Compute the following macroaverages:

$$\bar{\bar{u}}_{i:n} = \sum_{p=1}^{r} \bar{u}_{i:n}^{(p)}/r = \sum_{b=1}^{m} u_{i:n}^{(b)}/m, \quad i=1,\ldots,n.$$

$$\overline{\widehat{\mathbf{V}}}(U_{i:n}) = \sum_{p=1}^{r} \widehat{\mathbf{V}}^{(p)}(U_{i:n})/r, \quad i=1,\ldots,n.$$

$$\overline{\widehat{\mathbf{V}}}(Z_{i:n}) = \sum_{p=1}^{r} \widehat{\mathbf{V}}^{(p)}(Z_{i:n})/r, \quad i=1,\ldots,n.$$

$$\overline{\widehat{\mathbf{Cov}}}(U_{i:n}, U_{j:n}) = \sum_{p=1}^{r} \widehat{\mathbf{Cov}}^{(p)}(U_{i:n}, U_{j:n})/r, \quad i=1,\ldots,n.$$

$$\overline{\widehat{\mathbf{Cov}}}(Z_{i:n}, Z_{j:n}) = \sum_{p=1}^{r} \widehat{\mathbf{Cov}}^{(p)}(Z_{i:n}, Z_{j:n})/r, \quad i=1,\ldots,n.$$

4. Compute the following variance and covariance estimates:

$$\widehat{\mathrm{Cov}}[\hat{\mu}_{i:n}, \widehat{\mathbf{E}}(U_{i:n})] = \frac{[\sum_{p=1}^{r} \bar{z}_{i:n}^{(p)} \bar{u}_{i:n}^{(p)}] - r\bar{\bar{z}}_{i:n}\bar{\bar{u}}_{i:n}}{r(r-1)}, \quad i = 1, \ldots, n.$$

$$\widehat{\mathbf{V}}[\widehat{\mathbf{E}}(U_{i:n})] = \frac{\sum_{p=1}^{r} [\bar{u}_{i:n}^{(p)}]^2 - r[\bar{\bar{u}}_{i:n}]^2}{r(r-1)}, \quad i = 1, \ldots, n.$$

$$\widehat{\mathrm{Cov}}[\hat{\sigma}_{i,i:n}, \widehat{\mathbf{V}}(U_{i:n})] = \frac{[\sum_{p=1}^{r} \widehat{\mathbf{V}}^{(p)}(Z_{i:n})\, \widehat{\mathbf{V}}^{(p)}(U_{i:n})] - r\overline{\widehat{\mathbf{V}}}(Z_{i:n})\, \overline{\widehat{\mathbf{V}}}(U_{i:n})}{r(r-1)}, \quad i = 1, \ldots, n.$$

$$\widehat{\mathbf{V}}[\widehat{\mathbf{V}}(U_{i:n})] = \frac{\sum_{p=1}^{r} [\widehat{\mathbf{V}}^{(p)}(U_{i:n})]^2 - r[\overline{\widehat{\mathbf{V}}}(U_{i:n})]^2}{r(r-1)}, \quad i = 1, \ldots, n.$$

$$\widehat{\mathrm{Cov}}[\hat{\sigma}_{i,j:n}, \widehat{\mathrm{Cov}}(U_{i:n}, U_{j:n})] = [r(r-1)]^{-1} \left[\sum_{p=1}^{r} \widehat{\mathrm{Cov}}^{(p)}(Z_{i:n}, Z_{j:n})\, \widehat{\mathrm{Cov}}^{(p)}(U_{i:n}, U_{j:n}) - r\overline{\widehat{\mathrm{Cov}}}(Z_{i:n}, Z_{j:n})\overline{\widehat{\mathrm{Cov}}}(U_{i:n}, U_{j:n}) \right], \quad 1 \leqslant i < j \leqslant n.$$

$$\widehat{\mathbf{V}}[\widehat{\mathrm{Cov}}(U_{i:n}, U_{j:n})] = \frac{\sum_{p=1}^{r} [\widehat{\mathrm{Cov}}^{(p)}(U_{i:n}, U_{j:n})]^2 - r[\overline{\widehat{\mathrm{Cov}}}(U_{i:n}, U_{j:n})]^2}{r(r-1)}, \quad 1 \leqslant i < j \leqslant n.$$

5. Compute the control-variate coefficients:

$$\alpha_i = \widehat{\mathrm{Cov}}[\hat{\mu}_{i:n}, \widehat{\mathbf{E}}(U_{i:n})] / \widehat{\mathbf{V}}[\widehat{\mathbf{E}}(U_{i:n})], \quad i = 1, \ldots, n.$$

$$\beta_i = \widehat{\mathrm{Cov}}[\hat{\sigma}_{i,i:n}, \widehat{\mathbf{V}}(U_{i:n})] / \widehat{\mathbf{V}}[\widehat{\mathbf{V}}(U_{i:n})], \quad i = 1, \ldots, n.$$

$$\gamma_{i,j} = \widehat{\mathrm{Cov}}[\hat{\sigma}_{i,j:n}, \widehat{\mathrm{Cov}}(U_{i:n}, U_{j:n})] / \widehat{\mathbf{V}}[\widehat{\mathrm{Cov}}(U_{i:n}, U_{j:n})], \quad 1 \leqslant i < j \leqslant n.$$

6. Compute the uniform control-variate estimates using Eq. (11):

$$\hat{\mu}_{i:n}^{U} = \hat{\mu}_{i:n} - \alpha_i \left[\bar{\bar{u}}_{i:n} - \frac{i}{n+1} \right], \quad i = 1, \ldots, n.$$

$$\hat{\sigma}_{i,i:n}^{U} = \hat{\sigma}_{i,i:n} - \beta_i \left[\overline{\widehat{\mathbf{V}}}(U_{i:n}) - \frac{i(n-i+1)}{(n+1)^2(n+2)} \right], \quad i = 1, \ldots, n.$$

$$\hat{\sigma}_{i,j:n}^{U} = \hat{\sigma}_{i,j:n} - \gamma_{i,j} \left[\overline{\widehat{\mathrm{Cov}}}(U_{i:n}, U_{j:n}) - \frac{i(n-j+1)}{(n+1)^2(n+2)} \right], \quad 1 \leqslant i < j \leqslant n.$$

Appendix B. The exp(1) control-variate method

Given: n, k, r and l as in Appendix A.

Procedure:

0. Let $m = rl$.
1. For $p = 1, \ldots, r$, do steps 1.1 and 1.2:
   1.1 For $q = 1, \ldots, l$:
   (i)–(iii): Same as in Appendix A.
   (iv) Compute $y_{i:n}^{(b)} = -\ln(1 - u_{i:n}^{(b)})$, $i = 1, \ldots, n$.
   1.2 Compute microsample statistics $\bar{z}_{i:n}^{(p)}$, $\widehat{\mathbf{V}}^{(p)}(Z_{i:n})$, and $\widehat{\mathrm{Cov}}^{(p)}(Z_{i:n}, Z_{j:n})$ as in Appendix A, and

$$\bar{y}_{i:n}^{(p)} = \sum_{b=(p-1)l+1}^{pl} y_{i:n}^{(b)}/l, \quad i = 1, \ldots, n,$$

$$\widehat{\mathbf{V}}^{(p)}(Y_{i:n}) = \frac{[\sum_{b=(p-1)l+1}^{pl} (y_{i:n}^{(b)})^2] - l[\bar{y}_{i:n}^{(p)}]^2}{l-1}, \quad i = 1, \ldots, n,$$

$$\widehat{\mathrm{Cov}}^{(p)}(Y_{i:n}, Y_{j:n}) = \frac{[\sum_{b=(p-1)l+1}^{pl} y_{i:n}^{(b)} y_{j:n}^{(b)}] - l\bar{y}_{i:n}^{(p)} \bar{y}_{j:n}^{(p)}}{l-1}, \quad 1 \leqslant i < j \leqslant n.$$

2. Compute the crude estimates $\hat{\mu}_{i:n}$, $\hat{\sigma}_{i,i:n}$, and $\hat{\sigma}_{i,j:n}$ as in Step 2 of Appendix A.
3. Compute macroaverages $\overline{\widehat{\mathbf{V}}}(Z_{i:n})$ and $\overline{\widehat{\mathrm{Cov}}}(Z_{i:n}, Z_{j:n})$ as in Step 3 of Appendix A, and

$$\bar{\bar{y}}_{i:n} = \sum_{p=1}^{r} \bar{y}_{i:n}^{(p)}/r, \quad i = 1, \ldots, n,$$

$$\overline{\widehat{\mathbf{V}}}(Y_{i:n}) = \sum_{p=1}^{r} \widehat{\mathbf{V}}^{(p)}(Y_{i:n})/r, \quad i = 1, \ldots, n,$$

$$\overline{\widehat{\mathrm{Cov}}}(Y_{i:n}, Y_{j:n}) = \sum_{p=1}^{r} \widehat{\mathrm{Cov}}^{(p)}(Y_{i:n}, Y_{j:n})/r, \quad i = 1, \ldots, n.$$

4. Compute $\widehat{\mathrm{Cov}}[\hat{\mu}_{i:n}, \widehat{\mathbf{E}}(Y_{i:n})]$, $\widehat{\mathbf{V}}[\widehat{\mathbf{E}}(Y_{i:n})]$, $\widehat{\mathrm{Cov}}[\hat{\sigma}_{i,i:n}, \widehat{\mathbf{V}}(Y_{i:n})]$, $\widehat{\mathbf{V}}[\widehat{\mathbf{V}}(Y_{i:n})]$, $\widehat{\mathrm{Cov}}[\hat{\sigma}_{i,j:n}, \widehat{\mathrm{Cov}}(Y_{i:n}, Y_{j:n})]$, and $\widehat{\mathbf{V}}[\widehat{\mathrm{Cov}}(Y_{i:n}, Y_{j:n})]$ as in Step 4 of Appendix A, but replace $\bar{u}_{i:n}^{(p)}$, $\bar{\bar{u}}_{i:n}$, $\widehat{\mathbf{V}}^{(p)}(U_{i:n})$, $\overline{\widehat{\mathbf{V}}}(U_{i:n})$, $\widehat{\mathrm{Cov}}^{(p)}(U_{i:n}, U_{j:n})$, and $\overline{\widehat{\mathrm{Cov}}}(U_{i:n}, U_{j:n})$ by $\bar{y}_{i:n}^{(p)}$, $\bar{\bar{y}}_{i:n}$, $\widehat{\mathbf{V}}^{(p)}(Y_{i:n})$, $\overline{\widehat{\mathbf{V}}}(Y_{i:n})$, $\widehat{\mathrm{Cov}}^{(p)}(Y_{i:n}, Y_{j:n})$, and $\overline{\widehat{\mathrm{Cov}}}(Y_{i:n}, Y_{j:n})$, respectively.
5. Compute the control-variate coefficients:

$$\tilde{\alpha}_i = \widehat{\mathrm{Cov}}[\hat{\mu}_{i:n}, \widehat{\mathbf{E}}(Y_{i:n})]/\widehat{\mathbf{V}}[\widehat{\mathbf{E}}(Y_{i:n})], \quad i = 1, \ldots, n.$$

$$\tilde{\beta}_i = \widehat{\mathrm{Cov}}[\hat{\sigma}_{i,i:n}, \widehat{\mathbf{V}}(Y_{i:n})]/\widehat{\mathbf{V}}[\widehat{\mathbf{V}}(Y_{i:n})], \quad i = 1, \ldots, n.$$

$$\tilde{\gamma}_{i,j} = \widehat{\mathrm{Cov}}[\hat{\sigma}_{i,j:n}, \widehat{\mathrm{Cov}}(Y_{i:n}, Y_{j:n})]/\widehat{\mathbf{V}}[\widehat{\mathrm{Cov}}(Y_{i:n}, Y_{j:n})], \quad 1 \leqslant i < j \leqslant n.$$

6. Compute the exp(1) control-variate estimates using Eq. (12):

$$\hat{\mu}_{i:n}^{E} = \hat{\mu}_{i:n} - \tilde{\alpha}_i \left[\bar{\bar{y}}_{i:n} - \sum_{c=0}^{i-1} (n-c)^{-1} \right], \quad i = 1, \ldots, n.$$

$$\hat{\sigma}_{i,i:n}^{E} = \hat{\sigma}_{i,i:n} - \tilde{\beta}_i \left[\overline{\widehat{\mathbf{V}}}(Y_{i:n}) - \sum_{c=0}^{i-1} (n-c)^{-2} \right], \quad i = 1, \ldots, n.$$

$$\hat{\sigma}_{i,j:n}^{E} = \hat{\sigma}_{i,j:n} - \tilde{\gamma}_{i,j} \left[\overline{\widehat{\mathrm{Cov}}}(Y_{i:n}, Y_{j:n}) - \sum_{c=0}^{i-1} (n-c)^{-2} \right], \quad 1 \leqslant i < j \leqslant n.$$

References

Arnold, B.C., Balakrishnan, N., Nagaraja, H.N., 1992. A First Course in Order Statistics. John Wiley and Sons, New York.

Balakrishnan, N., Chen, W.W., 1997. CRC Handbook of Tables for Order Statistics from Inverse Gaussian Distributions with Applications. CRC Press, New York.

Banerjee, A.K., Bhattacharyya, G.K., 1976. A purchase incidence model with inverse Gaussian interpurchase times. Journal of the American Statistical Association 71, 823–829.

Bickel, P.J., Doksum, K.A., 2001. Mathematical Statistics: Basic Ideas and Selected Topics, Vol. 1, second ed. Holden-Day, Oakland, CA.

Cheng, L., 2000. Computing Means and Covariances of Inverse-Gaussian Order Statistics for Applications in Constructing BLUEs, Master Thesis, Department of Industrial Engineering, Dayeh University, Chang Hwa, Taiwan. Chinese.

Chhikara, R.S., Folks, J.L., 1977. The inverse Gaussian distribution as a lifetime model. Technometrics 19, 461–468.

Chhikara, R.S., Folks, J.L., 1989. The Inverse Gaussian Distribution: Theory, Methodology, and Applications. Marcel Dekker, New York.

Cohen, A.C., Whitten, B.J., 1988. Parameter Estimation in Reliability and Life Span Models. Marcel Dekker, New York.

Conte, S.D., de Boor, C., 1980. Elementary Numerical Analysis: An Algorithmic Approach. McGraw-Hill, New York.

Folks, J.L., Chhikara, R.S., 1978. The inverse Gaussian distribution and its statistical application—A review. Journal of the Royal Statistical Society B 40, 263–289.

Gacula Jr., M.C., Kubala, J.J., 1975. Statistical models for shelf life failures. Journal of Food Science 40, 404–409.

Gunes, H., Dietz, D.C., Auclair, P.F., Moore, A.H., 1997. Modified goodness-of-fit tests for the inverse Gaussian distribution. Computational Statistics and Data Analysis 24, 63–77.

Johnson, N.L., Kotz, S., Balakrishnan, N., 1994. Continuous Univariate Distributions 1, second ed. John Wiley and Sons, New York.

Lancaster, A., 1972. A stochastic model for the duration of a strike. Journal of the Royal Statistical Society A 135, 257–271.

Law, A.M., Kelton, W.D., 2000. Simulation Modeling and Analysis. McGraw-Hill, New York.

Lieblein, J., Zelen, M., 1956. Statistical investigation of the fatigue life of deep-groove ball bearings. Journal of Research of the National Bureau of Standards 57, 272–316.

Michael, J.R., Schucany, W.R., Haas, R.W., 1976. Generating random variables using transformation with multiple roots. The American Statistician 30, 88–90.

Parrish, R.S., 1992a. Computing expected values of normal order statistics. Communications in Statistic—Simulation and Computation 21, 57–70.

Parrish, R.S., 1992b. Computing variances and covariances of normal order statistics. Communications in Statistic—Simulation and Computation 21, 71–101.

Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 1997. Numerical Recipes in C—The Art of Scientific Computing, second ed. Cambridge University Press, New York.

Schrödinger, E., 1915. Zur Theorie der Fall- und Steigversuche an Teilchen mit Brownscher Bewegung. Physikalische Zeitschrift 16, 289–295.

Seshadri, V., 1993. The Inverse Gaussian Distribution: A Case Study in Exponential Families. Oxford University Press, Oxford.

Sichel, H.S., 1975. On a distribution law for word frequencies. Journal of the American Statistical Association 70, 542–547.

Smoluchowski, M.V., 1915. Notiz über die Berechning der Brownschen Molkularbewegung bei des Ehrenhaft-millikanchen Versuchsanordnung. Physikalische Zeitschrift 16, 318–321.

Tweedie, M.C.K., 1945. Inverse statistical variates. Nature 155, 453.

Tweedie, M.C.K., 1956. Some statistical properties of inverse Gaussian distributions. Virginia Journal of Science 7, 160–165.

Tweedie, M.C.K., 1957a. Statistical properties of inverse Gaussian distributions, I. Annals of Mathematical Statistics 28, 362–377.

Tweedie, M.C.K., 1957b. Statistical properties of inverse Gaussian distributions, II. Annals of Mathematical Statistics 28, 696–705.

Von Alven, W.H., 1964. Reliability Engineering. Prentice-Hall, Upper Saddle River (by ARINC Englewood Cliffs, editor).

Wald, A., 1947. Sequential Analysis. John Wiley and Sons, New York.

Wasan, M.T., 1968. On an inverse Gaussian process. Skand. Actuar. 60, 69–96.

Whitmore, G.A., 1979. An inverse Gaussian model for labour turnover. Journal of the Royal Statistical Society A 142, 468–478.

Wise, M.E., 1971. Skew probability curves with negative powers of time and related to random walks in series. Statistica Neerlandica 25, 159–180.

Wise, M.E., 1975. Skew distributions in biomedicine including some with negative powers of time. In: Patil, G.P. et al. (Eds.), Statistical Distributions in Scientific Work 2: Model Building and Model Selection. D. Reidel, Dordrecht, Holland, pp. 241–262.

Wise, M.E., Osborn, S.B., Anderson, J., Tomlinson, R.W.S., 1968. A stochastic model for turnover of radiocalcium based on the observed laws. Mathematical Biosciences 2, 199–224.