Improving the Efficiency of the Monte-Carlo Methods Using Ranked Simulated Approach

Hani Michel Samawi

Abstract This chapter explores the concept of using ranked simulated sampling approach (RSIS) to improve the well-known Monte-Carlo methods, introduced by Samawi (1999), and extended to steady-state ranked simulated sampling (SRSIS) by Al-Saleh and Samawi (2000). Both simulation sampling approaches are then extended to multivariate ranked simulated sampling (MVRSIS) and multivariate steady-state ranked simulated sampling approach (MVSRSIS) by Samawi and Al-Saleh (2007) and Samawi and Vogel (2013). These approaches have been demonstrated as providing unbiased estimators and improving the performance of some of the Monte-Carlo methods of single and multiple integrals approximation. Additionally, the MVSRSIS approach has been shown to improve the performance and efficiency of Gibbs sampling (Samawi et al. 2012). Samawi and colleagues showed that their approach resulted in a large savings in cost and time needed to attain a specified level of accuracy.

1 Introduction

The term Monte-Carlo refers to techniques that use random processes to approximate a non-stochastic k-dimensional integral of the form

$$\theta = \int_{R^k} g(\underline{u})d\underline{u}, \tag{1.1}$$

(Hammersley and Handscomb 1964).

The literature presents many approximation techniques, including Monte-Carlo methods. However, as the dimension of the integrals rises, the difficulty of the integration problem increases even for relatively low dimensions (see Evans and Swartz 1995). Given such complications, many researchers are confused about which method

H.M. Samawi (✉)
Department of Biostatistics, Jiann-Ping Hsu College Public Health,
Georgia Southern University, 30460 Statesboro, Georgia
e-mail: hsamawi@georgiasouthern.edu

D.-G. Chen and J.D. Chen (eds.), *Monte-Carlo Simulation-Based Statistical Modeling*, ICSA Book Series in Statistics, DOI 10.1007/978-981-10-3307-0_2

to use; however, the advantages and disadvantages of each method are not the primary concern of this chapter. The focus of this chapter is the use of Monte-Carlo methods in multiple integration approximation.

The motivation for this research is based on the concepts of ranked set sampling (RSS), introduced by McIntyre (1952). The motivation is based on the fact that the ith quantified unit of RSS is simply an observation from $f_{(i)}$, where $f_{(i)}$ is the density function of the ith order statistic of a random sample of size n. When the underlying density is the uniform distribution on (0, 1), $f_{(i)}$ follows a beta distribution with parameters $(i, n - i + 1)$.

Samawi (1999) was the first to explore the idea of RSS (Beta sampler) for integral approximation. He demonstrated that the procedure can improve the simulation efficiency based on the ratio of the variances. Samawi's ranked simulated sampling procedure RSIS generates an independent random sample $U_{(1)}, U_{(2)}, \ldots, U_{(n)}$, which is denoted by RSIS, where $U_{(i)} \sim \beta(i, n - i + 1), \{i = 1, 2, ..., n\}$ and $\beta(., .)$ denotes the beta distribution. The RSIS procedure constitutes an RSS based on random samples from the uniform distribution $U(0, 1)$. The idea is to use this RSIS to compute (1.1) with $k = 1$, instead of using an SRS of size n from $U(0, 1)$, when the range of the integral in (1.1) is (0, 1). In case of arbitrary range (a, b) of the integral in (1.1), Samawi (1999) used the sample: $X_{(1)}, X_{(2)}, ..., X_{(n)}$ and the importance sampling technique to evaluate (1.1), where $X_{(i)} = F_X^{-1}(U_{(i)})$ and $F_X(.)$ is the distribution function of a continuous random variable. He showed theoretically and through simulation studies that using the RSIS sampler for evaluating (1.1) substantially improved the efficiency when compared with the traditional uniform sampler (USS).

Al-Saleh and Zheng (2002) introduced the idea of bivariate ranked set sampling (BVRSS) and showed through theory and simulation that BVRSS outperforms the bivariate simple random sample for estimating the population means. The BVRSS is as follows:

Suppose (X, Y) is a bivariate random vector with the joint probability density function $f_{X,Y}(x, y)$. Then,

1. A random sample of size n^4 is identified from the population and randomly allocated into n^2 pools each of size n^2 so that each pool is a square matrix with n rows and n columns.
2. In the first pool, identify the minimum value by judgment with respect to the first characteristic X, for each of the n rows.
3. For the n minima obtained in Step 2, the actual quantification is done on the pair that corresponds to the minimum value of the second characteristic, Y, identified by judgment. This pair, given the label (1, 1), is the first element of the BVRSS sample.
4. Repeat Steps 2 and 3 for the second pool, but in Step 3, the pair corresponding to the second minimum value with respect to the second characteristic, Y, is chosen for actual quantification. This pair is given the label (1, 2).
5. The process continues until the label (n, n) is ascertained from the n^2th (last) pool.

The procedure described above produces a BVRSS of size n^2. Let $[(X_{[i](j)}, Y_{(i)[j]})$, $i = 1, 2, \ldots, n$ *and* $j = 1, 2, \ldots, n]$ denote the BVRSS sample from $f_{X,Y}(x, y)$ where $f_{X_{[i](j)}}, Y_{(i)[j]}(x, y)$ is the joint probability density function of $(X_{[i](j)} Y_{(i)[j]})$. From Al-Saleh and Zheng (2002),

$$f_{X_{[i](j)}}, Y_{(i)[j]}(x, y) = f_{Y_{(i)[j]}}(y) \frac{f_{X_{(j)}}(x) f_{Y|X}(y|x)}{f_{Y_{[j]}}(y)}, \tag{1.2}$$

where $f_{X_{(j)}}$ is the density of the jth order statistic for an SRS sample of size n from the marginal density of f_X and $f_{Y_{[j]}}(y)$ be the density of the corresponding $Y-$ value given by $f_{Y_{[j]}}(y) = \int\limits_{-\infty}^{\infty} f_{X_{(j)}}(x) f_{Y|X}(y|x)dx$, while $f_{Y_{(i)[j]}}(y)$ is the density of the ith order statistic of an iid sample from $f_{Y_{[j]}}(y)$, i.e.

$$f_{Y_{(i)[j]}}(y) = \mathrm{c.}(\mathrm{F}_{Y_{[j]}}(y))^{i-1}(1 - \mathrm{F}_{Y_{[j]}}(y))^{n-i} f_{Y_{[j]}}(y)$$

where $F_{Y_{[j]}}(y) = \int\limits_{-\infty}^{y} (\int\limits_{-\infty}^{\infty} f_{X_{(j)}}(x) f_{Y|X}(w|x)dx)dw$.

Combining these results, Eq. (1.2) can be written as

$$f_{X_{[i](j)}}, Y_{(i)[j]}(x, y) = c_1 (F_{Y_{[j]}}(y))^{i-1}(1 - F_{Y_{[j]}}(y))^{n-i}(F_X(x))^{j-1}(1 - F_X(x))^{n-j} f(x, y) \tag{1.3}$$

where

$$c_1 = \left(\frac{n!}{(i-1)!(n-i)!}\right)\left(\frac{n!}{(j-1)!(n-j)!}\right).$$

Furthermore, Al-Saleh and Zheng (2002) showed that,

$$\frac{1}{n^2} \sum_j^n \sum_i^n f_{X_{[i](j)}, Y_{(i)[j]}}(x, y) = f(x, y). \tag{1.4}$$

For a variety of choices of $f(u, v)$, one can have (U, V) bivariate uniform with a probability density function *f(u, v); 0 < u, v < 1*, such that $U \sim U(0, 1)$ and $V \sim U(0, 1)$ (See Johnson 1987). In that case, $[(U_{[i](j)}, V_{(i)[j]})$, $i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, n]$ should have a bivariate probability density function given by

$$f_{(j),(i)}(u, v) = \left[\frac{n!}{(i-1)!(n-i)!}\right]\left[\frac{n!}{(j-1)!(n-j)!}\right][F_{Y_{[j]}}(v)]^{i-1}[1 - F_{Y_{[j]}}(v)]^{n-i}[u]^{j-1} [1-u]^{n-j} f(u, v). \tag{1.5}$$

Samawi and Al-Saleh (2007) extended the work of Samawi (1999) and Al-Saleh and Zheng (2002) for the Monte-Carlo multiple integration approximation of (1.1) when k = 2.

Moreover, to further improve some of the Monte-Carlo methods of integration, Al-Saleh and Samawi (2000) used steady-state ranked set simulated sampling (SRSIS) as introduced by Al-Saleh and Al-Omari (1999). SRSIS has been shown to be simpler and more efficient than Samawi's (1999) method.

In Samawi and Vogel (2013) work, the SRSIS algorithm introduced by Al-Saleh and Samawi (2000) was extended to multivariate case for the approximation of multiple integrals using Monte-Carlo methods. However, to simplify the algorithms, we introduce only the bivariate integration problem; with this foundation, multiple integral problems are a simple extension.

2 Steady-State Ranked Simulated Sampling (SRSIS)

Al-Saleh and Al-Omari (1999) introduced the idea of multistage ranked set sampling (MRSS). To promote the use of MRSS in simulation and Monte-Carlo methods, let $\{X_i^{(s)}; i = 1, 2, \ldots, n, \}$ be an MRSS of size n at stage s. Assume that $X_i^{(s)}$ has probability density function $f_i^{(s)}$ and a cumulative distribution function $F_i^{(s)}$. Al-Saleh and Al-Omeri demonstrated the following properties of MRSS:

1.
$$f(x) = \frac{1}{n}\sum_{i=1}^{n} f_i^{(s)}(x), \tag{2.1}$$

2.
$$\text{If } s \to \infty, \text{ then } \mathrm{F}_i^{(s)}(x) \to \mathrm{F}_i^{(\infty)}(x) = \begin{cases} 0 & if \quad x < Q_{(i-1)/n} \\ nF(x) - (i-1) & if \quad Q_{(i-1)/n} \le x < Q_{(i)/n} \\ 1 & if \quad x \ge Q_{(i)/n} \end{cases}, \tag{2.2}$$

for $i = 1, 2, \ldots, n$, where Q_α is the 100αth percentile of $F(x)$.

3. If $X \sim U(0, 1)$, then for $i = 1, 2, ..., n$, we have

$$F_i^{(\infty)}(x) = \begin{cases} 0 & if \quad x < (i-1)/n \\ nx - (i-1) & if \quad (i-1)/n \le x < i/n \\ 1 & if \quad x \ge i/n \end{cases}, \tag{2.3}$$

and

$$f_i^{(\infty)}(x) = \begin{cases} n \; if \;\; (i-1)/n \le x < i/n \\ 0 \qquad otherwise. \end{cases} \tag{2.4}$$

These properties imply $\mathrm{X}_i^{(\infty)} \sim U(\frac{i-1}{n}, \frac{i}{n})$, when the underlying distribution function is U(0, 1).

Samawi and Vogel (2013) provided a modification of the Al-Saleh and Samawi (2000) steady-state ranked simulated samples procedure (SRSIS) to bivariate cases (BVSRSIS) as follows:

1. For each (i, j), $j = 1, 2, \ldots, n$ and $i = 1, 2, \ldots, n$ generate independently

   a. ($U_{i(j)}$from U $\left(\frac{j-1}{n}, \frac{j}{n}\right)$ and independent $W_{(i)j}$ from U $\left(\frac{i-1}{n}, \frac{i}{n}\right)$, $i = 1$, 2, ..., n).

2. Generate $Y_{(i)j} = F_Y^{-1}(W_{(i)j})$ and $X_{i(j)} = F_X^{-1}(U_{i(j)})$ from $F_Y(y)$ and $F_X(x)$ respectively.
3. To generate $(X_{[i](j)}, Y_{(i)[j]})$ from $f(x, y)$, generate $U'_{i(j)}$from U $\left(\frac{j-1}{n}, \frac{j}{n}\right)$ and independent $W'_{(i)j}$ from U $\left(\frac{i-1}{n}, \frac{i}{n}\right)$, then

$$X_{[i](j)}|Y_{(i)j} = F^{-1}_{X|Y}(U'_{i(j)}|Y_{(i)j}) \text{ and } Y_{(i)[j]}|X_{i(j)} = F^{-1}_{Y|X}(W'_{(i)j}|X_{i(j)}).$$

The joint density function of $(X_{[i](j)}, Y_{(i)[j]})$ is formed as follows:

$$f^{(\infty)}_{X_{[i](j)}Y_{(i)[j]}}(x, y) = f^{(\infty)}_{X_{[i](j)}}(x) f^{(\infty)}_{Yi|X_{[i](j)}}(y|X_{[i](j)}) = n^2 f_X(x) f_{Y|X_{[i](j)}}(y|X_{[i](j)}),$$
$$Q_{X(j-1)/n} \leq x < Q_{X(j)/n}, Q_{Y(i-1)/n} \leq y < Q_{Y(i)/n},$$

where $Q_{X(s)}$ and $Q_{Y(v)}$ are the 100 sth percentile of $F_X(x)$ and 100 vth percentile of $F_Y(y)$, respectively. However, for the first stage, both Stokes (1977) and David (1981) showed that $F_{Y|X_{[i](j)}}(y|x) = F_{Y|X}(y|x)$. Al-Saleh and Zheng (2003) demonstrated that joint density is valid for an arbitrary stage, and therefore, valid for a steady state. Therefore,

$$f^{(\infty)}_{X_{[i](j)}Y_{(i)[j]}}(x, y) = f^{(\infty)}_{X_{[i](j)}}(x) f^{(\infty)}_{Yi|X_{[i](j)}}(y|X_{[i](j)}) = n^2 f_X(x) f_{Y|X}(y|x) = n^2 f_{Y,X}(x, y),$$
$$Q_{X(j-1)/n} \leq x < Q_{X(j)/n}, Q_{Y(i-1)/n} \leq y < Q_{Y(i)/n}. \tag{2.5}$$

Thus, we can write:

$$\frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} f^{(\infty)}_{X_{[i](j)}Y_{(i)[j]}}(x, y) = \sum_{i=1}^{n} \sum_{j=1}^{n} f_{Y,X}(x, y).I[Q_{X(j-1)/n} \leq x < Q_{X(j)/n}] I[Q_{Y(i-1)/n} \leq y < Q_{Y(i)/n}]$$
$$= f(x, y), \tag{2.6}$$

where I is an indicator variable. Similarly, Eq. (2.5) can be extended by mathematical induction to the multivariate case as follows:

$f^{(\infty)}(x_1, x_2, ..., x_k) = n^k f(x_1, x_2, ..., x_k)$, $Q_{X_i(j-1)/n} \leq x_i < Q_{X_i(j)/n}$, $i = 1, \ldots, k$ and $j = 1, 2, .., n$.. In addition, the above algorithm can be extended for k > 2 as follows:

1. For each $(i_l, l = 1, 2, ..., k)$, $i_l = 1, 2, ..., n$ generate independently

$$U_{i_l(i_s)}\text{from U}\left(\frac{i_s - 1}{n}, \frac{i_s}{n}\right), l, s = 1, 2, \ldots, k \text{ and } i_l, i_s = 1, 2, \ldots, n.$$

2. Generate $X_{i_l(i_s)} = F_{X_{i_l}}^{-1}(U_{i_l(i_s)})l, s = 1, 2, \ldots, k$ and $i_l, i_s = 1, 2, \ldots, n,$ from $F_{X_{i_l}}(x), l = 1, 2, \ldots, k$, respectively.
3. Then, generate the multivariate version of the steady-state simulated sample by using any technique for conditional random number generation.

3 Monte-Carlo Methods for Multiple Integration Problems

Very good descriptions of the basics of the various Monte-Carlo methods have been provided by Hammersley and Handscomb (1964), Liu (2001), Morgan (1984), Robert and Casella (2004), and Shreider (1966). The Monte-Carlo methods described include crude, antithetic, importance, control variate, and stratified sampling approaches. However, when variables are related, Monte-Carlo methods cannot be used directly (i.e., similar to the manner that these methods are used in univariate integration problems) because using the bivariate uniform probability density function $f(u, v)$ as a sampler to evaluate Eq. (1.1) with $k = 2$, $f(u, v)$ is not consistent. However, in this context it is reasonable to use the importance sampling method, and therefore, it follows that other Monte-Carlo techniques can be used in conjunction with importance sampling. Thus, our primary concern is importance sampling.

3.1 Importance Sampling Method

In general, suppose that f is a density function on R^k such that the closure of the set of points where $g(.)$ is non-zero and the closure set of points where $f(.)$is non-zero. Let $[\underline{U}_{\underline{i}}\ i = 1, 2, ..., n]$ be a sample from $f(.)$. Then, because

$$\theta = \int \frac{g(\underline{u})}{f(\underline{u})} f(\underline{u}) d\underline{u},$$

Equation (1.1) can be estimated by

$$\hat{\theta} = \frac{1}{n} \sum_{i=1}^{n} \frac{g(\underline{u}_i)}{f(\underline{u}_i)} \quad . \tag{3.1}$$

Equation (3.1) is an unbiased estimator for (1.1), with variance given by

$$Var(\hat{\theta}) = \frac{1}{n}\left(\int_{R^k} \frac{g(\underline{u})^2}{f(\underline{u})} d\underline{u} - \theta^2\right).$$

In addition, from the point of view of the strong law of large numbers, it is clear that $\hat{\theta} \rightarrow \theta$ almost surely as $n \rightarrow \infty$.

A limited number of distributional families exist in a multidimensional context and are commonly used as importance samplers. For example, the multivariate Student's family is used extensively in the literature as an importance sampler. Evans and Swartz (1995) indicated a need for developing families of multivariate distribution that exhibit a wide variety of shapes. In addition, statisticians want distributional families to have efficient algorithms for random variable generation and the capacity to be easily fitted to a specific integrand.

This paper provides a new way of generating a bivariate sample based on the bivariate steady-state sampling (BVSRSIS) that has the potential to extend the existing sampling methods. We also provide a means for introducing new samplers and to substantially improve substantially the efficiency of the integration approximation based on those samplers.

3.2 Using Bivariate Steady-State Sampling (BVSRSIS)

Let

$$\theta = \int g(x, y)dx\, dy. \tag{3.2}$$

To estimate θ, generate a bivariate sample of size n^2 from *f(x, y),* which mimics $g(x, y)$ and has the same range, such as $[(X_{ij}, Y_{ij}), i = 1,\ 2, ..., \ n\, and$ $j = 1, 2,\ \ ..., n]$. Then

$$\hat{\theta} = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{g(x_{ij}, y_{ij})}{f(x_{ij}, y_{ij})}. \tag{3.3}$$

Equation (3.3) is an unbiased estimate for (3.2) with variance

$$Var(\hat{\theta}) = \frac{1}{n^2}(\int\int\frac{g^2(x, y)}{f(x, y)}dx\, dy - \theta^2). \tag{3.4}$$

To estimate (3.2) using BVSRSIS, generate a bivariate sample of size n^2, as described in above, say $[(X_{[i](j)},\ Y_{(i)[j]}),\ i = 1,\ 2,\ ...,\ n\ and\ j = 1,\ 2,\ ...,\ n]$. Then

$$\hat{\theta}_{BVSRSIS} = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{g(x_{[i](j)},\ y_{(i)[j]})}{f(x_{[i](j)},\ y_{(i)[j]})}. \tag{3.5}$$

Equation (3.5) is also an unbiased estimate for (3.2) using (2.5). Also, by using (2.5) the variance of (3.5) can be expressed as

$$Var(\hat{\theta}_{BVSRSIS}) = Var(\hat{\theta}) - \frac{1}{n^4}\sum_{i=1}^{n}\sum_{j=1}^{n}(\theta_{g/f}^{(i,j)} - \theta_{g/f})^2, \tag{3.6}$$

where, $\theta_{g/f}^{(i,j)} = E[g(X_{[i](j)}, Y_{(i)[j]})/f(X_{[i](j)}, Y_{(i)[j]})], \theta_{g/f} = E[g(X, Y)/f(X, Y)] = \theta$. The variance of the estimator in (3.6) is less than the variance of the estimator in (3.4).

3.3 Simulation Study

This section presents the results of a simulation study that compares the performance of the importance sampling method described above using BVSRSIS schemes with the performance of the bivariate simple random sample (BVUSS) and BVRSS schemes by Samawi and Al-Saleh (2007) as introduced by Samawi and Vogel (2013).

3.3.1 Illustration for Importance Sampling Method When Integral's Limits Are (0, 1)x(0, 1)

As in Samawi and Al-Saleh (2007), illustration of the impact of BVSRSIS on importance sampling is provided by evaluating the following integral

$$\theta = \int_0^1\int_0^1 (1+v).\exp(u(1+v))\,du\,dv = 3.671. \tag{3.7}$$

This example uses four bivariate sample sizes: n= 20, 30, 40 and 50. To estimate the variances using the simulation method, we use 2,000 simulated samples from BVUSS and BVSRSIS. Many choices of bivariate and multivariate distributions with uniform marginal on [0, 1] are available (Johnson 1987). However, for this simulation, we chose Plackett's uniform distribution (Plackett 1965), which is given by

$$f(u, v) = \frac{\psi\{(\psi - 1)(u + v - 2\,uv) + 1\}}{\{[1 + (u + v)(\psi - 1)]^2 - 4\psi(\psi - 1)uv\}^{3/2}},\; 0 < u,\; v < 1,\; \psi > 0. \tag{3.8}$$

The parameter ψ governs the dependence between the components (U, V) distributed according to f. Three cases explicitly indicate the role of ψ (Johnson 1987):

$$\begin{array}{ll} \psi \to 0 & U = 1 - V, \\ \psi = 1 & U \text{ and } V \text{ are independent}, \\ \psi \to \infty & U = V, \end{array}$$

Table 1 Efficiency of estimating (3.7) using BVSRSIS relative to BVUSS and BVRSS

$n\backslash\psi$	1	2
20	289.92 (**8.28**)	273.68 (**9.71**)
30	649.00 (**12.94**)	631.31 (**13.06**)
40	1165.31 (**16.91**)	1086.46 (**18.60**)
50	1725.25 (**21.67**)	1687.72 (**23.03**)

Note Values shown in *bold* were extracted from Samawi and Al-Saleh (2007)

Table 1 presents the relative efficiencies of our estimators using BVRSIS in comparison with using BVUSS and BVSRSIS relative to BVUSS for estimating (3.7).

As illustrated in Table 1, BVSRSIS is clearly more efficient than either BVUSS or BVRSIS when used for estimation.

3.3.2 Illustration When the Integral's Limits Are Arbitrary Subset of R^2

Recent work by Samawi and Al-Saleh (2007) and Samawi and Vogel (2013) used an identical example in which the range of the integral was not (0, 1), and the authors evaluated the bivariate normal distribution (e.g., $g(x, y)$ is the N_2 *(0, 0, 1, 1, ρ)* density.) For integrations with high dimensions and a requirement of low relative error, the evaluation of the multivariate normal distribution function remains one of the unsolved problems in simulation (e.g., Evans and Swartz 1995). To demonstrate how BVSRSIS increases the precision of evaluating the multivariate normal distribution, we illustrate the method by evaluating the bivariate normal distribution as follows:

$$\theta = \int_{-\infty}^{z_1} \int_{-\infty}^{z_2} g(x, y)\, dx\, dy, \tag{3.9}$$

where $g(x, y)$ is the $N_2(0,\ 0,\ 1,\ 1,\ \rho)$ density.

Given the similar shapes of the marginal of the normal and the marginal of the logistic probability density functions, it is natural to attempt to approximate the bivariate normal cumulative distribution function by the bivariate logistic cumulative distribution function. For the multivariate logistic distribution and its properties, see Johnson and Kotz (1972). The density of the bivariate logistic (Johnson and Kotz 1972) is chosen to be

$$f(x, y) = \frac{2!\, \pi^2 e^{-\pi(x+y)/\sqrt{3}} (1 + e^{-\pi z_1/\sqrt{3}} + e^{-\pi z_2/\sqrt{3}})}{3(1 + e^{-\pi x/\sqrt{3}} + e^{-\pi y/\sqrt{3}})^3}, \quad -\infty < x\ < z_1;\ \ -\infty < y < z_2. \tag{3.10}$$

It can be shown that the marginal of X is given by

$$f(x)=\frac{\pi e^{-\pi x/\sqrt{3}}(1+e^{-\pi z_1/\sqrt{3}}+e^{-\pi z_2/\sqrt{3}})}{\sqrt{3}\,(1+e^{-\pi x/\sqrt{3}}+e^{-\pi z_2/\sqrt{3}})^2},\quad -\infty<x\ <z_1. \tag{3.11}$$

Now let $W=Y+\frac{\sqrt{3}}{\pi}\ln(1+e^{-\pi X/\sqrt{3}}+e^{-\pi z_2/\sqrt{3}})$. Then it can be shown that

$$\begin{gathered} f(w|x)=\frac{2\,\pi\ e^{-\pi w/\sqrt{3}}}{\sqrt{3}\left(\frac{1+e^{-\pi x/\sqrt{3}}}{1+e^{-\pi x/\sqrt{3}}+e^{-\pi z_2/\sqrt{3}}}+e^{-\pi w/\sqrt{3}}\right)^3}, \\ -\infty<w<z_2+\frac{\sqrt{3}}{\pi}\ln\left(1+e^{-\pi x/\sqrt{3}}+e^{-\pi z_2/\sqrt{3}}\right). \end{gathered} \tag{3.12}$$

To generate from (3.10) proceed as follows:

1. Generate X from (3.11).
2. Generate W independently from (3.12)
3. Let $Y=W-\frac{\sqrt{3}}{\pi}\ln(1+e^{-\pi X/\sqrt{3}}+e^{-\pi z_2/\sqrt{3}})$.
4. Then the resulting pair (X,Y) has the correct probability density function, as defined in (3.10).

For this illustration, two bivariate sample sizes, $n=20$ and 40, and different values of ρ and $(z_{1,}z_2)$ are used. To estimate the variances using simulation, we use 2,000 simulated samples from BVUSS, BVRSIS, and BVSRSIS (Tables 2 and 3).

Notably, when Samawi and Vogel (2013) used identical examples to those used by Samawi and Al-Saleh (2007), a comparison of the simulations showed that Samawi and Vogel (2013) BVSRSIS approach improved the efficiency of estimating the multiple integrals by a factor ranging from 2 to 100.

As expected, the results of the simulation indicated that using BVSRSIS substantially improved the performance of the importance sampling method for integration

Table 2 Efficiency of using BVSRSIS to estimate Eq. (3.9) relative to BVUSS

(z_1,z_2)	$n=20$				$n=40$			
	$\rho=\pm 0.20$	$\rho=\pm0.50$	$\rho=\pm0.80$	$\rho=\pm0.95$	$\rho=\pm0.20$	$\rho=\pm0.50$	$\rho=\pm0.80$	$\rho=\pm0.95$
(0, 0)	5.39 **(6.42)**	9.89 **(21.70)**	5.98 **(144.31)**	49.65 **(171.22)**	9.26 **(12.2)**	22.50 **(63.74)**	15.80 **(526.10)**	158.77 **(612.50)**
$(-1,-1)$	22.73 **(75.82)**	29.43 **(182.27)**	22.48 **(87.42)**	95.10 **(23.71)**	55.99 **(255.75)**	90.24 **(688.59)**	69.85 **(336.58)**	380.87 **(100.40)**
$(-2,-2)$	148.30 **(200.46)**	133.21 **(143.62)**	125.37 **(30.45)**	205.99 **(42.75)**	506.63 **(759.81)**	408.66 **(568.77)**	411.19 **(128.94)**	802.73 **(45.55)**
$(-1,-2)$	173.07 **(91.08)**	281.60 **(42.89)**	216.86 **(08.47)**	24.11[a]	714.92 **(382.01)**	1041.39 **(148.16)**	882.23 **(43.19)**	98.76[a]

Note Values shown in *bold* are results of negative correlations coefficients

[a]Values cannot be obtained due the steep shape of the bivariate distribution for the large negative correlation

Table 3 Relative efficiency of estimating Eq. (3.9) using BVRSIS as compared with using BVUSS

(z_1, z_2)	n = 20				n = 40			
	ρ = 0.20	ρ = 0.50	ρ = 0.80	ρ = 0.95	ρ = 0.20	ρ = 0.50	ρ = 0.80	ρ = 0.95
(0, 0)	2.39	3.02	2.29	3.21	3.80	4.85	3.73	5.66
(−1, −1)	4.73	4.30	4.04	3.79	7.01	8.28	7.83	7.59
(−2, −2)	8.44	8.47	8.73	5.65	15.69	15.67	16.85	11.02

Source Extracted from Samawi and Al-Saleh (2007)

approximation. BVSRSIS also outperformed the BVRSIS method used by Samawi and Al-Saleh (2007). Moreover, increasing the sample size in both of the above illustrations increases the relative efficiencies of these methods. For instance, in the first illustration, by increasing the sample size from 20 to 50, the relative efficiency of using BVSRSIS as compared with BVUSS to estimate (3.10) is increased from 289.92 to 649.00, with the increase dependent on the dependency between U and V. A similar relationship between sample size and relative efficiencies of these two methods can be demonstrated in the second illustration.

Based on the above conclusions, BVSRSIS can be used in conjunction with other multivariate integration procedures to improve the performance of those methods, and thus providing researchers with a significant reduction in required sample size. With the use of BVSRSIS, researchers can perform integral estimation using substantially fewer simulated numbers. Since using BVSRSIS in simulation does not require any extra effort or programming, we recommend using BVSRSIS to improve the well-known Monte-Carlo method of numerical multiple integration problems. Using BVSRSIS will yield an unbiased and more efficient estimate of those integrals. Moreover, this sampling scheme can be applied successfully to other simulation problems. Last, we recommend using the BVSRSIS method for integrals with a dimension no greater than 3. For higher dimensional integrals, other methods in the literature can be used in conjunction with independent steady ranked simulated sampling.

4 Steady-State Ranked Gibbs Sampler

Many approximation techniques are found in the literature, including Monte-Carlo methods, asymptotic, and Markov chain Monte-Carlo (MCMC) methods such as the Gibbs sampler (Evans and Swartz 1995). Recently, many statisticians have become interested in MCMC methods to simulate complex, nonstandard multivariate distributions. Of the MCMC methods, the Gibbs sampling algorithm is one of the best known and most frequently used MCMC method. The impact of the Gibbs sampler method on Bayesian statistics has been detailed by many authors (e.g., Chib and Greenberg 1994; Tanner 1993) following the work of Tanner and wong (1987) and Gelfand and Smith (1990).

To understand the MCMC process, suppose that we need to evaluate the Monte-Carlo integration *E[f(X)],* where $f(.)$ is any user-defined function of a random variable X. The MCMC process is as follows: Generate a sequence of random variables, $\{X_0, X_1, X_2, \ldots\}$, such that at each time $t \geq 0$, the next state X_{t+1} is sampled from a distribution $P(X_{t+1}|X_t)$ which depends only on the current state of the chain, X_t. This sequence is called a Markov chain, and $P(.|.)$ is called the transition kernel of the chain. The transition kernel is a conditional distribution function that represents the probability of moving from X_t to the next point X_{t+1} in the support of X. Assume that the chain is time homogenous. Thus, after a sufficiently long burn-in of k iterations, $\{X_t; t = k + 1, \ldots, n\}$ will be dependent samples from the stationary distribution. Burn-in samples are usually discarded for this calculation, given an estimator,

$$\bar{f} \approx \frac{1}{n-k} \sum_{t=k+1}^{n} f(\underline{X}_t). \tag{4.1}$$

This average in (4.1) is called an ergodic average. Convergence to the required expectation is ensured by the ergodic theorem. More information and discussions on some of the issues in MCMC can be found in Roberts (1995) and Tierney (1995).

To understand how to construct a Markov chain so that its stationary distribution is precisely the distribution of interest $\pi(.)$, we outline Hastings' (1970) algorithm, which is a generalization of the method first proposed by Metropolis et al. (1953). The method is useful for obtaining a sequence of random samples from a probability distribution for which direct sampling is difficult. The method is as follows: At each time t, the next state X_{t+1} is chosen by first sampling a candidate point Y from a proposal distribution $q(.|X_t)$ (ergodic). Note that the proposal distribution may depend on the current point X_t. The candidate point Y is then accepted with probability $\alpha(X_t, Y)$ where

$$\alpha(X_t, Y) = \min\left(1, \frac{\pi(Y)q(X_t|Y)}{\pi(X_t)q(Y|X_t)}\right). \tag{4.2}$$

If the candidate point is accepted, the next state becomes $X_{t+1} = Y$. If the candidate is rejected, the chain does not move, that is, $X_{t+1} = X_t$. Thus the Metropolis–Hastings algorithm simply requires the following:

Initialize X_0; set t = 0.
Repeat {generate a candidate Y from $q(.|X_t)$
and a value u from a uniform (0, 1), if
$u \leq \alpha(X_t, Y)$ set $X_{t+1} = Y$
Otherwise set $X_{t+1} = X_t$
Increment t}.

A special case of the Metropolis–Hastings algorithm is the Gibbs sampling method proposed by Geman and Geman (1984) and introduced by Gelfand and Smith (1990). To date, most statistical applications of MCMC have used Gibbs sampling. In Gibbs sampling, variables are sampled one at a time from their full conditional distributions.

Gibbs sampling uses an algorithm to generate random variables from a marginal distribution indirectly, without calculating the density. Similar to Casella and George (1992), we demonstrate the usefulness and the validity of the steady-state Gibbs sampling algorithm by exploring simple cases. This example shows that steady-state Gibbs sampling is based only on elementary properties of Markov chains and the properties of BVSRSIS.

4.1 Traditional (standard) Gibbs Sampling Method

Suppose that $f(x, y_1, y_2, \ldots, y_g)$ is a joint density function on R^{g+1} and our purpose is to find the characteristics of the marginal density such as the mean and the variance.

$$f_X(x) = \int \ldots \int f(x, y_1, y_2, \ldots, y_g) dy_1, dy_2, \ldots, dy_g \tag{4.3}$$

In cases where (4.3) is extremely difficult or not feasible to perform either analytically or numerically, Gibbs sampling enables the statistician to efficiently generate a sample $X_1, \ldots, X_n \sim f_x(x)$, without requiring $f_x(x)$. If the sample size n is large enough, this method will provide a desirable degree of accuracy for estimating the mean and the variance of $f_x(x)$.

The following discussion of the Gibbs sampling method uses a two-variable case to make the method simpler to follow. A case with more than two variables is illustrated in the simulation study.

Given a pair of random variables (X, Y), Gibbs sampling generates a sample from $f_X(x)$ by sampling from the conditional distribution, $f_{X|Y}(x|y)$ and $f_{Y|X}(y|x)$, which are usually known in statistical models application. The procedure for generating a Gibbs sequence of random variables,

$$X'_0, Y'_0, X'_1, Y'_1, \ldots, X'_k, Y'_k, \tag{4.4}$$

is to start from an initial value $Y'_0 = y'_0$, (which is a known or specified value) and obtaining the rest of the sequence (4.4) iteratively by alternately generating values from

$$\begin{aligned} X'_j &\sim f_{X|Y}(x|Y'_j = y'_j) \\ Y'_{j+1} &\sim f_{Y|X}(y|X'_j = x'_j). \end{aligned} \tag{4.5}$$

For large k and under reasonable conditions (Gelfand and Smith 1990), the final observation in (4.5), namely $X'_j = x'_j$, is effectively a sample point from $f_X(x)$. A natural way to obtain an independent and identically distributed (i.i.d) sample from $f_X(x)$ is to follow the suggestion of Gelfand and Smith (1990) to use Gibbs sampling to find the kth, or final value, from n independent repetitions of the Gibbs sequence in (4.5). Alternatively, we can generate one long Gibbs sequence and use a systematic sampling technique to extract every rth observation. For large enough r, this method will also yield an approximate i.i.d sample from $f_X(x)$. For the advantage and disadvantage of this alternate method see, Gelman and Rubin (1991).

Next, we provide a brief explanation of why Gibbs sampling works under reasonable conditions. Suppose we know the conditional densities $f_{X|Y}(x|y)$ and $f_{Y|X}(y|x)$ of the two random variables X and Y, respectively. Then the marginal density of X, $f_X(x)$ can be determined as follows:

$$f_X(x) = \int f(x, y)dy,$$

where $f(x, y)$ is the unknown joint density of (X, Y). Using the fact that $f_{XY}(x, y) = f_Y(y).f_{Y|X}(x|y)$, then

$$f_X(x) = \int f_Y(y).f_{Y|X}(x|y)dy.$$

Using a similar argument for $f_Y(y)$, then

$$f_X(x) = \int \left[\int f_{X|Y}(x|y) f_{Y|X}(y|t)dy \right] f_X(t)dt = \int g(x, t) f_X(t)dt, \tag{4.6}$$

where $g(x, t) = \int f_{X|Y}(x|y) f_{Y|X}(y|t)dy$ As argued by Gelfand and Smith (1990), Eq. (4.6) defines a fixed-point integral equation for which $f_X(x)$ is the solution and the solution is unique.

4.2 *Steady-State Gibbs Sampling (SSGS): The Proposed Algorithms*

To guarantee an unbiased estimator for the mean, density, and the distribution function of $f_X(x)$, Samawi et al. (2012) introduced two methods for performing steady-state Gibbs sampling. The first method is as follows:

In standard Gibbs sampling, the Gibbs sequence is obtained using the conditional distribution, $f_{X|Y}(x|y)$ and $f_{Y|X}(y|x)$, to generate a sequence of random variables,

$$X'_0, Y'_0, X'_1, Y'_1, \ldots, X'_{k-1}, Y'_{k-1}, \tag{4.7}$$

starting from an initial, specified value $Y'_0 = y'_0$ and iteratively obtaining the rest of the sequence (4.7) by alternately generating values from

$$\begin{aligned} X'_j &\sim f_{X|Y}(x|Y'_j = y'_j) \\ Y'_{j+1} &\sim f_{Y|X}(y|X'_j = x'_j). \end{aligned} \tag{4.8}$$

However, in steady state Gibbs sampling (SSGS), the Gibbs sequence is obtained as follows:

One step before the kth step in the standard Gibbs sampling method, take the last step as

$$\begin{aligned} X'_{i(j)} &\sim F^{-1}_{X|Y}(U_{i(j)}|Y'_{k-1} = y'_{k-1}) \\ Y'_{(i)j} &\sim F^{-1}_{Y|X}(W_{(i)j}|X'_{k-1} = x'_{k-1}), \\ X_{[i](j)} &\sim F^{-1}_{X|Y}(U'_{i(j)}|Y'_{(i)j} = y'_{(i)j}) \\ Y_{(i)[j]} &\sim F^{-1}_{Y|X}(W'_{(i)j}|X'_{i(j)} = x'_{i(j)}) \end{aligned} \tag{4.9}$$

where $\{U_{i(j)}, U'_{i(j)}\}$ from $U\left(\frac{j-1}{n}, \frac{j}{n}\right)$ and $\{W_{(i)j}, W'_{(i)j}\}$ from $U\left(\frac{i-1}{n}, \frac{i}{n}\right)$ as described above. Clearly, this step does not require extra computer time since we generate the Gibbs sequences from uniform distributions only. Repeat this step independently for $i = 1, 2, \ldots, n$ *and* $j = 1, 2, \ldots, n$ to get an independent sample of size n^2, namely $[(X_{[i](j)}, Y_{(i)[j]}), i = 1, 2, \ldots, n$ *and* $j = 1, 2, \ldots, n]$. For large k and under reasonable conditions (Gelfand and Smith 1990), the final observation in Eq. (4.9), namely $(X_{[i](j)} = x_{[i](j)}, Y_{(i)[j]} = y_{(i)[j]})$ is effectively a sample point from (2.5). Using the properties of SRSIS, $[(X_{[i](j)}, Y_{(i)[j]}), i = 1, 2, \ldots, n$ *and* $j = 1, 2, \ldots, n]$, will produce unbiased estimators for the marginal means and distribution functions. Alternatively, we can generate one long standard Gibbs sequence and use a systematic sampling technique to extract every rth observation using a similar method as described above. Again, a SSGS sample will be obtained as

$$\begin{aligned} X'_{i(j)} &\sim F^{-1}_{X|Y}(U_{i(j)}|Y'_{r-1} = y'_{r-1}), \\ Y'_{(i)j} &\sim F^{-1}_{Y|X}(W_{(i)j}|X'_{r-1} = x'_{r-1}), \\ X_{[i](j)} &\sim F^{-1}_{X|Y}(U'_{i(j)}|Y'_{(i)j} = y'_{(i)j}), \\ Y_{(i)[j]} &\sim F^{-1}_{Y|X}(W'_{(i)j}|X'_{i(j)} = x'_{i(j)}), \end{aligned} \tag{4.10}$$

where $\{U_{i(j)}, U'_{i(j)}\}$ from $U\left(\frac{j-1}{n}, \frac{j}{n}\right)$ and $\{W_{(i)j}, W'_{(i)j}\}$ from $U\left(\frac{i-1}{n}, \frac{i}{n}\right)$ to obtain an independent sample of size n^2, that is, $[(X_{[i](j)}, Y_{(i)[j]}), i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, n]$.

Using the same arguments as in (4.1), suppose we know the conditional densities $f_{X|Y}(x|y)$ and $f_{Y|X}(y|x)$ of the two random variables X and Y, respectively. Equation (4.6) is the limiting form of the Gibbs iteration scheme, showing how sampling from conditionals produces a marginal distribution. As in Gelfand and Smith (1990) for $k \to \infty$, $X'_{k-1} \sim f_X(x)$ and $Y'_{k-1} \sim f_Y(y)$ and hence $F^{-1}_{Y|X}(W_{(i)j}|x'_{k-1}) = Y'_{(i)j} \sim f^{\infty}_{Y_{(i)}}(y)$, $F^{-1}_{Y|X}(U_{i(j)}|x'_{k-1}) = X'_{(i)j} \sim f^{\infty}_{X_{(j)}}(x)$ where $W_{(i)j} \sim U\left(\frac{i-1}{n}, \frac{i}{n}\right)$ and $U_{i(j)} \sim U\left(\frac{j-1}{n}, \frac{j}{n}\right)$. Therefore,

$$F^{-1}_{X|Y'_{(i)j}}(U'_{i(j)}|Y'_{(i)j}) = X_{[i](j)}|Y'_{(i)j} \sim f^{\infty}_{X_{[i](j)}|Y'_{(i)j}}(x|Y'_{(i)j}) \text{ and}$$
$$F^{-1}_{Y|X'_{i(j)}}(W'_{(i)j}|X'_{i(j)}) = Y_{(i)[j]}|X'_{i(j)} \sim f^{\infty}_{Y_{(i)[j]}|X'_{i(j)}}(y|X'_{i(j)}).$$

This step produces an independent bivariate steady-state sample, $[(X_{[i](j)}, Y_{(i)[j]}), i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, n]$, where some characteristic of the marginal distributions are to be investigated. To see how to apply this bivariate steady-state sample Gibbs sampling, using (2.5) we get

$$\begin{aligned} f_{X[i](j)}(x) &= \int_{Q_{\frac{(i-1)}{n}}}^{Q_{\frac{(i)}{n}}} n^2 f_Y(y).f_{X|Y}(x|y)dy = \int_{Q_{\frac{(i-1)}{n}}}^{Q_{\frac{(i)}{n}}} n f_{X|Y}(x|y) n \int_{Q_{\frac{(j-1)}{n}}}^{Q_{\frac{(j)}{n}}} f_X(t) f_{Y|X}(y|t)dtdy \\ &= \int_{Q_{\frac{(j-1)}{n}}}^{Q_{\frac{(j)}{n}}} \left[\int_{Q_{\frac{(i-1)}{n}}}^{Q_{\frac{(i)}{n}}} n f_{Y|X}(y|t) f_{X|Y}(x|y)dy \right] n f_X(t)dt \\ &= \int_{Q_{\frac{(j-1)}{n}}}^{Q_{\frac{(j)}{n}}} \left[\int_{Q_{\frac{(i-1)}{n}}}^{Q_{\frac{(i)}{n}}} n f_{Y|X}(y|t) f_{X|Y}(x|y)dy \right] f_{X[i](j)}(t)dt. \end{aligned} \tag{4.11}$$

As argued by Gelfand and Smith (1990), Eq. (4.11) defines a fixed-point integral equation for which $f_{X[i](j)}(x)$ is the solution and the solution is unique.

We next show how SSGS can improve the efficiency of estimating the sample means of a probability density function $f(x)$.

Theorem 4.1 (Samawi et al. 2012). *Under the same conditions of the standard Gibbs sampling, the bivariate SSGS sample above* $[(X_{[i](j)}, Y_{(i)[j]}), i = 1, 2, \ldots, n$ *and* $j = 1, 2, \ldots, n]$ *from* $f(x, y)$ *provides the following:*

1. *Unbiased estimator of the marginal means of X and/or Y. Hence* $E(\bar{X}_{SSGS}) = \mu_x$, *where* $\mu_x = E(X)$.
2. $Var(\bar{X}_{SSGS}) \leq Var(\bar{X})$, *where* $\bar{X} = \frac{\sum_{i=1}^{n^2} X_i}{n^2}$.

Proof Using (2.5),

$$E(\bar{X}_{SSGS}) = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} E(X_{[i](j)}) = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} \int_{Q_{Y(i-1)/n}}^{Q_{Y(i)/n}} \int_{Q_{X(j-1)/n}}^{Q_{X(j)/n}} x\, f^{(\infty)}_{X_{[i](j)Y(i)[j]}}(x, y)dxdy$$

$$= \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} \int_{Q_{Y(i-1)/n}}^{Q_{Y(i)/n}} \int_{Q_{X(j-1)/n}}^{Q_{X(j)/n}} xn^2 f(x, y)dxdy$$

$$= \sum_{j=1}^{n} \int_{Q_{X(j-1)/n}}^{Q_{X(j)/n}} x\, f_x(x)dx \sum_{i=1}^{n} \int_{Q_{Y(i-1)/n}}^{Q_{Y(i)/n}} f_{Y|X}(y|x)dy = E(X) = \mu_x.$$

Similarly,

$$\mathrm{var}(\bar{X}_{SSGS}) = \frac{1}{n^4}\sum_{i=1}^{n}\sum_{j=1}^{n} Var(X_{[i](j)}) = \frac{1}{n^4}\sum_{i=1}^{n}\sum_{j=1}^{n} \int_{Q_{Y(i-1)/n}}^{Q_{Y(i)/n}} \int_{Q_{X(j-1)/n}}^{Q_{X(j)/n}} (x - \mu_{x_{[i](j)}})^2 f^{(\infty)} X_{[i](j)Y(i)[j]}(x, y)dxdy$$

$$= \frac{1}{n^4}\sum_{i=1}^{n}\sum_{j=1}^{n} \int_{Q_{Y(i-1)/n}}^{Q_{Y(i)/n}} \int_{Q_{X(j-1)/n}}^{Q_{X(j)/n}} (x - \mu_{x_{[i](j)}} \pm \mu_x)^2 n^2 f(x, y)dxdy$$

$$= \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} \int_{Q_{Y(i-1)/n}}^{Q_{Y(i)/n}} \int_{Q_{X(j-1)/n}}^{Q_{X(j)/n}} [(x - \mu_x) - (\mu_{[i](j)} - \mu_x]^2 f(x, y)dxdy$$

and,

$$\mathrm{var}(\bar{X}_{SSGS}) = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} \int_{Q_{Y(i-1)/n}}^{Q_{Y(i)/n}} \int_{Q_{X(j-1)/n}}^{Q_{X(j)/n}} \{(x - \mu_x)^2 - 2(x - \mu_x)(\mu_{x_{[i](j)}} - \mu_x)$$

$$+ (\mu_{x_{[i](j)}} - \mu_x)^2\} f(x, y)dxdy$$

$$= \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} \int_{Q_{Y(i-1)/n}}^{Q_{Y(i)/n}} \int_{Q_{X(j-1)/n}}^{Q_{X(j)/n}} (x - \mu_x)^2 f(x, y)dxdy - \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}$$

$$\int_{Q_{Y(i-1)/n}}^{Q_{Y(i)/n}} \int_{Q_{X(j-1)/n}}^{Q_{X(j)/n}} 2(x - \mu_x)(\mu_{x_{[i](j)}} - \mu_x) f(x, y)dxdy$$

$$+ \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} \int_{Q_{Y(i-1)/n}}^{Q_{Y(i)/n}} \int_{Q_{X(j-1)/n}}^{Q_{X(j)/n}} (\mu_{x_{[i](j)}} - \mu_x)^2 f(x, y)dxdy$$

$$= \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} \int_{Q_{Y(i-1)/n}}^{Q_{Y(i)/n}} \int_{Q_{X(j-1)/n}}^{Q_{X(j)/n}} (x - \mu_x)^2 f(x, y)dxdy - \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}$$

$$\int_{Q_{Y(i-1)/n}}^{Q_{Y(i)/n}} \int_{Q_{X(j-1)/n}}^{Q_{X(j)/n}} (\mu_{x_{[i](j)}} - \mu_x)^2 f(x,y)dxdy$$

$$= \frac{\sigma_X^2}{n^2} - \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} \int_{Q_{Y(i-1)/n}}^{Q_{Y(i)/n}} \int_{Q_{X(j-1)/n}}^{Q_{X(j)/n}} (\mu_{x_{[i](j)}} - \mu_x)^2 f(x,y)dxdy \leq V(\bar{X}) = \frac{\sigma_X^2}{n^2},$$

where $\sigma_X^2 = Var(X)$. Similar results can be obtained for the marginal mean of Y and the marginal distributions of X and Y. Note that as compared with using standard Gibbs sampling, using SSGS not only provides a gain in efficiency by but also reduces the sample size required to achieve a certain accuracy in estimating the marginal means and distributions. For very complex situations, the smaller required sample size can substantially reduce computation time. To provide insight to the gain in efficiency by using SSGS, we next conduct a simulation study.

4.3 Simulation Study and Illustrations

This section presents the results of a simulation study comparing the performance of the SSGS with the standard Gibbs sampling methods. To compare the performance of our proposed algorithm, we used the same illustrations as Casella and George (1992). For these examples, four bivariate samples of sizes, $n = 10, 20$, and 50 and Gibbs sequence length $k = 20, 50$ and 100 and $r = 20, 50$, and 100 in the long sequence Gibbs sampler. To estimate the variances of the estimators using the simulation method, we completed 5,000 replications. Using the 5,000 replications, we estimate the efficiency of our procedure relative to the traditional (i.e., standard) Gibbs sampling method by $\text{eff}(\hat{\theta}, \hat{\theta}_{SSGS}) = \frac{Var(\hat{\theta})}{Var(\hat{\theta}_{SSGS})}$, where θ is the parameter of interest.

Example 1 Casella and George (1992).
X and Y have the following joint distribution, $f(x, y) \propto \binom{m}{x} y^{x+\alpha-1}(1-y)^{m-x+\beta-1}$, $x = 0, 1, \ldots, m, 0 \leq y \leq 1$. Assume our purpose is to determine certain characteristics of the marginal distribution $f(x)$ of X. In Gibbs sampling method, we use the conditional distributions $f(x|y) \sim Binomial(m, y)$ and $f(y|x) \sim Beta(x + \alpha, m - x + \beta)$.

Tables 4 and 5 show that, relative to the standard Gibbs sampling method, SSGS improves the efficiency of estimating the marginal means. The amount of improvement depends on two factors: (1) which parameters we intend to estimate, and (2) the conditional distributions used in the process. Moreover, using the short or long Gibbs sampling sequence has only a slight effect on the relative efficiency.

Table 4 Standard Gibbs sampling method compared With the Steady-State Gibbs Sampling (SSGS) method (Beta-Binomial distribution)

n^2	k	Sample mean Gibbs sampling of X	Sample mean SSGS of X	Relative efficiency	Sample mean Gibbs sampling of Y	Sample mean SSGS of Y	Relative efficiency
$m = 5, \alpha = 2,$ and $\beta = 4$							
100	20	1.672	1.668	**3.443**	0.340	0.334	**3.787**
	50	1.667	1.666	**3.404**	0.333	0.333	**3.750**
	100	1.667	1.666	**3.328**	0.333	0.333	**3.679**
400	20	1.666	1.666	**3.642**	0.333	0.333	**3.861**
	50	1.669	1.667	**3.495**	0.333	0.333	**3.955**
	100	1.668	1.667	**3.605**	0.333	0.333	**4.002**
2500	20	1.666	1.666	**3.760**	0.333	0.333	**4.063**
	50	1.668	1.667	**3.786**	0.333	0.333	**3.991**
	100	1.667	1.667	**3.774**	0.333	0.333	**4.007**
$m = 16, \alpha = 2,$ and $\beta = 4$							
100	20	5.321	5.324	**1.776**	0.333	0.333	**1.766**
	50	5.334	5.334	**1.771**	0.333	0.333	**1.766**
	100	5.340	5.337	**1.771**	0.334	0.334	**1.769**
400	20	5.324	5.327	**1.805**	0.333	0.333	**1.811**
	50	5.333	5.333	**1.816**	0.333	0.333	**1.809**
	100	5.330	5.331	**1.803**	0.333	0.333	**1.806**
2500	20	5.322	5.325	**1.820**	0.333	0.333	**1.828**
	50	5.334	5.334	**1.812**	0.333	0.333	**1.827**
	100	5.334	5.333	**1.798**	0.333	0.333	**1.820**

Note The exact mean of x is equal to 5/3 and the exact mean of y is equal to 1/3 for the first case

Example 2 Casella and George (1992).
Let X and Y has the following conditional distributions that are exponential distributions, restricted to the interval (0, B), that is $f(x|y) \propto ye^{-yx}, 0 < x < B < \infty$ and $f(y|x) \propto xe^{-yx}, 0 < y < B < \infty$.

Similarly, Table 6 shows that SSGS improves the efficiency of marginal means estimation relative to standard Gibbs sampling. Again, using short or long Gibbs sampler sequence has only a slight effect on the relative efficiency.

Example 3 Casella and George (1992).
In this example, a generalization of the joint distribution is $f(x, y, m) \propto \binom{m}{x} y^{x+\alpha-1}$ $(1-y)^{nx+\beta-1}, e^{-\lambda}\frac{\lambda^m}{m!}, \lambda > 0, x = 0, 1, \ldots, m, 0 \leq y \leq 1, m = 1, 2, \ldots.$

Again, suppose we are interested in calculating some characteristics of the marginal distribution $f(x)$ of X. In Gibbs sampling method, we use the conditional distributions $f(x|y, m) \sim Binomial(m, y)$, $f(y|x, m) \sim Beta(x + \alpha, m - x + \beta)$ and

Table 5 Comparison of the Long Gibbs sampling method and the Steady-State Gibbs Sampling (SSGS) method (Beta-Binomial distribution)

n^2	r	Sample mean Gibbs sampling of X	Sample mean SSGS of X	Relative efficiency	Sample mean Gibbs sampling of Y	Sample mean SSGS of Y	Relative efficiency
$m = 5, \alpha = 2,$ and $\beta = 4$							
100	20	1.667	1.666	**3.404**	0.333	0.333	**3.655**
	50	1.665	1.665	**3.506**	0.333	0.333	**3.670**
	100	1.668	1.667	**3.432**	0.334	0.333	**3.705**
400	20	1.667	1.667	**3.623**	0.333	0.333	**4.014**
	50	1.666	1.666	**3.606**	0.333	0.333	**3.945**
	100	1.667	1.667	**3.677**	0.333	0.333	**3.997**
2500	20	1.667	1.666	**3.814**	0.333	0.333	**4.011**
	50	1.667	1.667	**3.760**	0.333	0.333	**4.125**
	100	1.667	1.667	**3.786**	0.333	0.333	**4.114**
$m = 16, \alpha = 2,$ and $\beta = 4$							
100	20	5.338	5.338	**1.770**	0.334	0.334	**1.785**
	50	5.335	5.334	**1.767**	0.334	0.334	**1.791**
	100	5.335	5.334	**1.744**	0.334	0.333	**1.763**
400	20	5.332	5.332	**1.788**	0.333	0.333	**1.820**
	50	5.337	5.336	**1.798**	0.333	0.333	**1.815**
	100	5.332	5.333	**1.821**	0.333	0.333	**1.820**
2500	20	5.332	5.332	**1.809**	0.333	0.333	**1.821**
	50	5.335	5.335	**1.832**	0.333	0.333	**1.827**
	100	5.333	5.333	**1.825**	0.333	0.333	**1.806**

Note The exact mean of x is equal to 5/3 and the exact mean of y is equal to 1/3

$f(m|x, y) \propto e^{-(1-y)\lambda} \frac{[(1-y)\lambda]^{m-x}}{(m-x)!}$, $m = x, x+1, \ldots$. For this example, we used the following parameters: $m = 5, \alpha = 2,$ and $\beta = 4$.

Similarly, Table 7 illustrates the improved efficiency of using SSGS for marginal means estimation, relative to standard Gibbs sampling. Again using a short or long Gibbs sampling sequence has only a slight effect on the relative efficiency. Note that this example is a three-dimensional problem, which shows the improved efficiency depends on the parameters under consideration.

We show that SSGS converges in the same manner as in the standard Gibbs sampling method. However, Sects. 3 and 4 indicate that SSGS is more efficient than standard Gibbs sampling for estimating the means of the marginal distributions using the same sample size. In the examples provided above, the SSGS efficiency (versus standard Gibbs) ranged from 1.77 to 6.6, depending on whether Gibbs sampling used the long or short sequence method and the type of conditional distributions used in the process. Using SSGS yielded a reduced sample size, and thus, reduces

Table 6 Relative efficiency of Gibbs sampling method and Steady-State Gibbs Sampling (SSGS) method (Exponential Distribution)

Standard Gibbs Algorithm $B = 5$							
n^2	k	Sample mean Gibbs sampling of X	Sample mean SSGS of X	Relative efficiency	Sample mean Gibbs sampling of Y	Sample mean SSGS of Y	Relative efficiency
100	20	1.265	1.264	**4.255**	1.264	1.263	**4.132**
	50	1.267	1.267	**4.200**	1.265	1.264	**4.203**
	100	1.267	1.267	**4.100**	1.263	1.265	**4.241**
400	20	1.263	1.264	**4.510**	1.266	1.265	**4.651**
	50	1.265	1.265	**4.341**	1.262	1.263	**4.504**
	100	1.263	1.264	**4.436**	1.265	1.265	**4.345**
2500	20	1.264	1.264	**4.461**	1.264	1.264	**4.639**
	50	1.264	1.264	**4.466**	1.265	1.265	**4.409**
	100	1.265	1.264	**4.524**	1.265	1.264	**4.525**
Long Gibbs Algorithm							
n^2	r	$B = 5$					
100	20	1.265	1.265	**4.305**	1.264	1.265	**4.349**
	50	1.267	1.264	**4.254**	1.261	1.264	**4.129**
	100	1.264	1.265	**4.340**	1.265	1.265	**4.272**
400	20	1.265	1.264	**4.342**	1.263	1.264	**4.543**
	50	1.265	1.265	**4.434**	1.265	1.265	**4.446**
	100	1.266	1.265	**4.387**	1.264	1.264	**4.375**
2500	20	1.264	1.264	**4.403**	1.264	1.264	**4.660**
	50	1.265	1.265	**4.665**	1.264	1.264	**4.414**
	100	1.265	1.265	**4.494**	1.264	1.264	**4.659**

computing time. For example, if the efficiency of using SSGS is 4, then the sample size needed for estimating the simulation's distribution mean, or other distribution characteristics, when using the ordinary Gibbs sampling method is 4 times greater than when using SSGS to achieve the same accuracy and convergence rate. Additionally, our SSGS sample produces unbiased estimators, as shown by theorem 4.1. Moreover, the bivariate steady-state simulation depends on n^2 simulated sample size to produce an unbiased estimate. However, in k dimensional problem, multivariate steady-state simulation depends on n^k simulated sample size to produce an unbiased estimate. Clearly, this sample size is not practical and will increase the simulated sample size required. To overcome this problem in high dimensional cases, we can use the independent simulation method described by Samawi (1999) that needs only a simulated sample of size n regardless of the number of dimensions. This approach slightly reduces the efficiency of using steady-state simulation. In conclusion, SSGS

Table 7 Gibbs Sampling Method and Steady-State Gibbs Sampling (SSGS) method (Beta-Binomial Distribution and Poisson Distribution)

Standard Gibbs Algorithm $\lambda = 5, \alpha = 2,$ and $\beta = 4$

n^3	k	Gibbs mean of X	SSGS mean X	Relative efficiency	Gibbs mean of Y	SSGS mean Y	Relative efficiency	Gibbs mean of Z	SSGS mean Z	Relative efficiency
1000	20	1.667	1.757	**1.883**	0.333	0.343	**6.468**	5.001	5.043	**5.333**
	50	1.665	1.756	**1.891**	0.333	0.343	**6.298**	4.999	5.043	**5.341**
	100	1.666	1.756	**1.881**	0.333	0.343	**6.352**	5.000	5.043	**5.331**
8000	20	1.667	1.758	**1.899**	0.333	0.343	**6.510**	5.000	5.042	**5.338**
	50	1.666	1.757	**1.897**	0.333	0.343	**6.477**	5.001	5.039	**5.321**
	100	1.666	1.755	**1.894**	0.333	0.343	**6.480**	5.000	5.039	**5.318**
27000	20	1.668	1.756	**1.909**	0.333	0.343	**6.591**	4.999	5.039	**5.340**
	50	1.667	1.755	**1.899**	0.333	0.343	**6.584**	5.001	5.039	**5.338**
	100	1.667	1.757	**1.892**	0.333	0.343	**6.583**	5.000	5.039	**5.339**

Long Gibbs Algorithm

n^3	r	$\lambda = 5, \alpha = 2,$ and $\beta = 4$								
1000	20	1.670	1.757	**1.883**	0.333	0.343	**6.468**	5.001	5.043	**5.333**
	50	1.665	1.756	**1.891**	0.333	0.343	**6.298**	4.999	5.043	**5.341**
	100	1.666	1.756	**1.881**	0.333	0.343	**6.352**	5.000	5.043	**5.331**
8000	20	1.665	1.758	**1.899**	0.333	0.343	**6.510**	5.000	5.042	**5.338**
	50	1.666	1.757	**1.897**	0.333	0.343	**6.477**	5.001	5.039	**5.321**
	100	1.667	1.755	**1.894**	0.333	0.343	**6.480**	5.000	5.039	**5.318**
27000	20	1.668	1.756	**1.909**	0.333	0.343	**6.591**	4.999	5.039	**5.340**
	50	1.667	1.755	**1.899**	0.333	0.343	**6.584**	5.001	5.039	**5.338**
	100	1.667	1.757	**1.892**	0.333	0.343	**6.583**	5.000	5.039	**5.339**

The exact mean of x is equal to 5/3 and the exact mean of y is equal to 1/3 performs at least as well as standard Gibbs sampling and SSGS offers greater accuracy. Thus, we recommend using SSGS whenever a Gibbs sampling procedure is needed. Further investigation is needed to explore additional applications and more options for using the SSGS approach.

References

Al-Saleh, M. F., & Al-Omari, A. I., (1999). Multistage ranked set sampling. *Journal of Statistical Planning and Inference*, *102*(2), 273–286.

Al-Saleh, M. F., & Samawi, H. M. (2000). On the efficiency of Monte-Carlo methods using steady state ranked simulated samples. *Communication in Statistics- Simulation and Computation*, *29*(3), 941–954. doi:10.1080/03610910008813647.

Al-Saleh, M. F., & Zheng, G. (2002). Estimation of multiple characteristics using ranked set sampling. *Australian & New Zealand Journal of Statistics*, *44*, 221–232. doi:10.1111/1467-842X.00224.

Al-Saleh, M. F., & Zheng, G. (2003). Controlled sampling using ranked set sampling. *Journal of Nonparametric Statistics*, *15*, 505–516. doi:10.1080/10485250310001604640.

Casella, G., & George, I. E. (1992). Explaining the Gibbs sampler. *American Statistician*, *46*(3), 167–174.

Chib, S., & Greenberg, E. (1994). Bayes inference for regression models with ARMA (p, q) errors. *Journal of Econometrics*, *64*, 183–206. doi:10.1016/0304-4076(94)90063-9.

David, H. A. (1981). *Order statistics*(2nd ed.). New York, NY: Wiley.

Evans, M., & Swartz, T. (1995). Methods for approximating integrals in statistics with special emphasis on Bayesian integration problems. *Statistical Science*, *10*(2), 254–272. doi:10.1214/ss/1177009938.

Gelfand, A. E., & Smith, A. F. M. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, *85*, 398–409. doi:10.1080/01621459.1990.10476213.

Gelman, A., & Rubin, D. (1991). *An overview and approach to inference from iterative simulation [Technical report]*. Berkeley, CA: University of California-Berkeley, Department of Statistics.

Geman, S., & Geman, D. (1984). Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *6*, 721–741. doi:10.1109/TPAMI.1984.4767596.

Hammersley, J. M., & Handscomb, D. C. (1964). Monte-Carlo methods. London, UK: Chapman & Hall. doi:10.1007/978-94-009-5819-7.

Hastings, W. K. (1970). Monte-Carlo sampling methods using Markov chains and their applications. *Biometrika*, *57*, 97–109. doi:10.1093/biomet/57.1.97.

Johnson, M. E. (1987). *Multivariate statistical simulation*. New York, NY: Wiley. doi:10.1002/9781118150740.

Johnson, N. L., & Kotz, S. (1972). *Distribution in statistics: Continuous multivariate distributions*. New York, NY: Wiley.

Liu, J. S. (2001). *Monte-Carlo strategies in scientific computing*. New York, NY: Springer.

McIntyre, G. A. (1952). A method of unbiased selective sampling, using ranked sets. *Australian Journal of Agricultural Research*, *3*, 385–390. doi:10.1071/AR9520385.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equations of state calculations by fast computing machine. *Journal of Chemical Physics*, *21*, 1087–1091. doi:10.1063/1.1699114.

Morgan, B. J. T. (1984). *Elements of simulation*. London, UK: Chapman & Hall. doi:10.1007/978-1-4899-3282-2.

Plackett, R. L. (1965). A class of bivariate distributions. *Journal of the American Statistical Association*, *60*, 516–522. doi:10.1080/01621459.1965.10480807.

Robert, C., & Casella, G. (2004). *Monte-Carlo statistical methods* (2nd ed.). New York, NY: Springer. doi:10.1007/978-1-4757-4145-2.

Roberts, G. O. (1995). Markov chain concepts related to sampling algorithms. In W. R. Gilks, S. Richardson, & D. J. Spiegelhalter (Eds.), *Markov Chain Monte-Carlo in practice* (pp. 45–57). London, UK: Chapman & Hall.

Samawi, H. M. (1999). More efficient Monte-Carlo methods obtained by using ranked set simulated samples. *Communication in Statistics-Simulation and Computation*, *28*, 699–713. doi:10.1080/03610919908813573.

Samawi H. M., & Al-Saleh, M. F. (2007). On the approximation of multiple integrals using multivariate ranked simulated sampling. *Applied Mathematics and Computation*, *188*, 345–352. doi:10.1016/j.amc.2006.09.121.

Samawi, H. M., Dunbar, M., & Chen, D. (2012). Steady state ranked Gibbs sampler. *Journal of Statistical Simulation and Computation*, *82*, 1223–1238. doi:10.1080/00949655.2011.575378.

Samawi, H. M., & Vogel, R. (2013). More efficient approximation of multiple integration using steady state ranked simulated sampling. *Communications in Statistics-Simulation and Computation*, *42*, 370–381. doi:10.1080/03610918.2011.636856.

Shreider, Y. A. (1966). *The Monte-Carlo method*. Oxford, UK: Pergamom Press.

Stokes, S. L. (1977). Ranked set sampling with concomitant variables. *Communications in Statistics—Theory and Methods*, *6*, 1207–1211. doi:10.1080/03610927708827563.

Tanner, M. A. (1993). *Tools for statistical inference* (2nd ed.). New York, NY: Wiley. doi:10.1007/978-1-4684-0192-9.

Tanner, M. A., & Wong, W. (1987). The calculation of posterior distribution by data augmentation (with discussion). *Journal of the American Statistical Association*, *82*, 528–550. http://dx.doi.org/10.1080/01621459.1987.10478458.

Tierney, L. (1995). Introduction to general state-space Markov chain theory. In W. R. Gilks, S. Richardson, & D. J. Spiegelhalter (Eds.), *Markov Chain Monte-Carlo in practice* (pp. 59–74). London, UK: Chapman & Hall.

Springer

http://www.springer.com/978-981-10-3306-3

Monte-Carlo Simulation-Based Statistical Modeling
Chen, D.-G.; Chen, J.D. (Eds.)
2017, XX, 430 p. 64 illus., 33 illus. in color., Hardcover
ISBN: 978-981-10-3306-3