Available online at www.sciencedirect.com
ScienceDirect
Journal of Multivariate Analysis 98 (2007) 350–369
Journal of Multivariate Analysis
www.elsevier.com/locate/jmva

Densities for random balanced sampling

Peter Bubenik[a,*,1], John Holbrook[b,2]

[a] *Department of Mathematics, Cleveland State University, 2121 Euclid Ave. RT 1515, Cleveland OH 44115, USA*
[b] *Department of Mathematics and Statistics, University of Guelph, Guelph, Ont., Canada N1G2W1*

Received 23 March 2005
Available online 10 March 2006

Abstract

A random balanced sample (RBS) is a multivariate distribution with n components X_k, each uniformly distributed on $[-1, 1]$, such that the sum of these components is precisely 0. The corresponding vectors $\vec{X}$ lie in an $(n-1)$-dimensional polytope $M(n)$. We present new methods for the construction of such RBS via densities over $M(n)$ and these apply for arbitrary n. While simple densities had been known previously for small values of n (namely 2,3, and 4), for larger n the known distributions with large support were fractal distributions (with fractal dimension asymptotic to n as $n \to \infty$). Applications of RBS distributions include sampling with antithetic coupling to reduce variance, and the isolation of nonlinearities. We also show that the previously known densities (for $n \leqslant 4$) are in fact the only solutions in a natural and very large class of potential RBS densities. This finding clarifies the need for new methods, such as those presented here.

AMS 2000 subject classification: 62M05 (62H99)

Keywords: Multivariate distributions; Sampling methodology; Antithetic variates; Densities on polytopes; Fractal geometry; Monte Carlo

1. Introduction

The aim of this article is to improve our understanding of "random balanced samples" (RBS), multivariate distributions having specified marginals as well as a "balanced" property; the precise

* Corresponding author. Fax: +1 216 523 7340.
E-mail addresses: p.bubenik@csuohio.edu (P. Bubenik), jholbroo@uoguelph.ca (J. Holbrook).
URL: http://academic.csuohio.edu/bubenik_p/ (P. Bubenik).

1 The first author thanks NSERC of Canada for support under its Summer Undergraduate Scholarship program; his work during summer 1995 was the basis for some of the results reported here.

2 The second author thanks NSERC for support through its Discovery Grants program.

0047-259X/$ - see front matter
doi:10.1016/j.jmva.2006.01.007

definition is given below. RBS have found several applications, described in the literature, but the construction of nontrivial RBS distributions has presented interesting mathematical challenges. We describe new methods that generate RBS *densities* with respect to the underlying Lebesgue measure. In contrast, previous work on this problem has relied on fractal geometry and iterated function systems. Ironically, the earliest examples of RBS distributions were given by simple explicit densities. In Section 3, however, we show that those earlier techniques *cannot* be extended to higher dimensions.

Definition 1. A RBS of size n is a system of random variables

$$\vec{X} = (X_1, X_2, \dots, X_n)$$

such that $\vec{X}$ is *balanced*, i.e.

$$\sum_{k=1}^{n} X_k = 0,$$

and each X_k is uniformly distributed over $[-1, 1]$. Other marginal distributions might be considered, but we focus here on the uniform case.

Our main concern here is to present new methods for constructing RBS densities (see Section 2), and to point out the severe limitations of older methods (see Section 3). Nevertheless, to set the context we briefly review some situations where RBS distributions may be usefully applied.

A classical technique for reducing Monte Carlo variance is the use of "antithetic variates". First introduced by Hammersley and Morton in 1956 [11], it has been extended to larger groupings of variables by Arvidsen and Johnsson [1], and applied to the bootstrap by Hall [10]. For more recent applications, see the papers of Craiu and Meng [2–4]. In the following example, we motivate the use of antithetic variates, and we show that in the exchangeable case, RBS distributions are *extremely antithetic*. That is, the variables are as negatively correlated as possible.

Example 2. We may wish to estimate the mean $\frac{1}{2}\int_{-1}^{1} f(x)\,\mathrm{d}x$ of an unknown function $f{:}[-1, 1] \to \mathbb{R}$ by means of

$$\overline{f}_n = \frac{1}{n}\sum_{k=1}^{n} f(X_k).$$

The variance of $\overline{f}_n$ will be reduced if $\mathrm{cov}(f(X_k), f(X_j)) < 0$ for $k \neq j$. This reduction can be achieved for a variety of functions f by means of *antithetic coupling* of the sampling variables X_k, i.e. by insisting that $\mathrm{cov}(X_k, X_j)$ be negative for $k \neq j$. Note that for X_k uniformly distributed over $[-1, 1]$ we have

$$E\left(\left|\sum_{k=1}^{n} X_k\right|^2\right) = \frac{n}{3} + \sum_{k \neq j} E(X_k X_j),$$

so that if $\mathrm{cov}(X_k, X_j) = \alpha$ for all $k \neq j$ (e.g. in the exchangeable case) then

$$\alpha = \frac{1}{n(n-1)}\left(E\left(\left|\sum_{k=1}^{n} X_k\right|^2\right) - \frac{n}{3}\right).$$

Thus, the minimum value of α is $-\frac{1}{3(n-1)}$, achieved precisely when we have a RBS.

Example 3. In the same setting as Example 2, the estimate $\overline{f}_n$ via a RBS allows a cleaner distinction between the linear and nonlinear parts of *f*. Indeed, if $f(x) = Cx + g(x)$, where g is nonlinear then a RBS entirely eliminates the linear component *Cx* (even though the constant *C* may not be known) from the estimate $\overline{f}_n$. Applications of this nature are discussed in Gerow et al. [5,6,8,9]; see also [14].

The structure of RBS distributions is by no means a simple matter; there is a bewildering variety of RBS distributions for sample size $n \geqslant 3$. Each may be regarded as a probability distribution on the regular polytope

$$M(n) = \left\{\vec{x} \in [-1, 1]^n : \sum_{k=1}^{n} x_k = 0\right\},$$

with the additional requirement of marginal uniformity. For small values of n, the polytopes $M(n)$ are familiar geometric objects; for example, one easily sees that $M(3)$ is the regular hexagon with vertices $(\pm 1, \mp 1, 0)$, $(\pm 1, 0, \mp 1)$, and $(0, \pm 1, \mp 1)$. It seems that the earliest solution to this puzzle (aside from the trivial case where $n = 2$) is due to D. Robson. He noted that a simple piecewise linear density on $M(3)$ meets the requirements. See Fig. 1. K Gerow, in [5], extended Robson's method to $n = 4$, and detected problems for this method in case $n > 4$. As it turned out, even a broad generalization of the Gerow–Robson method could not produce RBS densities for $n \geqslant 5$. This phenomenon is studied in detail in Section 3.

As alternatives to the Robson density on $M(3)$, many other RBS distributions were constructed; part of [8] is devoted to a survey of these, many of which have attractive geometric features. In particular, certain *fractal* measures define RBS distributions for arbitrary sample size n. See Figs. 2 and 3. While such fractal measures cannot have all of $M(n)$ as their support, the fractal dimension of their supports can be made large with respect to the topological dimension $n - 1$ of $M(n)$ (see [8,13]). For some time, there appeared to be a dichotomy in the possible RBS constructions: for small n we had simple RBS densities, but for larger n only *singular* RBS distributions (the fractal constructions, which produce measures singular with respect to the natural $(n-1)$-volume on $M(n)$).

More recently, however, we have found constructions that yield RBS densities for any sample size n. These are described in Section 2. They introduce a technique of *redistribution* that may have other applications as well. Like the earlier fractal constructions (which can be implemented via iterated function systems), these new methods are algorithmically effective. That is, workers in the field can easily generate sampling vectors $\vec{X}$ by following the procedures outlined in Section 2.

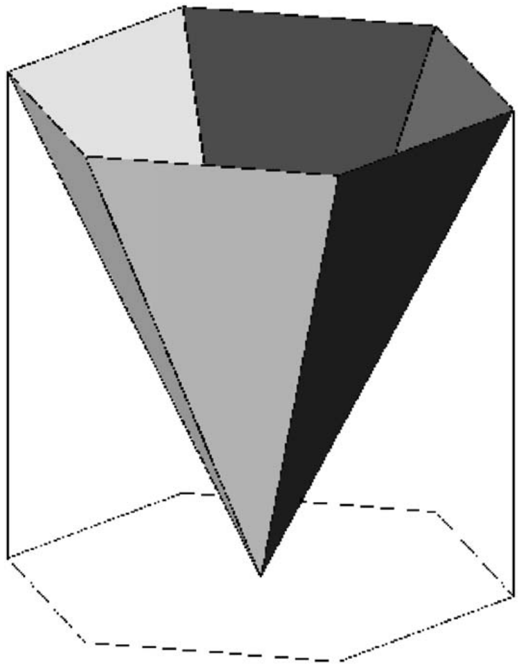

Fig. 1. The Robson density graphed as a piecewise linear density over $M(3)$, represented as a regular hexagon. Gerow discovered that a similar construction (density proportional to max $|X_k|$) works on $M(4)$.

2. Optimal densities

2.1. Goals

Given a RBS, it may be optimized in several directions, and these goals are to some degree incompatible. We may wish, for example, to achieve as much mutual independence as possible among the coordinate random variables X_k. In this way, we would stay near the more familiar situation of i.i.d. sampling. It is clear, however, that for a given value of n not more than $n/2$ of the X_k can be mutually independent. Suppose, for example, that $X_1, X_2, \ldots, X_m$ are mutually independent; then there is a positive probability (namely $(\varepsilon/2)^m$, if $\varepsilon > 0$ is small) that

$$X_k \in [1-\varepsilon, 1] \quad (k = 1, \ldots, m). \tag{1}$$

In this case, we would have

$$m(1-\varepsilon) \leqslant \sum_{k=1}^{m} X_k = -\sum_{k=m+1}^{n} X_k \leqslant (n-m), \tag{2}$$

which is not possible (for sufficiently small ε) if $2m > n$. On the other hand, if we compromise in other directions, the maximal mutual independence is easy to obtain. Consider, for example, an even sample size $n = 2m$ and a RBS defined by choosing $X_1, X_2, \ldots, X_m$ independently (and

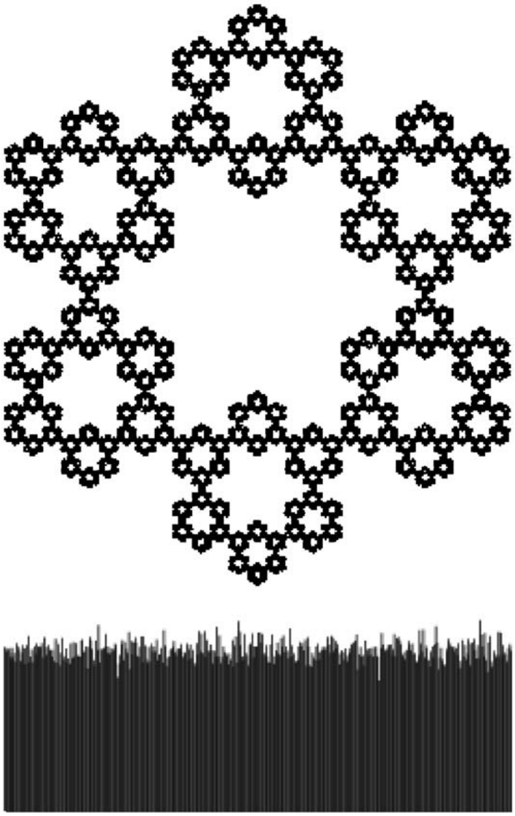

Fig. 2. A fractal "superstar" RBS distribution on the hexagon $M(3)$. The histogram generated by this simulation is shown at the bottom, visually verifying the uniform distribution of X_k.

each X_k uniform in $[-1, 1]$, of course), then setting

$$X_{m+k} = -X_k \quad (k = 1, \ldots, m). \tag{3}$$

Here, we have the coordinates partitioned into two subsets of $m = n/2$ mutually independent random variables (for example, the first m and the last m of the X_k). This RBS, however, is *degenerate* in the sense that the distribution of $\vec{X}$ is supported on a $(n/2)$-dimensional subset of the natural range $M(n)$ for RBS of size n. Recall that

$$M(n) = \left\{ \vec{x} \in [-1, 1]^n : \sum_{k=1}^{n} x_k = 0 \right\}, \tag{4}$$

so that $M(n)$ has the much larger dimension $n - 1$. A sampling procedure based on such a RBS lacks robustness in a certain sense: it may overlook significant structure in the distribution under investigation because that structure lies outside the support of the RBS. Following this line of thought, the dimension of the support of a RBS distribution has been viewed as a measure of *robustness*; in some interesting cases this dimension must be computed as a *fractal* dimension (see [8], for example; a fractal construction is also displayed in [3]). In Section 2.2, we shall see

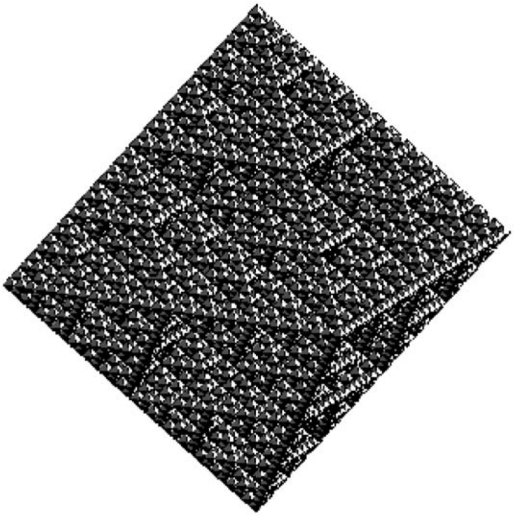

Fig. 3. As in Fig. 2, a fractal RBS may be generated in the octahedron $M(4)$ by means of an IFS (iterated function system) mapping all of $M(n)$ closer to its various vertices (six in this case).

how the degenerate RBS defined by (3) can be modified to achieve maximal robustness $(n-1)$ without much loss in mutual independence of the coordinates.

A natural $(n-1)$-dimensional model for $M(n)$ is obtained by first constructing n unit vectors $u_1, u_2, \ldots, u_n$ in $\mathbb{R}^{n-1}$ such that $(u_i, u_j) = -1/(n-1)$ whenever $i \neq j$. Then $M(n)$ may be identified with $\{v \in \mathbb{R}^{n-1} : (\forall k)\ -1 \leqslant (v, u_k) \leqslant 1\}$. Corresponding to each such v we have $\vec{X}$ with $X_k = (v, u_k)$ $(k = 1, 2, \ldots, n)$. See [8] for further details. These models of $M(n)$ yield regular polytopes in $\mathbb{R}^{n-1}$. For example, $M(3)$ is seen as a regular hexagon, as in Figs. 1 and 2, while $M(4)$ is the regular octahedron outlined in Fig. 3. Fig. 4 presents one of the three-dimensional faces of $M(5)$, which is itself a four-dimensional object.

Along with independence and robustness, a third desirable feature of a RBS procedure might be *algorithmic efficiency*, the efficiency with which sampling variables may be generated for use in experiments. The RBS defined by (3), for example, is very efficient, since it involves very little beyond m calls to a random number generator to produce a sampling vector $\vec{X}$. We shall see that the modifications proposed in Section 2.2 retain most of this efficiency as they increase robustness.

2.2. Redistribution

Here, we illustrate, in a simple setting, a general procedure of *redistribution* that may be used to increase the robustness of a RBS. In this initial example, we confine ourselves to the redistribution of *pairs* of independent coordinates chosen from $\vec{X}$.

Lemma 4. *Given independent* X_1, X_2, *each uniformly distributed on* $[-1, 1]$, *let* $S = X_1 + X_2$ *and define the new variables*

$$Y_1 = \frac{S}{2} + \left(1 - \frac{|S|}{2}\right)T, \quad Y_2 = \frac{S}{2} - \left(1 - \frac{|S|}{2}\right)T, \tag{5}$$

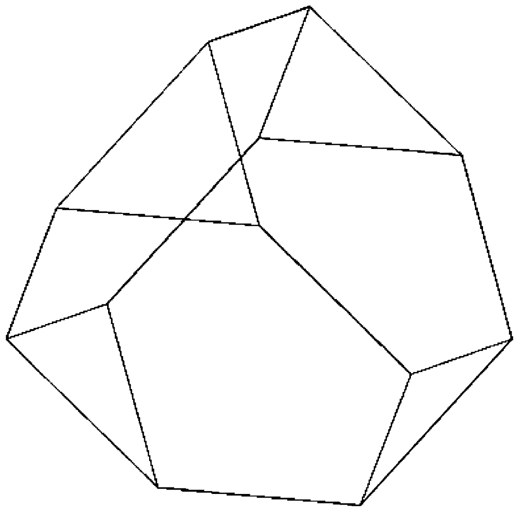

Fig. 4. While the four-dimensional $M(5)$ cannot be easily visualized, each of its faces is similar to the elegant three-dimensional object seen here. It may be regarded as a truncated tetrahedron; the result has four regular hexagons and four equilateral triangles as *its* faces.

where T is uniform on $[-1, 1]$ *and independent of the* X_k; *then* Y_1, Y_2 *are also independent and uniform on* $[-1, 1]$.

Remark 5. Since $Y_1 + Y_2 = S = X_1 + X_2$, we may think of the procedure in this lemma as *redistributing* the part of a RBS captured by $X_1 + X_2$. The geometry behind this construction is revealed by writing $\vec{Y} = (Y_1, Y_2) = \vec{D} + T\vec{Q}$, where $\vec{D} = (S/2, S/2)$ is a point on the diagonal of $[-1, 1]^2$ and

$$\vec{Q} = \left(1 - \frac{|S|}{2}, -\left(1 - \frac{|S|}{2}\right)\right),$$

so that $\vec{A} = \vec{D} + \vec{Q}$ and $\vec{B} = \vec{D} - \vec{Q}$ are the endpoints of the segment in $[-1, 1]^2$ passing through $\vec{D}$ and perpendicular to the diagonal (see Fig. 5). Thus, $\vec{Y}$ is chosen in two steps: first, $\vec{D}$ is chosen with a density proportional to the length of $\vec{A}\vec{B}$, then $\vec{Y}$ is placed at a point chosen uniformly along $\vec{A}\vec{B}$. This procedure strongly suggests that $\vec{Y}$ will be uniformly distributed over the square $[-1, 1]^2$. We present a more formal proof below, computing the joint density of (Y_1, Y_2).

Proof. Let $D = S/2$. The joint density $f(t, d)$ of (T, D) on $[-1, 1]^2$ is given by $\frac{1}{2}(1 - |d|)$, since T and D are independent and it is easy to compute the density of D on $[-1, 1]$ as $1 - |d|$. Also

$$(Y_1, Y_2) = (D + (1 - |D|)T, \quad D - (1 - |D|)T),$$

a transformation with Jacobian matrix

$$J(t, d) = \begin{bmatrix} (1 - |d|) & 1 - \operatorname{sign}(d)t \\ -(1 - |d|) & 1 + \operatorname{sign}(d)t \end{bmatrix}$$

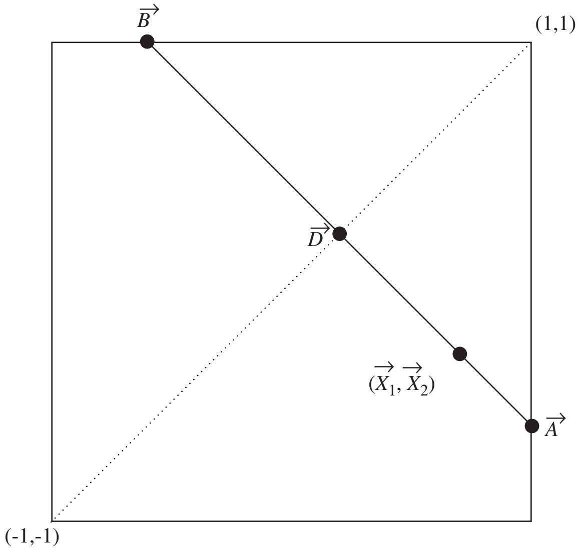

Fig. 5. A graphical view of the redistribution algorithm for the case $\vec{X}_1 = 0.7$, $\vec{X}_2 = -0.3$.

(compute for the two cases $d \geqslant 0$ and $d < 0$). It follows that $\det J(t, d) = 2(1 - |d|)$. The transformation is injective from $[-1, 1]^2$ to $[-1, 1]^2$, so that we may compute (compare the discussion of "change of variable" in [12, p. 213]) the joint density of (Y_1, Y_2) at the image of (T, D) as $f(t, d)/\det J(t, d) = \frac{1}{4}$, the uniform density on $[-1, 1]^2$. Thus, Y_1, Y_2 are uniform over $[-1, 1]$ and are independent. $\square$

The following theorem shows how to redistribute selected pairs of coordinates from the degenerate RBS defined by (3) so that we obtain a RBS that is maximally robust.

Theorem 6. *Let $X_1, \ldots, X_m, T_1, \ldots, T_m$ be mutually independent random variables, each uniform on $[-1, 1]$. Assume $m > 1$ and let*

$$S_k = X_k - X_{k+1} \quad (k = 1, \ldots, m), \tag{6}$$

with the understanding that $X_{m+1} = X_1$. Next (redistributing the coordinates combined in S_k as in the lemma), let

$$Y_{2k-1} = \frac{S_k}{2} + \left(1 - \frac{|S_k|}{2}\right) T_k, \quad Y_{2k} = \frac{S_k}{2} - \left(1 - \frac{|S_k|}{2}\right) T_k, \tag{7}$$

for $k = 1, \ldots, m$. Then $\vec{Y}$ defines a RBS of size $n = 2m$ that is maximally robust, i.e. that has a set of dimension $n - 1$ as its support. Moreover, the distribution of $\vec{Y}$ is given by a density on $M(n)$.

Proof. Since X_k and $-X_{k+1}$ are independent and uniform on $[-1, 1]$, Lemma 4 (with appropriate change of notation) ensures that each coordinate [3] of $\vec{Y}$ is uniform on $[-1, 1]$. Clearly

[3] We remark that it may be of interest to assess the degree to which these coordinates are mutually independent.

$\sum_1^n Y_k = \sum_1^m S_k = 0$, so that $\vec{Y}$ indeed defines a RBS. To see that it is maximally robust, note first that every $\vec{y} \in M(n)$ with sufficiently small $|y_k|$ can occur as a value of $\vec{Y}$. Indeed, if $|y_k| < \frac{1}{n}$ for each k, we set (for $k = 1, 2, \ldots, m$)

$$x_k = -\sum_{j=1}^{k-1} r_j, \tag{8}$$

where $r_j = y_{2j-1} + y_{2j}$. Note that $x_1 = x_{m+1} = 0$ and that $|x_k| \leqslant \frac{2(m-1)}{n} < 1$. We have $x_k - x_{k+1} = r_k$ for each k; for $k = m$ this is a consequence of the fact that $\sum_1^m r_k = \sum_1^n y_k = 0$. Solving for appropriate values of t_k, we find that we require only that

$$t_k = \frac{y_{2k-1} - y_{2k}}{2 - |y_{2k-1} + y_{2k}|}.$$

It is easy to check that $|t_k| \leqslant 1$, simply because $|y_{2k-1}|, |y_{2k}| \leqslant 1$; if $y_{2k-1} = y_{2k} = \pm 1$, any value of t_k will do. Thus, with $\vec{X} = \vec{x}$ and $\vec{T} = \vec{t}$, we have $\vec{Y} = \vec{y}$, as claimed.

Next, consider a neighborhood $\vec{y}(\varepsilon)$ of such a $\vec{y}$ in $M(n)$; we must show that the procedure of the theorem places sampling vectors in $\vec{y}(\varepsilon)$ with positive probability. By the continuity of the procedure for obtaining $\vec{Y}$ from $(\vec{X}, \vec{T})$, there is a neighborhood $\vec{u}(\delta)$ of $(\vec{x}, \vec{t})$ in $[-1, 1]^n$ that is mapped by the procedure into $\vec{y}(\varepsilon)$. The probability of $\vec{u}(\delta)$ is already positive; indeed it is just the normalized n-volume of $\vec{u}(\delta)$, since the coordinates of $(\vec{X}, \vec{T})$ are chosen independently and uniformly in $[-1, 1]$.

To see that the distribution of $\vec{Y}$ is given by a density on $M(n)$, we introduce the (linear) mapping $L : [-1, 1]^m \to 2M(m)$ defined by

$$L(x_1, \ldots, x_m) = (x_1 - x_2, x_2 - x_3, \ldots, x_m - x_1). \tag{9}$$

The analysis above shows that $\vec{y} \in M(n)$ occurs as a value of $\vec{Y}$ exactly when $L(\vec{x}) = \vec{r}$ for some $\vec{x} \in [-1, 1]^m$. Given $\varepsilon > 0$, consider the map

$$f : (1 - \varepsilon)M(n) \to 2M(m) \times [-1, 1]^m$$

defined by $f(\vec{y}) = (f_1(\vec{y}), f_2(\vec{y}))$, where $f_1(\vec{y}) = \vec{r}$ and $f_2(\vec{y}) = \vec{t}$. The denominators in the expressions for the t_k are bounded away from zero since $2 - |y_{2k-1} + y_{2k}| \geqslant 2 - 2(1 - \varepsilon) = 2\varepsilon$, so that f is Lipschitz on $(1 - \varepsilon)M(n)$ for each fixed ε. Given a Borel subset B, of $(1 - \varepsilon)M(n)$, the probability assigned to B by the distribution of $\vec{Y}$ is the (normalized) n-volume (or Lebesgue measure) of $\{(\vec{x}, \vec{t}) : (L(\vec{x}), \vec{t}) \in f(B)\}$. Since L is linear and $[-1, 1]^n$ is bounded, this is at most a constant times the $(n - 1)$-volume of $f(B)$. As f is Lipschitz, this is in turn bounded by a constant times the $(n - 1)$-volume of B. Certainly, then, the distribution is absolutely continuous with respect to normalized Lebesgue measure on $(1-\varepsilon)M(n)$. Considering a sequence of ε-values tending to 0, we see that the distribution is absolutely continuous on $M(n)$ itself and so given by a density with respect to the normalized $(n - 1)$-volume on $M(n)$. □

2.3. *Using all of $M(n)$*

The procedure of the last section retains algorithmic efficiency and it is robust in the sense of dimension, but in most cases the sampling values do not fill all of $M(n)$. Here, we show how to modify the construction to obtain RBS procedures that have *all* of $M(n)$ as support. First, let us clarify the reasons for the failure of the construction in Section 2.2 to cover $M(n)$; we reuse

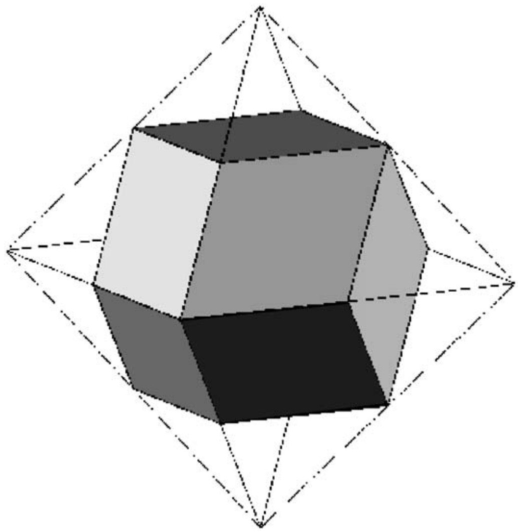

Fig. 6. The fact that the rhombic dodecahedron does not fill up all of the octahedron $2M(4)$ reveals the need for symmetrization in order to generate a RBS supported on *all* of $M(8)$.

the notations of the proof above. As $\vec{y}$ ranges over $M(n)$ the corresponding $\vec{r}$ fills all of $2M(m)$ (recall that $n = 2m$). The procedure will yield every $\vec{y} \in M(n)$ as a possible value exactly when $2M(m) = L([-1, 1]^m)$. Since $L([-1, 1]^m)$ is the convex hull of the images of the 2^m extreme points of $[-1, 1]^m$, we can simply check whether these images include the extreme points of $2M(m)$, which are easy to identify (at most one coordinate can differ from ± 2).

For $m = 2$, for example, the extreme points of $2M(2)$ are

$$(2, -2) = L(1, -1) \quad \text{and} \quad (-2, 2) = L(-1, 1), \tag{10}$$

so that the procedure of Section 2.2 for $m = 2$ gives a RBS distribution on $M(4)$ having a density that is positive everywhere on the octahedron. One can check that this density is unbounded, in contrast to the piecewise linear density found by Gerow (see Section 3,). For $m = 3$, we again get a RBS distribution with all of $M(6)$ as support, because the extreme points of $2M(3)$ are

$$\pm (2, -2, 0) = L(\pm(1, -1, 1)), \pm(0, 2, -2) = L(\pm(1, 1, -1)), \text{ etc.} \tag{11}$$

For $m = 4$, however, one can compute in a similar fashion that $L([-1, 1]^4)$ is a sort of rhombic dodecahedron lying *strictly* inside the octahedron $2M(4)$. See Fig. 6. Thus, except in a few simple cases, the procedure of Section 2.2, while it is maximally robust in terms of dimension, does not fill $M(n)$. As we will now see, the remedy is simply to symmetrize this procedure.

Lemma 7. *Given any $\vec{w} \in M(n)$, there is a permutation σ of $\{1, 2, \ldots, n\}$ such that for $\vec{z} = \sigma(\vec{w}) := (w_{\sigma(1)}, \ldots, w_{\sigma(n)})$ all partial sums*

$$\sum_{j=1}^{k} z_j \quad (k = 1, \ldots, n) \tag{12}$$

lie in $[-1, 1]$.

Proof. We can choose the values of $\sigma(i)$ inductively. Suppose that (distinct) $\sigma(1), \ldots, \sigma(i)$ have been chosen so that the partial sums (12) lie in $[-1, 1]$ for all $k \leqslant i$. Since

$$\sum_{j=1}^{i} w_{\sigma(j)} + \sum_{j \notin \sigma(\{1,\ldots,i\})} w_j = 0, \tag{13}$$

we can choose $\sigma(i+1)$ such that $w_{\sigma(i+1)}$ has sign opposite from that of the first sum in (13); it follows that the partial sum (12) lies in $[-1, 1]$ for $k = i + 1$ as well. □

Theorem 8. *Given $n = 2m$, let the RBS $\vec{Y}$ be defined as in Theorem* 6, *and let σ be a randomly chosen permutation of* $\{1, \ldots, n\}$. *Then $\vec{W} = \sigma(\vec{Y})$ defines a RBS with all of $M(n)$ as support.*

Proof. For any $\vec{w} \in M(n)$ using the σ from Lemma 7, $\vec{w} = \sigma^{-1}(\vec{z})$ where all the partial sums of (12) lie in $[-1, 1]$. The proof of Theorem 6 showed that any such $\vec{z}$ could occur as a RBS $\vec{Y}$ defined in Section 2.2. Thus, $\vec{w}$ can occur as $\tau(\vec{Y})$ for the permutation $\tau = \sigma^{-1}$. □

2.4. Odd sample sizes

For odd sample sizes, $n = 2m + 1$ we can likewise construct an algorithmically efficient degenerate RBS with two disjoint maximal subsets of mutually independent variables. This can be done by choosing $X_1, \ldots, X_m$ (with each X_k uniform in $[-1, 1]$) and $B \in \{-1, 1\}$ independently, and then setting

$$X_{m+k} = -X_k \quad (k = 1, \ldots, m-1),$$

$$X_{2m} = -\tfrac{1}{2}(X_m + B),$$

$$X_{2m+1} = -\tfrac{1}{2}(X_m - B).$$

It is easy to check that this gives a RBS. As in the even case, we can make this distribution robust by redistributing pairs of the variables.

Theorem 9. *Let $X_1, \ldots, X_m, T_1, \ldots, T_m$ be mutually independent random variables, each uniform on $[-1, 1]$ and B be a discrete random variable, uniform on $\{-1, 1\}$. Assume $m > 1$ and let*

$$S_k = X_k - X_{k+1} \quad (k = 1, \ldots m-2),$$

$$S_{m-1} = X_{m-1} - \tfrac{1}{2}(X_m + B),$$

$$S_m = -\tfrac{1}{2}(X_m - B) - X_1.$$

Next, redistribute the variables as in Lemma 4:

$$Y_{2k-1} = \frac{S_k}{2} + \left(1 - \frac{|S_k|}{2}\right) T_k \quad (k = 1, \ldots, m),$$

$$Y_{2k} = \frac{S_k}{2} - \left(1 - \frac{|S_k|}{2}\right) T_k \quad (k = 1, \ldots, m),$$

$$Y_{2m+1} = X_m.$$

Then $\vec{Y}$ is an RBS of size $n = 2m + 1$ that is maximally robust. That is, it has a set of dimension $n - 1$ as its support.

Proof. First, note that $-\frac{1}{2}(X_m + B)$ and $-\frac{1}{2}(X_m - B)$ are uniformly distributed on $[-1, 1]$. As a result, for each k, S_k is the sum of two independent uniform variables, and by Lemma 4, each coordinate of $\vec{Y}$ is uniform on $[-1, 1]$. Also, $\sum_{k=1}^{n} Y_k = \sum_{k=1}^{m} S_k + Y_{2m+1} = -\frac{1}{2}(X_m + B) - \frac{1}{2}(X_m - B) + X_m = 0$, so $\vec{Y}$ does indeed define an RBS.

To see that this RBS is maximally robust, we will show that every $\vec{y} \in M(n)$ with sufficiently small coordinates can occur as $\vec{Y}$. If $|y_k| < \frac{1}{2m}$ for each k, construct $\vec{x}$ and $\vec{t}$ as follows. Fix $b \in \{-1, 1\}$ and let $r_k = y_{2k-1} + y_{2k}$ for $k = 1, \dots, m$, and let

$$x_m = -\sum_{i=1}^{m} r_i,$$

$$x_k = \frac{1}{2}(x_m + b) + \sum_{i=k}^{m-1} r_i \quad (k = 1, \dots, m - 1).$$

Then for $k < m$, $x_k = \frac{1}{2}(\sum_{i=1}^{m}(-1)^{a_i} r_i + b)$ for some choice of a_i. So since $|r_k| < \frac{1}{m}$, $|x_k| < 1$ for all k. Also for $k \leqslant m - 2$,

$$s_k = x_k - x_{k+1} = r_k, \quad s_{m-1} = x_{m-1} - \tfrac{1}{2}(x_m + b) = r_{m-1}$$

and

$$s_m = -\frac{1}{2}(x_m - b) - x_1 = -x_m - \sum_{i=1}^{m-1} r_i = r_m.$$

Thus, we have a $\vec{x} \in [-1, 1]^m$ such that the corresponding $s_k = r_k$ for each k. We can then solve for $t_k \in [-1, 1]$ such that $Y_{2k-1} = y_{2k-1}$ and $Y_{2k} = y_{2k}$. Also, $Y_{2m+1} = x_m = -\sum_{i=1}^{m} r_i = -\sum_{k=1}^{2m} y_k = y_{2m+1}$.

Finally, consider the neighborhood $y(\varepsilon)$ of $\vec{y}$ in $M(n)$; we must show that the procedure of the theorem places sampling vectors in $y(\varepsilon)$ with positive probability. By the continuity of the procedure there is a neighborhood $\vec{u}(\delta)$ of $(\vec{x}, \vec{t})$ in $[-1, 1]^n$ that is mapped by the procedure into $y(\varepsilon)$. The probability of $\vec{u}(\delta)$ is already positive; indeed it is just the normalized n-volume, since the coordinates are chosen independently and uniformly in $[-1, 1]$. $\square$

Note that the procedure above will yield a $\vec{y} \in M(n)$ if and only if $x_k = \sum_{i=2k-1}^{n-3} y_i + \frac{1}{2}(y_n + b)$ lies in $[-1, 1]$ for $k = 1, \dots, n - 3$. By symmetrizing the procedure and proving the following lemma which is slightly stronger than necessary, we will get a RBS that fills all of $M(n)$.

Lemma 10. *Given any $\vec{w} \in M(n)$ where n is odd, there exists a permutation σ of $\{1, 2, \dots, n\}$ and $b \in \{-1, 1\}$ such that for $\vec{z} = \sigma(\vec{w}) := (w_{\sigma(1)}, \dots, w_{\sigma(1)})$ all of the sums*

$$\sum_{i=k}^{n-3} z_i + \frac{1}{2}(z_n + b) \quad (k = 1, \dots, n - 3) \tag{14}$$

lie in $[-1, 1]$.

Proof. Choose $\sigma(n)$ such that $z_n = w_{\sigma(n)} \geqslant 0$. Let $b = -1$. Let $a = -\frac{1}{2}(z_n + b) \in [0, \frac{1}{2}]$. The condition now becomes

$$S_k := \sum_{i=k}^{n-3} z_i \in [-1+a, 1+a] \quad (k = 1, \ldots, n-3).$$

Taking $S_{n-2} = 0$, we will prove the above condition by induction. Assume $\sigma(n-3), \ldots, \sigma(k)$ are defined such that $S_k \in [-1+a, 1+a]$.

If $S_k \in [a, 1+a]$ then $\sum_{i=k}^{n-3} w_{\sigma(i)} + w_{\sigma(n)} \geqslant 0$. Since

$$\sum_{i=k}^{n-3} w_{\sigma(i)} + w_{\sigma(n)} + \sum_{i \notin \sigma(\{k,\ldots,n-3,n\})} w_i = 0$$

we can choose $\sigma(k-1)$ such that $w_{\sigma(k-1)} \leqslant 0$. Then $S_{k-1} \in [-1+a, 1+a]$.

If $S_{k-1} \in [-1+a, a]$ then if possible choose $\sigma(k-1)$ such that $w_{\sigma(k-1)} \geqslant 0$. Then $S_{k-1} \in [-1+a, 1+a]$. Otherwise choose any $\sigma(k-1)$ and $\forall i \notin \sigma(\{k-1, \ldots, n-3, n\})$, $w_{\sigma(i)} \leqslant 0$. So $\sum_{i=k-1}^{n-3} w_{\sigma(i)} + w_{\sigma(n)} \geqslant 0$ and thus $S_{k-1} \geqslant -w_{\sigma(n)} = -1 + 2a \geqslant -1 + a$. Therefore $S_{k-1} \in [-1+a, 1+a]$. □

Theorem 11. *Given $n = 2m+1$, let the RBS $\vec{Y}$ be defined as in the above theorem and let σ be a randomly chosen permutation of $\{1, \ldots, n\}$. Then $\vec{w} = \sigma(\vec{Y})$ defines a RBS with all of $M(n)$ as support.*

Proof. The proof follows from Theorem 9 and Lemma 10 in much the same way as Theorem 8 followed from Theorem 6 and Lemma 7. That is, for any $\vec{w} \in M(n)$ using the σ and b from Lemma 10, $\vec{w} = \sigma^{-1}(\vec{z})$ where all the sums in (14) lie in $[-1, 1]$. The proof of Theorem 9 showed that any such $\vec{z}$ could occur as a RBS $\vec{Y}$ as defined in the statement of that theorem. Thus, $\vec{w}$ can occur as $\tau(\vec{Y})$ for the permutation $\tau = \sigma^{-1}$. □

3. Gerow–Robson densities

We are interested in probability densities on $M(n)$ with respect to the measure v_{n-1} defined as follows. Let $\vec{1}$ denote the vector $(1, 1, \ldots, 1) \in \mathbb{R}^n$, and let $\vec{1}^{\perp}$ denote the hyperplane consisting of all vectors perpendicular to $\vec{1}$. For each Borel subset S of $\vec{1}^{\perp}$, let $v_{n-1}(S)$ denote the $(n-1)$-dimensional Lebesgue measure of S, regarding $\vec{1}^{\perp}$ as an isometric copy of $\mathbb{R}^{n-1}$. Let $V_{n-1} = v_{n-1}(M(n))$.

By a *Gerow–Robson* (G–R) density on $M(n)$ we shall mean a probability density h with respect to v_{n-1} such that

$$h(\vec{X}) = f(\|\vec{X}\|_\infty) \quad (\vec{X} \in M(n)) \tag{15}$$

for some $f : [0, 1] \to [0, \infty)$, where

$$\|\vec{X}\|_\infty = \max_{1 \leqslant k \leqslant n} |X_k|. \tag{16}$$

We remark that on $M(3)$ and $M(4)$ the subspaces with constant ∞-norm are represented by hexagons and octahedra, respectively. This definition is suggested by the densities discovered by Robson and Gerow for the cases $n = 3, 4$. They noted that with $f(s)$ proportional to s, i.e. with

$f_3(s) = C_3 s$ and $f_4(s) = C_4 s$ for certain constants C_k, the corresponding G–R densities define RBS distributions on $M(3)$, $M(4)$, respectively (see Fig. 1). These facts will also follow from the general analysis of G–R densities given below. It was natural to try to extend this construction to larger values of n. We will prove below a rather surprising fact: not only do densities with $f_n(s) = C_n s$ fail to generate RBS distributions when $n > 4$, but (at least for $n \leqslant 250$) no other choices of f_n yield RBS distributions. Concisely, it seems that the *only* G–R densities that yield RBS distributions are those discovered by Robson and Gerow.

As a historical note, Gerow was already aware via numerical simulations reported in [5] that $f(s) = C_5 s$ did not yield a RBS distribution. In [7] this was verified theoretically and it was shown that, in fact, *no* choice of f_n gives a RBS distribution for $n = 5$.

Here, we will extend this result to larger values of n. For $n \geqslant 6$, we will derive a sufficient condition for the nonexistence of a G–R density on $M(n)$. This condition can be verified numerically for $n \leqslant 250$. We will also show that the densities given by Gerow and Robson are the only such densities on $M(3)$ and $M(4)$ and we will reprove the nonexistence of a G–R density on $M(5)$.

We conjecture that the condition we will derive for the nonexistence of a G–R density on $M(n)$ holds for all $n \geqslant 6$. While we are unable to prove this, our computational results settle all cases likely to be of practical interest. For very large samples sizes, constructing RBS is perhaps less important, since for large values of n a uniform sample from $[-1, 1]^n$ is, in view of the law of large numbers, likely to be very nearly balanced.

For any probability measure P on $M(n)$ we have the corresponding distribution function defined on $[0, 1]$ by

$$G(r) = P(rM(n)) = P(\{\vec{X} \in M(n) : \|\vec{X}\|_\infty \leqslant r\}). \tag{17}$$

In those cases where G has a density on $[0, 1]$ we denote it by g:

$$G(r) = P(rM(n)) = \int_0^r g(s)\,ds. \tag{18}$$

If P is defined by a G–R density h corresponding to f, we must have

$$G(r) = \int_0^r f(s)(v_{n-1}(sM(n)))'\,ds, \tag{19}$$

and since $v_{n-1}(sM(n)) = s^{n-1} v_{n-1}(M(n)) = s^{n-1} V_{n-1}$,

$$G(r) = \int_0^r f(s)(n-1)V_{n-1}s^{n-2}\,ds. \tag{20}$$

Thus, $g(s) = (n-1)V_{n-1}s^{n-2}f(s)$ for a G–R density. This relation, along with the fact that $G(1) = 1$, allows us to properly normalize f. For example, in the cases treated by Robson and Gerow we have $f_n(s) = C_n s$; normalizing g we see that $g_n(s) = ns^{n-1}$. Thus, the normalizing constants C_n are given by $\frac{n}{(n-1)V_{n-1}}$. A little computation reveals that $V_2 = 3\sqrt{3}$ and $V_3 = \frac{32}{3}$, so that, in our models, the densities found by Robson and Gerow correspond to $f_3(s) = s/2\sqrt{3}$ and $f_4(s) = s/8$.

Conversely, given any probability density g on $[0, 1]$, we shall see how to construct a probability measure on $M(n)$ with the corresponding G–R density: for $\vec{X} \in M(n)$,

$$h(\vec{X}) = f(\|\vec{X}\|_\infty) = \frac{g(\|\vec{X}\|_\infty)}{(n-1)V_{n-1}\|\vec{X}\|_\infty^{n-2}}. \tag{21}$$

It is natural to ask for which g and for which n we obtain RBS distributions. The answer, as claimed above, is given by the following result.

Theorem 12. *For $n \leqslant 250$, the G–R density on $M(n)$ corresponding to a density g on* [0, 1] *defines a RBS distribution in exactly two cases*: $n = 3$ *with* $g_3(s) = 3s^2$, *and* $n = 4$ *with* $g_4(s) = 4s^3$.

Before turning to the proof of this theorem, we introduce a specific model for the distribution of $\vec{X} \in M(n)$ according to the G–R density implied by a density g on [0, 1]. Consider auxiliary random variables Y_1 and $Z_2, \ldots, Z_n$ defined as follows: $0 \leqslant Y_1 \leqslant 1$ has the given density g, and (independently) $(Z_2, \ldots, Z_n)$ is uniformly distributed with respect to the $(n-2)$-volume on

$$\left\{ \vec{Z} \in [-1, 1]^{n-1} : \sum_{k=2}^{n} Z_k = -1 \right\}. \tag{22}$$

Let $Y_k = Y_1 Z_k$ for $k = 2, \ldots, n$, so that $Y_1 \geqslant |Y_k|$ and $\sum_{k=1}^{n} Y_k = 0$. Then $\vec{X} = \vec{Y}$ yields sample points in that part of $M(n)$ where $\|\vec{X}\|_\infty = X_1$. Finally, we symmetrize over $M(n)$:

$$\vec{X} = \pm(Y_{\sigma(1)}, \ldots, Y_{\sigma(n)}), \tag{23}$$

where the sign $\pm$ and the permutation σ are chosen at random.

Proposition 13. *Relative to $(n-1)$-volume on $M(n)$, this $\vec{X}$ has a density of the G–R form* (21).

Proof. It suffices to consider a point $\vec{X} \in M(n)$ where

$$\|\vec{X}\|_\infty = X_1 > |X_k| \quad (k = 2, \ldots, n),$$

and the behavior as $\Delta t \to 0$ of the probability $P(\vec{X} + \Delta t M(n))$ that a sample point falls in the neighborhood $\vec{X} + \Delta t M(n)$. Denote this probability, briefly, by $P(\Delta t)$. Then

$$P(\Delta t) = P\{\pm = +\} P\{\sigma(1) = 1\} P\{Y_k \in [X_k - \Delta t, X_k + \Delta t] \quad (k = 1, \ldots, n)\}.$$

For small values of Δt, $P\{Y_1 \in [X_1 - \Delta t, X_1 + \Delta t]\}$ is close to $2\Delta t g(X_1)$, and given Y_1 the probability

$$\begin{aligned} &P\{Y_k \in [X_k - \Delta t, X_k + \Delta t] \ (k = 2, \ldots, n)\} \\ &\quad = P\left\{ Z_k \in \left[\frac{X_k}{Y_1} - \frac{\Delta t}{Y_1}, \frac{X_k}{Y_1} + \frac{\Delta t}{Y_1} \right] \ (k = 2, \ldots, n) \right\} \end{aligned}$$

is nearly proportional to $(\Delta t / X_1)^{n-2}$, since $Y_1 \approx X_1$ and the hyperplane $\{\sum_2^n Z_k = -1\}$ makes the same angle with each coordinate axis. Thus, $P(\Delta t)$ is nearly proportional to

$$\frac{1}{2} \cdot \frac{1}{n} \cdot 2\Delta t g(X_1) \cdot (\Delta t / X_1)^{n-2}$$

as $\Delta t \to 0$. It follows that the density of P relative to $(n-1)$-volume, i.e.

$$\lim_{\Delta t \downarrow 0} \frac{P(\Delta t)}{(\Delta t)^{n-1} V_{n-1}},$$

is proportional to

$$\frac{g(X_1)}{X_1^{n-2}} = \frac{g(\|\vec{X}\|_\infty)}{\|\vec{X}\|_\infty^{n-2}}.$$

We have, therefore, a G–R density with $f(s) = Kg(s)/s^{n-2}$, for some constant K, and we must have (21), i.e. $K = 1/(n-1)V_{n-1}$, because of the general relation between f and g for G–R densities. $\square$

To see whether a given density g on $[0, 1]$ generates a RBS distribution on $M(n)$ we must examine the values of $P\{X_k \in [a, b]\}$ in the model defined above. It is sufficient to compute $P\{X_1 \in [0, t]\}$ since the coordinates X_k have been symmetrized, i.e. they are interchangeable and $-X_k$ has the same distribution as X_k. Thus, the sampling vector $\vec{X}$ will have a RBS distribution on $M(n)$ exactly when

$$P\{X_1 \in [0, t]\} = \frac{t}{2} \quad (t \in [0, 1]). \tag{24}$$

Recalling the equidistribution of $Y_2, \ldots, Y_n$, we compute:

$$\begin{aligned} P\{X_1 \in [0, t]\} &= P\{\pm = +\}P\{\sigma(1) = 1\}P\{Y_1 \in [0, t]\} \\ &\quad + P\{\pm = +\}P\{\sigma(1) \neq 1\}P\{Y_2 \in [0, t]\} \\ &\quad + P\{\pm = -\}P\{\sigma(1) \neq 1\}P\{Y_2 \in [-t, 0]\} \\ &= \frac{1}{2n}\int_0^t g(s)\,ds + \frac{1}{2}\left(1 - \frac{1}{n}\right) P\{Y_1 Z_2 \in [-t, t]\} \\ &= \frac{1}{2n}\int_0^t g(s)\,ds + \frac{1}{2}\left(1 - \frac{1}{n}\right) \int_0^1 g(s) P\left\{Z_2 \in \left[\frac{-t}{s}, \frac{t}{s}\right]\right\} ds. \end{aligned}$$

Let $P_n(r) = P\{|Z_2| \leqslant r\}$. Since $P_n(t/s) = 1$ when $s < t$, our condition for a RBS distribution becomes: for all $t \in [0, 1]$,

$$\frac{t}{2} = \frac{1}{2}\int_0^t g(s)\,ds + \frac{1}{2}\left(1 - \frac{1}{n}\right)\int_t^1 g(s) P_n\left(\frac{t}{s}\right) ds. \tag{25}$$

Differentiating with respect to t, we obtain the condition:

$$1 \equiv \frac{1}{n} g(t) + \left(1 - \frac{1}{n}\right)\int_t^1 g(s) P_n'\left(\frac{t}{s}\right) \frac{ds}{s} \quad (t \in [0, 1]). \tag{26}$$

A little geometry reveals that $P_3(r) = r$ and $P_4(r) = r$, so that we may verify that (26) is satisfied with $n = 3$, $g_3(s) = 3s^2$ and with $n = 4$, $g_4(s) = 4s^3$. This is one way to verify the discoveries of Robson and Gerow that $f_n(s) = C_n s$ yields RBS distributions for $n = 3, 4$. By a more involved geometric argument one may obtain $P_5(r) = (24s - s^3)/23$ and see that $g_5(s) = 5s^4$ does not satisfy (26) (with $n = 5$), i.e. that the most “natural” generalization of the constructions of Robson and Gerow does not yield a RBS distribution for sample size 5. To use (26) more systematically, we must first find some general expressions for $P_n'(r)$.

Note that the $(n-2)$-volume on $\mathbb{R} \times [-1, 1]^{n-2} \cap \{\sum_2^n Z_k = -1\}$ can be sampled uniformly by choosing $Z_3, \ldots, Z_n$ independently and uniformly on $[-1, 1]$ and then setting $Z_2 = -1 - \sum_3^n Z_k$.

Thus, the $(n-2)$-volume on $[-1,1]^{n-1} \cap \{\sum_2^n Z_k = -1\}$ can be sampled uniformly by repeating the above procedure until $Z_2 \in [-1,1]$. That is, until $\sum_3^n Z_k \in [0,2]$. Thus,

$$P_n(s) := P\{|Z_2| \leqslant s\} = \frac{P\{\sum_3^n Z_k \in [1-s, 1+s]\}}{P\{\sum_3^n Z_k \in [0,2]\}} \quad (s \in [0,1]), \tag{27}$$

where the Z_k are now independent and uniform over $[-1,1]$. Let ϕ be the uniform density over $[-1,1]$, i.e. $\phi = \frac{1}{2} I_{[-1,1]}$. Then the density ϕ_n of $\sum_3^n Z_k$ is the $(n-2)$-fold convolution of ϕ with itself. Note that ϕ_n is an even function. It follows that, for $s \in [0,1]$,

$$P_n(s) = \frac{\int_{1-s}^{1+s} \phi_n(u)\,du}{\int_0^2 \phi_n(u)\,du},$$

and that

$$P_n'(s) = \frac{\phi_n(1+s) + \phi_n(1-s)}{\int_0^2 \phi_n(u)\,du} = \frac{\phi_n(1+s) + \phi_n(1-s)}{2\phi_{n+1}(1)}.$$

These functions, for $s \in [0,1]$, are polynomials in s. One way to see this, and to obtain explicit expressions for $P_n'(s)$, is to compute in terms of Laplace transforms. By the Laplace transform $\mathcal{L}\psi(s)$ of a function $\psi(t)$ (with left-bounded support) we mean

$$\mathcal{L}\psi(s) = \int_{-\infty}^{\infty} e^{-st}\psi(t)\,dt \quad (s > 0).$$

We shall use several well-known properties of $\mathcal{L}$; for example, $\mathcal{L}$ converts convolution products $\psi_1 \star \psi_2$ into ordinary pointwise products:

$$\mathcal{L}\{\psi_1 \star \psi_2\}(s) = \mathcal{L}\psi_1(s)\mathcal{L}\psi_2(s).$$

Since $\mathcal{L}\phi(s) = (e^s - e^{-s})/2s$, it follows that

$$\mathcal{L}\phi_n(s) = \frac{(e^s - e^{-s})^{n-2}}{2^{n-2}s^{n-2}}.$$

Now $e^{rs}/s^{n-2} = \mathcal{L}\{(t+r)_+^{n-3}/(n-3)!\}(s)$, where t_+ denotes $H(t)t$, $H(t)$ being the Heaviside function. Since $\mathcal{L}$ is injective,

$$\phi_n(t) = \frac{1}{(n-3)!2^{n-2}} \sum_{k=0}^{n-2} \binom{n-2}{k} (-1)^k (t+n-2-2k)_+^{n-3}. \tag{28}$$

In evaluating $P_n'(s)$ (for $s \in [0,1]$), we need only apply (28) for $t \in [0,2]$ and we obtain

$$P_n'(s) = C_n \sum_{k=0}^{n-2} \binom{n-2}{k} (-1)^k \{(n-1-2k+s)_+^{n-3} + (n-1-2k-s)_+^{n-3}\}, \tag{29}$$

where $C_n = 1/((n-3)!2^{n-1}\phi_{n+1}(1))$.

Let us consider the case of even n, say $n = 2m$. Then in (29), $2m-1-2k \pm s \leqslant 0$ if and only if $k \geqslant m$ (always with $s \in [0,1]$). So

$$P_n'(s) = C_n \sum_{k=0}^{m-1} \binom{2(m-1)}{k} (-1)^k \{(2m-1-2k+s)^{n-3} + (2m-1-2k-s)^{n-3}\}. \tag{30}$$

The odd case is the same but with one additional term. In both cases, the odd powers of s cancel, and the polynomial has the form

$$P_n'(s) = C_n \sum_{j=0}^{\alpha_n} c_{j,n} s^{2j}, \tag{31}$$

for certain coefficients $c_{j,n}$, where $\alpha_n = \lfloor (n-3)/2 \rfloor$. These coefficients may be evaluated explicitly in any specific case, by reference to (29); later we will take a somewhat different point of view to derive some general properties of the $c_{j,n}$.

Proposition 14. *The G–R density corresponding to a probability density g_n on* [0,1] *defines a RBS distribution on $M(n)$ exactly when*

$$\mathcal{L}q_n(s) = \frac{n}{s(1+(n-1)\mathcal{L}Q_n(s))}, \tag{32}$$

where $q_n(t) = H(t)g_n(e^{-t})$ and $Q_n(t) = H(t)P_n'(e^{-t})$.

Proof. The condition (26), with $g = g_n$, may be rewritten in the form

$$n \equiv q_n(t) + (n-1)\int_{e^{-t}}^{1} g_n(s) P_n'\left(\frac{e^{-t}}{s}\right) \frac{ds}{s} \quad (t \geqslant 0).$$

With the change of variable $s = e^{-u}$, this becomes

$$n \equiv q_n(t) + (n-1)\int_0^t q_n(u) P_n'(e^{-(t-u)})\, du \quad (t \geqslant 0),$$

i.e.

$$nH(t) = q_n(t) + (n-1)\int_{-\infty}^{\infty} q_n(u) Q_n(t-u)\, du,$$

i.e.

$$nH = q_n + (n-1) q_n \star Q_n.$$

Applying the Laplace transform we obtain

$$\frac{n}{s} = \mathcal{L}q_n(s) + (n-1)\mathcal{L}q_n(s)\mathcal{L}Q_n(s). \qquad \square$$

Now in terms of (31), we have

$$Q_n(t) = H(t) C_n \sum_{j=0}^{\alpha_n} c_{j,n} e^{-2jt},$$

so that

$$\mathcal{L}Q_n(s) = C_n \sum_{j=0}^{\alpha_n} \frac{c_{j,n}}{s+2j}.$$

From Proposition 14 it follows that, in order for g_n to generate a RBS distribution, we must have

$$\mathcal{L}q_n(s) = n \Bigg/ \left(s \left(1 + (n-1)C_n \sum_{j=0}^{\alpha_n} \frac{c_{j,n}}{s+2j} \right) \right). \tag{33}$$

Using this condition one can check that the densities discovered by Gerow and Robson are the only G–R densities for $n = 3, 4$. Indeed evaluating (33), $\mathcal{L}q_3(s) = \frac{3}{s+2}$ and $\mathcal{L}q_4(s) = \frac{4}{s+3}$. Thus, since $\mathcal{L}$ is injective, $q_3(s) = 3e^{-2t}$ and $q_3(s) = 4e^{-3t}$. So $g_3(t) = 3t^2$ and $g_4(t) = 4t^3$.

Continuing, using (33) one can check that $\mathcal{L}q_5(s) = \frac{115(s+2)}{23s^2+130s+192}$ and thus $q_5(t) = \frac{5}{191}e^{-\frac{65}{23}t}$ $(191\cos(\frac{\sqrt{191}}{23}t) - 19\sqrt{191}\sin(\frac{\sqrt{191}}{23}t))$. However $q_5(1.5) < 0$. This contradicts the fact that $q(t) = g(e^{-t})$ and g is a probability density on [0, 1]. Therefore, there is no G–R density on $M(5)$.

We can generalize this nonexistence result to higher n without explicitly calculating $\mathcal{L}q_n$ and q_n.

Let

$$B_n(s) = \left(\prod_{j=0}^{\alpha_n} (s+2j) \right) \left(1 + (n-1)C_n \sum_{j=0}^{\alpha_n} \frac{c_{j,n}}{s+2j} \right).$$

Then $B_n(s)$ is a polynomial of degree $\alpha_n + 1$ and

$$\mathcal{L}q_n(s) = \frac{n\prod_{j=1}^{\alpha_n}(s+2j)}{B_n(s)}.$$

Assume that $B_n(s)$ has $\alpha_n + 1$ real distinct roots $a_0 > a_1 > \cdots a_{\alpha_n}$. Then for some $k \neq 0$ and some $\{b_j \neq 0\}$,

$$\mathcal{L}q_n(s) = \frac{n\prod_{j=1}^{\alpha_n}(s+2j)}{k\prod_{j=0}^{\alpha_n}(s-a_j)} = \sum_{j=0}^{\alpha_n} \frac{b_j}{s-a_j}.$$

Since $\mathcal{L}$ is injective,

$$q_n(t) = \sum_{j=0}^{\alpha_n} b_j e^{a_j t}.$$

Since $a_0 > a_1 > \cdots > a_{\alpha_n}$, for large values of t the sign of $q_n(t)$ equals the sign of b_0. Since g is a probability density, $q(t) \geqslant 0$ for all $t \geqslant 0$. So $b_0 \geqslant 0$. We have shown the following.

Theorem 15. *Let $n \geqslant 6$. If $B_n(s)$ has $\alpha_n + 1$ distinct real roots and $b_0 < 0$ then there does not exist a G–R density on $M(n)$.*

With a few hours of computation on a desktop computer using a computational program such as Maple one can verify that $B_n(s)$ has $\alpha_n + 1$ distinct real roots for $6 \leqslant n \leqslant 250$. We remark that we can show if $B_n(s)$ has $\alpha_n + 1$ distinct real roots then $-3 < a_0 < -2$ and $b_0 < 0$. However, to prove Theorem 12 we do not need to use this fact. The second condition can be verified computationally for $6 \leqslant n \leqslant 250$. For example, for a given n one can check that $-3 < a_0 < -2$. Since for all

$s \in (-3, -2)$, the sign of $B_n(s)$ is the opposite of the sign of $\mathcal{L}q_n(s)$, it follows that $b_0 < 0$ if and only if $B_n(-3) < 0$ and $B_n(-2) > 0$. This proves Theorem 12.

References

[1] N. Arvidsen, T. Johnsson, Variance reduction through negative correlation, a simulation study, J. Statist. Comput. Simulation (1982) 119–127.
[2] R.V. Craiu, X.-L. Meng, Antithetic coupling for perfect sampling, Bayesian Methods with Applications to Science, Policy, and Official Statistics—Selected Papers from ISBA 2000.
[3] R.V. Craiu, X.-L. Meng, Chance and fractals, Chance 14 (2001) 47–52.
[4] R.V. Craiu, X.-L. Meng, Multiprocess parallel antithetic coupling for backward and forward Markov chain Monte Carlo, Ann. Statist. 33 (2) (2005) 661–697.
[5] K. Gerow, An algorithm for random balance sampling, Master's Thesis, University of Guelph, 1984.
[6] K. Gerow, Model-unbiased, unbiased-in-general estimation of a regression function, Ph.D. Thesis, Cornell University, 1993.
[7] K. Gerow, J. Holbrook, Construction of random balanced samples, Working Paper, 1990.
[8] K. Gerow, J. Holbrook, Statistical sampling and fractal distributions, Math. Intelligencer 18 (2) (1996) 12–22.
[9] K. Gerow, C.E. McCulloch, Simultaneously model-unbiased, design-unbiased estimation, Biometrics 56 (3) (2000) 873–878.
[10] P. Hall, Antithetic resampling for the bootstrap, Biometrika 76 (4) (1989) 713–724.
[11] J.M. Hammersley, K.W. Morton, A new Monte Carlo technique: antithetic variates, Proc. Cambridge Philos. Soc. 52 (1956) 449–475.
[12] P. Hoel, Introduction to Mathematical Statistics, Wiley, New York, 1954.
[13] R. Ramlochan, Iterated function systems and fractal sampling, Master's Thesis, University of Guelph, 1990.
[14] R. Royall, W. Cumberland, An empirical study of the ratio estimator and estimators of its variance, Discussion paper in J. Amer. Statist. Soc. 76 (373) (1981) 66–88.