LUND UNIVERSITY

Generating random variates from a bicompositional Dirichlet distribution

Bergman, Jakob

Total number of authors:
1

RESEARCH ARTICLE

Generating random variates from a bicompositional Dirichlet distribution

Jakob Bergman*

Department of Statistics, Lund University, Box 743, SE-220 07 Lund, Sweden

A composition is a vector of positive components summing to a constant. The sample space of a composition is the simplex and the sample space of two compositions, a bicomposition, is a Cartesian product of two simplices. We present a way of generating random variates from a bicompositional Dirichlet distribution defined on the Cartesian product of two simplices using the rejection method. We derive a general solution for finding a dominating density function and a rejection constant, and also compare this solution to using a uniform dominating density function. Finally some examples of generated bicompositional random variates, with varying number of components, are presented.

1. Introduction

A composition is a vector of positive components summing to a constant. The components of a composition are what we usually think of as proportions (at least when the vector sums to 1). Compositions arise in many different areas; the geo-chemical compositions of different rock specimens, the proportion of expenditures on different commodity groups in household budgets, and the party preferences in a party preference survey are all examples of compositions from three different scientific areas. For more examples of compositions, see for instance [1].

The sample space of a composition is the simplex. Without loss of generality we will always take the summing constant to be 1, and we define the D-dimensional simplex $\mathscr{S}^D$ as

$$\mathscr{S}^D = \left\{ \boldsymbol{x} = (x_1, \ldots, x_D)^{\mathrm{T}} \in \mathscr{R}_+^D : \sum_{j=1}^{D} x_j = 1 \right\},$$

where $\mathscr{R}_+$ is the positive real space. The joint sample space of two compositions is the Cartesian product of two simplices $\mathscr{S}^D \times \mathscr{S}^D$. It should be noted that, unlike the case for real Cartesian product spaces, $\mathscr{S}^D \times \mathscr{S}^D \neq \mathscr{S}^{D+D}$ and that

*Email: jakob.bergman@stat.lu.se

$\mathscr{S}^D \times \mathscr{S}^D$ is not even a simplex, but a manifold with two constraints, i.e. the subspace of $\mathscr{R}_+^{D+D}$ where

$$\sum_{j=1}^{D} x_j = \sum_{j=D+1}^{D+D} x_j = 1.$$

The bicompositional Dirichlet distribution, for modelling random vectors on $\mathscr{S}^D \times \mathscr{S}^D$, was proposed in [2]. The proposed distribution has the probability density function

$$f(\boldsymbol{x}, \boldsymbol{y}) = A(\boldsymbol{\alpha}, \boldsymbol{\beta}, \gamma) \left(\prod_{j=1}^{D} x_j^{\alpha_j - 1} y_j^{\beta_j - 1} \right) (\boldsymbol{x}^{\mathrm{T}} \boldsymbol{y})^{\gamma}, \tag{1}$$

where $\boldsymbol{x}, \boldsymbol{y} \in \mathscr{S}^D$ and $\alpha_j, \beta_j \in \mathscr{R}_+ (j = 1, \ldots, D)$. The parameter space of γ depends on $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_D)^{\mathrm{T}}$ and $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_D)^{\mathrm{T}}$; however, all non-negative values are always included. Expressions for the normalization constant A are given in [2]. For instance, when $D = 2$ the distribution exists if $\gamma > -\min(\alpha_1 + \beta_2, \alpha_2 + \beta_1)$ and A is determined by

$$\frac{1}{A} = \frac{1}{2^\gamma} \sum_{i=0}^{\infty} \binom{\gamma}{i} \left(\sum_{j=0}^{i} \binom{i}{j} (-1)^{i-j} B(\alpha_1 + j, \alpha_2 + i - j) \right)$$
$$\cdot \left(\sum_{k=0}^{i} \binom{i}{k} (-1)^{i-k} B(\beta_1 + k, \beta_2 + i - k) \right)$$

where $B(p, q)$ is the Beta function, and when $D > 2$ and γ is a non-negative integer A is determined by

$$\frac{1}{A} = \sum_{\substack{\boldsymbol{k} \geq 0 \\ k. = \gamma}} \binom{\gamma}{\boldsymbol{k}} \frac{\prod_{i=1}^{D} \Gamma(\alpha_i + k_i)}{\Gamma(\alpha. + \gamma)} \frac{\prod_{i=1}^{D} \Gamma(\beta_i + k_i)}{\Gamma(\beta. + \gamma)},$$

where $\alpha. = \alpha_1 + \cdots + \alpha_D$, $\beta. = \beta_1 + \cdots + \beta_D$, $k. = k_1 + \cdots + k_D$, $\boldsymbol{k} = (k_1, \ldots, k_D)$ and

$$\binom{\gamma}{\boldsymbol{k}} = \frac{\gamma!}{k_1! \cdots k_D!}.$$

If $\gamma = 0$, the probability density function (1) is the product of two Dirichlet probability density functions with parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ respectively, and hence $\boldsymbol{X}$ and $\boldsymbol{Y}$ are independent in that case.

When $\boldsymbol{X}, \boldsymbol{Y} \in \mathscr{S}^2$ we shall refer to this as the *bicomponent* case, and similarly to $\mathscr{S}^3$ as the *tricomponent* case, and to $\mathscr{S}^D (D > 2)$ as the *multicomponent* case.

Apart from the multinormal and Wishart distributions, papers on generating bivariate and multivariate random variates are rare and most suggested general methods have disadvantages [3]. The only universal algorithm for generating multivariate random variates is the algorithm presented in [4], which is a generalization of algorithms for the univariate and bivariate case given in different versions in [5] and [6]. However, in [3] it is noted that this algorithm is very slow and an alternative algorithm, which requires a function of the density to be concave, is suggested.

The class of distributions given in (1) is not necessarily unimodal and may not even be bounded, and it is therefore hard to find a function of the density that is concave for all parameter values. Hence we will use the rejection method to construct a specialized method for generating bicompositional random variates.

Let f be the density from which we wish to generate random variates. Let $c \geq 1$ be a constant and g be a density such that

$$f(\boldsymbol{x}, \boldsymbol{y}) \leq cg(\boldsymbol{x}, \boldsymbol{y}) \tag{2}$$

for all $(\boldsymbol{x}, \boldsymbol{y})$. We now generate a random variate $(\boldsymbol{X}, \boldsymbol{Y})$ with density g and a random number U uniformly distributed on the unit interval. The variate $(\boldsymbol{X}, \boldsymbol{Y})$ is accepted if

$$U \leq \frac{f(\boldsymbol{X}, \boldsymbol{Y})}{cg(\boldsymbol{X}, \boldsymbol{Y})},$$

otherwise we reject $(\boldsymbol{X}, \boldsymbol{Y})$ and generate new $(\boldsymbol{X}, \boldsymbol{Y})$ and U until acceptance.

We thus need to find a dominating density g and constant c, and preferably such choices of g and c that will give high probabilities of acceptance and hence make the random variate generation efficient.

2. Generating random bicompositions

2.1. *The case when $\gamma = 0$*

A Dirichlet distributed random variate is easily generated using Gamma distributed variates. Let V_i be a Gamma distributed variate with parameter $(\alpha_i, 1)$ and $X_i = V_i / \sum_{j=1}^{D} V_j$ $(i = 1, \ldots, D)$, then $\boldsymbol{X} = (X_1, \ldots, X_D)$ is Dirichlet distributed with parameter $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_D)$ [7, pp. 593–596]. There are also other more or less efficient ways to generate Dirichlet distributed variates. These are reviewed and compared in [8].

Hence, to generate a random bicompositional Dirichlet distributed variate $(\boldsymbol{x}, \boldsymbol{y})$ with parameter $(\boldsymbol{\alpha}, \boldsymbol{\beta}, 0)$, we need only to generate a Dirichlet distributed variate $\boldsymbol{x}$ with parameter $\boldsymbol{\alpha}$ and a Dirichlet distributed variate $\boldsymbol{y}$ with parameter $\boldsymbol{\beta}$.

2.2. *The case when $\gamma > 0$*

When $\gamma > 0$, we may use the product of two Dirichlet distributions, i.e. a bicompositional Dirichlet distribution with $\gamma = 0$, as a dominating density, since $0 < \boldsymbol{x}^{\mathrm{T}} \boldsymbol{y} < 1$ and thus

$$A(\boldsymbol{\alpha}, \boldsymbol{\beta}, \gamma) \left(\prod_{j=1}^{D} x_j^{\alpha_j - 1} y_j^{\beta_j - 1} \right) (\boldsymbol{x}^{\mathrm{T}} \boldsymbol{y})^{\gamma} \leq A(\boldsymbol{\alpha}, \boldsymbol{\beta}, \gamma) \left(\prod_{j=1}^{D} x_j^{\alpha_j - 1} y_j^{\beta_j - 1} \right).$$

The inequality (2) now becomes

$$A(\boldsymbol{\alpha}, \boldsymbol{\beta}, \gamma) \left(\prod_{j=1}^{D} x_j^{\alpha_j - 1} y_j^{\beta_j - 1} \right) (\boldsymbol{x}^{\mathrm{T}} \boldsymbol{y})^{\gamma} \leq cA(\boldsymbol{\alpha}, \boldsymbol{\beta}, 0) \left(\prod_{j=1}^{D} x_j^{\alpha_j - 1} y_j^{\beta_j - 1} \right),$$

which holds if we choose

$$c = \frac{A(\boldsymbol{\alpha}, \boldsymbol{\beta}, \gamma)}{A(\boldsymbol{\alpha}, \boldsymbol{\beta}, 0)} > 1. \tag{3}$$

Here $A(\boldsymbol{\alpha}, \boldsymbol{\beta}, \gamma)$ is inversely proportional to the integral of $x_1^{\alpha_1 - 1} \cdots x_D^{\alpha_D - 1} y_1^{\beta_1 - 1} \cdots y_D^{\beta_D - 1} (\boldsymbol{x}^{\mathrm{T}} \boldsymbol{y})^{\gamma}$. Since the crossproduct $\boldsymbol{x}^{\mathrm{T}} \boldsymbol{y}$ is bounded to $(0, 1)$, it follows that $A(\boldsymbol{\alpha}, \boldsymbol{\beta}, \gamma)$ is a non-negative increasing function of $\gamma \geq 0$ for fixed $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. Hence the last inequality of (3) is thus true.

A random variate $(\boldsymbol{x}, \boldsymbol{y})$ with a bicompositional Dirichlet distribution with parameter $(\boldsymbol{\alpha}, \boldsymbol{\beta}, 0)$ is generated as described in Section 2.1. We accept the variate $(\boldsymbol{x}, \boldsymbol{y})$ if

$$U \leq \frac{A(\boldsymbol{\alpha}, \boldsymbol{\beta}, \gamma) \left(\prod_{j=1}^{D} x_j^{\alpha_j - 1} y_j^{\beta_j - 1} \right) (\boldsymbol{x}^{\mathrm{T}} \boldsymbol{y})^{\gamma}}{\frac{A(\boldsymbol{\alpha}, \boldsymbol{\beta}, \gamma)}{A(\boldsymbol{\alpha}, \boldsymbol{\beta}, 0)} A(\boldsymbol{\alpha}, \boldsymbol{\beta}, 0) \left(\prod_{j=1}^{D} x_j^{\alpha_j - 1} y_j^{\beta_j - 1} \right)},$$

i.e. if $U \leq (\boldsymbol{x}^{\mathrm{T}} \boldsymbol{y})^{\gamma}$; otherwise it is rejected and new $(\boldsymbol{x}, \boldsymbol{y})$ and U are generated until acceptance.

We note that this procedure does not require the calculation of $A(\boldsymbol{\alpha}, \boldsymbol{\beta}, \gamma)$ and hence is applicable for all non-negative γ. We thus have the slightly surprising situation that we may generate random variates from distributions whose densities we cannot calculate, since at present expressions for A are not known for non-integer γ in the multicomponent case [2].

Using a product of two Dirichlet distributions as dominating density is however not always very efficient, as $(\boldsymbol{x}^{\mathrm{T}} \boldsymbol{y})^{\gamma}$ may be close to 0 when γ is large. When $\gamma \geq 0$, and $\alpha_j, \beta_j \geq 1 \ (j = 1, \ldots, D)$, it is easily seen that the density (1) will have an upper bound. We may therefore use a uniform density as g, with $c = \max_{\boldsymbol{x}, \boldsymbol{y}} f(\boldsymbol{x}, \boldsymbol{y})$. This is though only applicable for non-negative integers γ, since it is necessary to calculate $A(\boldsymbol{\alpha}, \boldsymbol{\beta}, \gamma)$.

2.3. The case when $\gamma < 0$ and $D = 2$

The bicomponent case is simpler as $\boldsymbol{x} = (x, 1-x)^{\mathrm{T}}$ and $\boldsymbol{y} = (y, 1-y)^{\mathrm{T}}$. This has enabled the distribution to be defined also for $\gamma < 0$. We will in this section view the density as a function of x and y.

The bicomponent bicompositional Dirichlet density exists if and only if $\gamma > -\min(\alpha_1 + \beta_2, \alpha_2 + \beta_1)$ [2]. If $\gamma < 0$, the factor $(\boldsymbol{x}^{\mathrm{T}} \boldsymbol{y})^{\gamma}$ will tend to infinity when x is close to 0 and y is close to 1, and also when x is close to 1 and y is close to 0. We therefore divide the sample space $\mathscr{S}^2 \times \mathscr{S}^2$ into four quadrants, denoted Q_1-Q_4 counter-clockwise from the origin. Figure 1 shows the $\mathscr{S}^2 \times \mathscr{S}^2$ with the four quadrants.

To generate a random variate from a bicomponent bicompositional Dirichlet distribution with parameters $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ and $-\min(\alpha_2, \beta_2) < \gamma < 0$, we first randomly choose a quadrant $Q_k \ (k = 1, 2, 3, 4)$ with probability

$$p_k = \iint_{Q_k} f(x, y) \mathrm{d}x \mathrm{d}y \qquad (k = 1, 2, 3, 4),$$

where $f(x, y)$ is the bicomponent bicompositional Dirichlet probability density function (1) viewed as a function of x and y. Expressions for the cumulative dis-
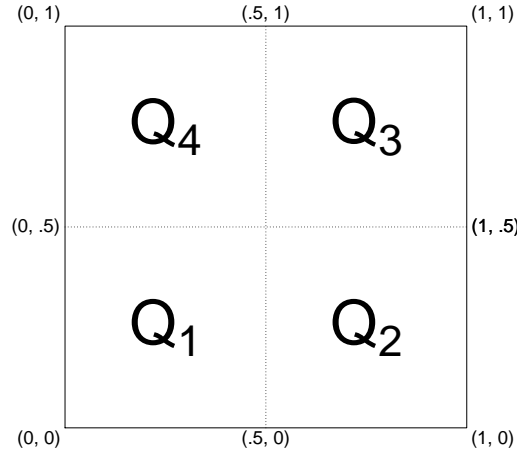
Figure 1. The four quadrants Q_1-Q_4 of the sample space $\mathscr{S}^2 \times \mathscr{S}^2$; the horizontal axis represents x and the vertical axis represents y.

tribution function are given in [2], which may be used in calculating p_k:

$$
F_{X,Y}(x,y) = \frac{A}{2^\gamma} \sum_{i=0}^{\infty} \binom{\gamma}{i} \left(\sum_{j=0}^{i} \binom{i}{j}(-1)^{i-j} B_x(\alpha_1 + j, \alpha_2 + i - j) \right)
$$
$$
\cdot \left(\sum_{k=0}^{i} \binom{i}{k}(-1)^{i-k} B_y(\beta_1 + k, \beta_2 + i - k) \right)
$$

where $B_x(p,q)$ is the incomplete Beta function. Depending on which quadrant is chosen, we then choose a dominating density g and a constant c in the following manner.

$\boldsymbol{Q_1}$ & $\boldsymbol{Q_3}$ In quadrants Q_1 and Q_3, $\boldsymbol{x}^\mathsf{T}\boldsymbol{y} > 1/2$ and we may hence use a product of two Dirichlet (or equivalently Beta) distributions with parameters $\boldsymbol{\alpha}$ respectively $\boldsymbol{\beta}$ as g and a constant

$$
c = \frac{A(\boldsymbol{\alpha}, \boldsymbol{\beta}, \gamma)}{A(\boldsymbol{\alpha}, \boldsymbol{\beta}, 0)2^\gamma}.
$$

$\boldsymbol{Q_2}$ In quadrant Q_2, we may find a lower bound for $\boldsymbol{x}^\mathsf{T}\boldsymbol{y}$ if we introduce a weight $0 < \theta < 1$:

$$
\begin{aligned}
\boldsymbol{x}^\mathsf{T}\boldsymbol{y} &= xy + (1-x)(1-y) \\
&> \frac{1-x}{2} + \frac{y}{2} \\
&= \theta\frac{1-x}{2\theta} + (1-\theta)\frac{y}{2(1-\theta)} \\
&\geq \left(\frac{1-x}{2\theta}\right)^\theta \left(\frac{y}{2(1-\theta)}\right)^{1-\theta} \\
&= \frac{1}{2\theta^\theta(1-\theta)^{1-\theta}}(1-x)^\theta y^{1-\theta}
\end{aligned}
$$

Since $\gamma < 0$,

$$
(\boldsymbol{x}^\mathsf{T}\boldsymbol{y})^\gamma \leq \frac{1}{2^\gamma\theta^{\gamma\theta}(1-\theta)^{\gamma(1-\theta)}}(1-x)^{\gamma\theta}y^{\gamma(1-\theta)}
$$

and

$$
\begin{aligned}
f(x,y;\boldsymbol{\alpha},\boldsymbol{\beta},\gamma) &= A(\boldsymbol{\alpha},\boldsymbol{\beta},\gamma)x^{\alpha_1-1}(1-x)^{\alpha_2-1}y^{\beta_1-1}(1-y)^{\beta_2-1}(\boldsymbol{x}^\mathsf{T}\boldsymbol{y})^\gamma \\
&\leq \frac{A(\boldsymbol{\alpha},\boldsymbol{\beta},\gamma)x^{\alpha_1-1}(1-x)^{\alpha_2+\gamma\theta-1}y^{\beta_1+\gamma(1-\theta)-1}(1-y)^{\beta_2-1}}{2^\gamma\theta^{\gamma\theta}(1-\theta)^{\gamma(1-\theta)}}
\end{aligned}
$$

A dominating density could hence be a the product of two independent Dirichlet densities with parameters $(\alpha_1, \alpha_2 + \gamma\theta)$ and $(\beta_1 + \gamma(1-\theta), \beta_2)$, respectively. For this distribution to exist

$$
\begin{cases}
\alpha_2 + \gamma\theta > 0 \\
\beta_1 + \gamma(1-\theta) > 0
\end{cases}
$$

or equivalently

$$
\begin{cases}
\gamma > -\frac{\alpha_2}{\theta} \\
\gamma > -\frac{\beta_1}{1-\theta}
\end{cases}
$$

To solve these inequalities we assume

$$
\frac{\alpha_2}{\theta} = \frac{\beta_1}{1-\theta}
$$

and solve for θ, which yields the solution

$$
\theta = \frac{\alpha_2}{\alpha_2 + \beta_1}.
$$

We therefore use a product of two Dirichlet distributions with parameters $(\alpha_1, \alpha_2 + \xi\alpha_2)$ and $(\beta_1 + \xi\beta_1, \beta_2)$, respectively, as the density g and the constant c given by

$$
c = \frac{A(\boldsymbol{\alpha},\boldsymbol{\beta},\gamma)(\alpha_2 + \beta_1)^\gamma}{A(\alpha_1, \alpha_2 + \xi\alpha_2, \beta_1 + \xi\beta_1, \beta_2, 0)2^\gamma\alpha_2^{\alpha_2\xi}\beta_1^{\beta_1\xi}},
$$

Table 1. Comparisons of the estimated acceptance probabilities depending on choice of dominating density. We clearly see that the product of two Dirichlet densities can be very inefficient for large values of γ, but also that it may be much more efficient than a uniform density for some distributions.

Parameter values					Dominating density	
α_1	α_2	β_1	β_2	γ	Dirichlet	Uniform
2.1	3.1	5.5	2.3	0.3	0.769	0.222
2.1	3.1	5.5	2.3	3.2	0.103	0.200
2.1	3.1	5.5	2.3	7.7	0.007	0.110
2.1	3.1	5.5	2.3	-1.2	0.208	0.208
2.1	3.1	0.7	2.3	3.2	0.185	NA
7.1	4.2	6.3	8.5	0.3	0.769	0.119
7.1	4.2	6.3	8.5	3.2	0.100	0.125
7.1	4.2	6.3	8.5	7.7	0.005	0.135
7.1	1.2	12.5	3.1	3.2	0.357	0.031

where

$$\xi = \frac{\gamma}{\alpha_2 + \beta_1}.$$

$\boldsymbol{Q_4}$ Analogously to quadrant Q_2, we may in quadrant Q_4 use a product of two Dirichlet distributions with parameters $(\alpha_1 + \eta\alpha_1, \alpha_2)$ and $(\beta_1, \beta_2 + \eta\beta_2)$, respectively, as the density g and the constant c given by

$$c = \frac{A(\boldsymbol{\alpha}, \boldsymbol{\beta}, \gamma)(\alpha_1 + \beta_2)^\gamma}{A(\alpha_1 + \eta\alpha_1, \alpha_2, \beta_1, \beta_2 + \eta\beta_2, 0)2^\gamma \alpha_1^{\alpha_1\eta}\beta_2^{\beta_2\eta}},$$

where

$$\eta = \frac{\gamma}{\alpha_1 + \beta_2}.$$

In all four cases, we must though assure that the generated variates with density g are restricted to that particular quadrant.

3. Comparison of the two dominating densities

The efficiency of the generation process will usually depend on the choice of dominating density. In most cases we have a possibility to choose between two different dominating densities: a product of two independent Dirichlet densities or a uniform density. In general, the product of two Dirichlet distributions will often be more efficient when γ is close to 0, but may however be highly inefficient when γ is large.

To compare the efficiencies of the two dominating densities we generated 25,000 random variates for each of the dominating densities from a number of different bicomponent bicompositional Dirichlet distributions, and calculated the average number of trials to generate one random variate. Table 1 shows the results presented as the estimated probability of acceptance (the reciprocal of the average number of trials) as well as the results for a distribution where only a Dirichlet product is available as dominating density as the distribution density function does not have an upper bound. We note that the probability of acceptance with a uniform density can be much (almost 30 times) larger than the probability of acceptance with a with a Dirichlet density. On the other hand we also see that there are distributions for which the probability of acceptance with a with a Dirichlet density is more
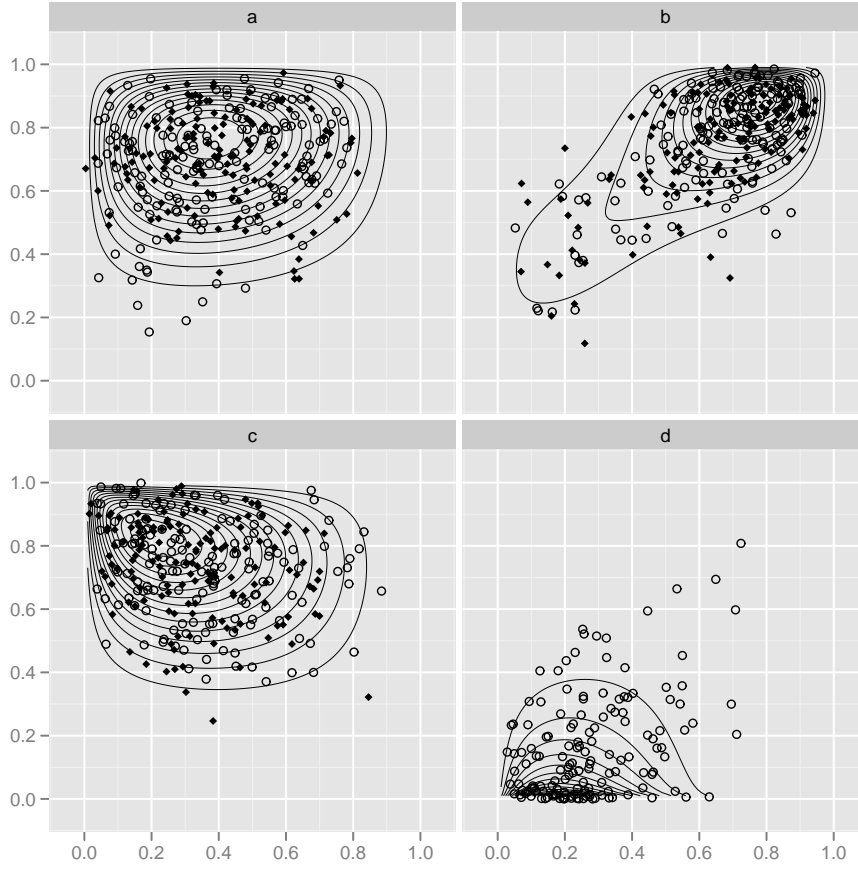
Figure 2. 150 random variates generated from four different bicomponent bicompositional Dirichlet distributions with $(\boldsymbol{\alpha}; \boldsymbol{\beta}; \gamma)$ parameters $(2.1, 3.1; 5.5, 2.3; 0.3)$ (a), $(2.1, 3.1; 5.5, 2.3; 7.7)$ (b), $(2.1, 3.1; 5.5, 2.3; -1.2)$ (c), and $(2.1, 3.1; 0.7, 2.3; 3.2)$ (d), using the product of two Dirichlet densities ($\circ$) and a uniform density ($\blacklozenge$) as dominating density. Since the distribution in (d) does not have an upper bound, a uniform density may not be utilized. As a reference, the contour curves of the true densities are also drawn.

Table 2. Comparisons of the estimated acceptance probabilities for some multicomponent bicompositional Dirichlet distributions using a Dirichlet and a uniform dominating density.

Parameter values			Density	
$\boldsymbol{\alpha}$	$\boldsymbol{\beta}$	γ	Dir.	Unif.
(2, 2, 2)	(2, 2, 2)	1	0.333	0.145
(2, 2, 2)	(2, 2, 2)	7	0.001	0.085
(2.1, 1.2, 3.2, 4.1, 2.8)	(3.2, 2.2, 5.3, 1.8, 2.9)	1	0.204	0.000
(2.1, 1.2, 3.2, 4.1, 2.8)	(3.2, 2.2, 5.3, 1.8, 2.9)	3	0.009	0.000

than 10 times the probability of acceptance with a uniform density. As a graphical illustration of the differences between the distributions, 150 generated random variates from four of the distributions in Table 1 are plotted for each of the two dominating densities in Figure 2 together with contour curves of the density.

The differences in efficiency between the two dominating densities is even more obvious for the multicomponent bicompositional Dirichlet distribution examples presented in Table 2. Here again, we generated 25,000 random variates, this time from four different multicomponent bicompositional Dirichlet distributions using both of the two dominating densities. For the tricomponent distributions, when $\gamma = 1$, the Dirichlet density has a probability of acceptance of more than twice

that of the uniform density, but when $\gamma = 7$ the probability of acceptance of the uniform density is more than 80 times that of the the Dirichlet density. For the two distributions with five components, we see that the Dirichlet density is much more effective for both cases. In general, as the dimension D increases the rejection constant often deteriorates quickly when a uniform density is used [7, p. 557].

4.   Conclusions

The choice of the dominating density is evidently crucial to the efficiency of this random variate generation. When γ is close to 0 or the number of components is large, a product of two Dirichlet density functions seems the most efficient, otherwise a uniform density function (if possible) is recommended. What is meant by close is however dependent of the other parameters $(\boldsymbol{\alpha}, \boldsymbol{\beta})$, so when in doubt, the recommendation would be to generate a small number of variates with each dominating density and see which is the most efficient for the particular parameter values in question. We note that the efficiency of the method seems to degrade as the dimension (i.e. the number of components) increases, and that further research is needed to find more efficient dominating densities for distributions with a large number of components and for large γ values.

It remains yet to find a way of generating random numbers for the bicomponent case when $-\min(\alpha_1 + \beta_2, \alpha_2 + \beta_1) < \gamma < -\min(\alpha_2, \beta_2)$ and the density function does not have an upper bound.

The random variate generation might further be made more efficient for at least the bicomponent case, by adopting the quadrant scheme also for positive γ; especially when the probability mass is concentrated in one or two of the quadrants, which is often the case for large γ, this might speed up the generation process considerably.

Acknowledgement

References

[1] J. Aitchison *The Statistical Analysis of Compositional Data*, The Blackburn Press, Caldwell, NJ, 2003.

[2] J. Bergman, *A Bicompositional Dirichlet Distribution*, Technical Report 3, Department of Statistics, Lund University, 2009. Available at http://www.lu.se/o.o.i.s?id=12730&postid=1514940

[3] J. Leydold, *A rejection technique for sampling from log-concave multivariate distributions*, ACM Transactions on Modeling and Computer Simulation 8 (1998), pp. 254–280.

[4] J. Leydold and W. Hörmann, *A sweep-plane algorithm for generating random tuples in simple polytopes*, Math. Comp. 67 (1998), pp. 1617–1635.

[5] W.R. Gilks and P. Wild, *Adaptive Rejection Sampling for Gibbs Sampling*, Applied Statistics 41 (1992), pp. 337–348.

[6] W. Hörmann, *A rejection technique for sampling from T-concave distributions*, ACM Transactions on Mathematical Software 21 (1995), pp. 182–193.

[7] L. Devroye *Non-uniform random variate generation*, Springer-Verlag, New York, 1986.

[8] Y.C. Hung, N. Balakrishnan, and C.W. Cheng, *Evaluation of algorithms for generating Dirichlet random vectors*, J. Stat. Comput. Simul. (2010). DOI: 10.1080/00949650903409999