3.3 Adaptive Dirichlet distributions

As stated earlier, Krzysztofowicz and Reese (1993) proposed the so-called adaptive Dirichlet (AD) distributions on the unit simplex. They presented a number of examples illustrating their approach, and showed that the Connor-Mosimann distribution can be derived as a special case of their model. As noted by its developers, "The family of distributions characterized herein constitutes the ultimate generalization of the Dirichlet distribution that can be obtained through an independent bifurcation process. This generalization does not totally relax the constraints on the correlation structure of fractions that the Connor-Mosimann and standard Dirichlet distributions impose. It nonetheless offers a much richer model of the stochastic dependence among fractions." However, we have pointed out that the adaptive Dirichlet family is itself a subclass of the multiple Dickey family. In chapter 5, we will relax the independent bifurcation assumption considered by Krzysztofowicz and Reese, using bifurcations with dependent ratios to come up with a new family of distributions that relaxes the constraints on correlation structure imposed by AD distributions.

To illustrate the idea behind adaptive Dirichlet distributions, let us start with the Connor-Mosimann distribution. The process that results in the Connor-Mosimann distribution is a bifurcation process that divides a unit into two parts according to a Bernoulli process with parameter y_1, where y_1 is taken to be beta distributed; the resultant right part is divided into two parts according to another Bernoulli process that is independent of the first one with parameter y_2; and so on until n bifurcations produce the fractions of interest $x_1, x_2, \ldots, x_{n+1}$. The y_i are called ratios, since they are essentially ratios of our original fractions.

Krzysztofowicz and Reese represented this bifurcation process graphically in the form of a *topology*, as shown in Figure 3.2. Adopting their terminology, the circles represent bifurcation nodes, and the branches are labeled with the ratio terms. Krzysztofowicz and Reese called this kind of topology the *cascaded topology* and the resultant distribution the ***Dirichlet type C*** distribution. Note that we have previously defined the notation $CM_n(\alpha, \lambda)$ for this distribution.

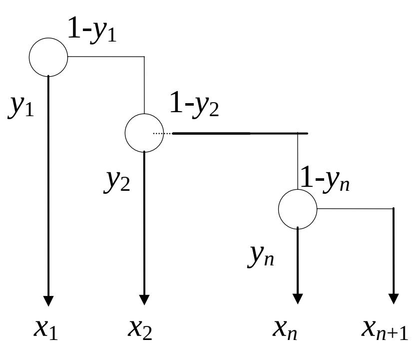

Figure 3.2. Cascaded bifurcation topology underlying the Dirichlet type C distribution (taken from Krzysztofowicz and Reese, 1993).

Now consider the case where the first node bifurcates the unit into two parts, but then each part is bifurcated through a cascaded subtopology, as shown in Figure 3.3. The resultant topology and distribution are the *double-cascaded topology* and the ***Dirichlet type B*** ($AD_{n+1}^{(B)}(\alpha, \lambda, \mathrm{k})$) distribution, respectively, as named by Krzysztofowicz and Reese.

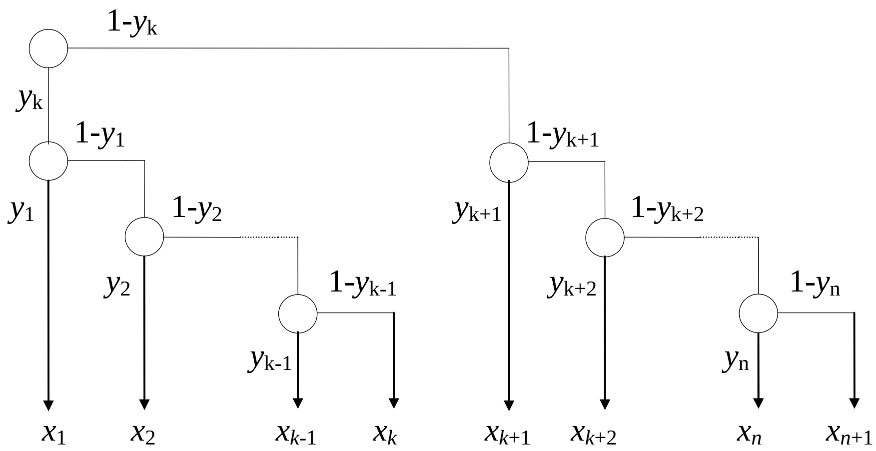

Figure 3.3. Double-cascaded bifurcation topology underlying the Dirichlet type B distribution (taken from Krzysztofowicz and Reese, 1993).

The ***Dirichlet type A*** $(AD_{n+1}^{(A)}(\alpha,\lambda))$ distribution results from generalizing the above bifurcation processes to allow arbitrary bifurcations, rather than only cascaded or double-cascaded topologies (see Krzysztofowicz and Reese for details). This generalization gives us the set of all possible bifurcation topologies. An example of a topology that is not cascaded is shown in Figure 3.4.

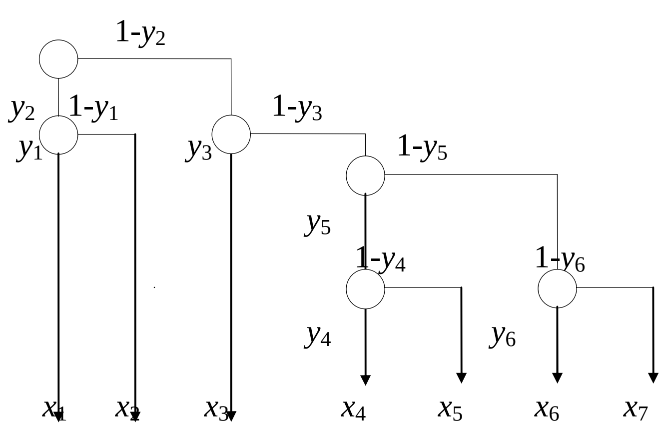

Figure 3.4. A basic bifurcation topology for seven fractions (taken from Krzysztofowicz and Reese, 1993).

Krzysztofowicz and Reese mention that any adaptive Dirichlet distribution can be completely characterized by: 1) its bifurcation topology; 2) the permutation of fractions; and 3) the parameters of the distributions for the ratios y_i. For more details, see their paper.

In what follows, we will give for the Dirichlet type B distribution: the one-to-one correspondence between the fractions and the ratios; the density function; and the moments, covariances, and correlation sign structure in closed form. We will also define all terminology needed for this purpose. All of these results are due to Krzysztofowicz and Reese.

Let $T_d(n+1)$ be the set of all double-cascaded topologies with n+1 fractions. Within this set, a particular topology is identified by the parameter k, which denotes the dividing line between the left and right cascades, where $1 \le k \le (n+1)/2$. Let $t_d(n+1, k)$ denote a topology from $T_d(n+1)$. Assuming that the bifurcation process is independent, Krzysztofowicz and Reese construct expressions for the densities and moments of the fractions x_i for all topologies in $T_d(n+1)$. The resulting expressions are given below. First we define some additional notation. In particular, let

$$\begin{aligned} \tau_i &= 1 && \text{if } i \le k, \\ &= k+1 && \text{if } i > k, \end{aligned} \tag{3.13}$$

and

$$\begin{aligned} K_i &= y_k && \text{if } i < k, \\ &= 1 && \text{if } i = k, \\ &= 1 - y_k && \text{if } i > k, \end{aligned} \tag{3.14}$$

where $y_k = x_1 + x_2 + \ldots + x_k$, and $y_{n+1} = 1$. Finally, we adopt the convention that $\prod_{i=a}^{b} w_i = 1$ if $a > b$.

Moments. A fraction is given by

$$x_i = K_i y_i \prod_{m=\tau_i}^{i-1} (1 - y_m). \tag{3.15}$$

The moments of the x_i can be computed by noting that $y_i \sim Be(\alpha_i, \lambda_i)$ for all *i*, and that the y_i are independent.

Covariances. To compute the covariance of x_i and x_j, we first define

$$Q_j = -\mathrm{var}(y_k)\mathrm{E}(y_j) \prod_{n=k+1}^{j-1} \mathrm{E}(1 - y_n), \quad j = k+1, \ldots, n+1 \tag{3.16}$$

and

$$R_i = [\mathrm{var}(K_i) + \mathrm{E}^2(K_i)][\mathrm{E}(y_i)\mathrm{E}(1 - y_i) - \mathrm{var}(y_i)] \prod_{m=\tau_i}^{i-1} [\mathrm{var}(y_m) + \mathrm{E}^2(1 - y_m)] \cdot$$

$$- \mathrm{E}^2(K_i)\mathrm{E}(y_i)\mathrm{E}(1 - y_i) \prod_{m=\tau_i}^{i-1} \mathrm{E}^2(1 - y_m), \; i = 1, \ldots, k-1, k+1, \ldots, n. \tag{3.17}$$

The covariance is then given by:

$$\mathrm{cov}(x_i, x_j) = \begin{cases} Q_j \mathrm{E}(y_i) \prod_{m=1}^{i-1} \mathrm{E}(1 - y_m) & \text{if } i < k < j \\ Q_j \prod_{m=1}^{i-1} \mathrm{E}(1 - y_m) & \text{if } i = k < j \\ R_i \prod_{n=i+1}^{j-1} \mathrm{E}(1 - y_n) & \text{if } i < k = j \\ R_i \mathrm{E}(y_j) \prod_{n=i+1}^{j-1} \mathrm{E}(1 - y_n) & \text{if } i < j < k \text{ or } k < i < j. \end{cases} \tag{3.18}$$

Equations (3.19) and (3.20) below (due to Krzysztofowicz and Reese) give additional insight into the structure of the covariance matrix:

$$\mathrm{cov}(x_i, x_{j+1}) = \begin{cases} \mathrm{cov}(x_i, x_j)\dfrac{Q_{j+1}}{Q_j} & if\ \ i \le k < j \\ \\ \mathrm{cov}(x_i, x_j)\mathrm{E}(y_{j+1})\dfrac{\mathrm{E}(1-y_j)}{\mathrm{E}(y_j)} & if\, i < j < k \quad or \quad k < i < j, \end{cases} \tag{3.19}$$

and

$$\mathrm{cov}(x_{i+1}, x_j) = \mathrm{cov}(x_i, x_j)\frac{R_{i+1}}{R_i \mathrm{E}(1-y_{i+1})}, \quad if\ \ i < k = j. \tag{3.20}$$

Density. Since we have $y_i \sim Be(\alpha_i, \lambda_i)$ $(i=1,...,n)$ and the y_i are independent, the density function of any Dirichlet type B distribution is of the form

$$g(\mathbf{x}) = A\left[x_k^{\lambda_{k-1}-1}\prod_{i=1}^{k-1} x_i^{\alpha_i-1}(\sum_{j=i}^{k} x_j)^{\lambda_{i-1}-\alpha_i-\lambda_i}\right]\left[x_{n+1}^{\lambda_n-1}\prod_{i=k+1}^{n} x_i^{\alpha_i-1}(\sum_{j=i}^{n+1} x_j)^{\lambda_{i-1}-\alpha_i-\lambda_i}\right], \tag{3.21}$$

where $A = \prod_{i=1}^{n} B^{-1}(\alpha_i, \lambda_i)$ and $\lambda_0 = \alpha_k$.

Krzysztofowicz and Reese also discussed the correlation sign structure of the Dirichlet type B distribution. As they noted, all terms in the expressions for covariances in equation (3.18) are positive except for the Q_j, which are always negative, and the R_i, which may be either negative or positive. Thus, the upper triangular correlation matrix shown in Figure 3.5 has three regions: a rectangular region (where $i \le k < j$) in which all correlations are negative; and two triangular regions in which the signs of the correlations may be either negative or positive. Note that all correlations in any given row of the triangular regions must have the same sign, but the signs may vary from row to row.

Krzysztofowicz and Reese note that, unlike the Dirichlet and Connor-Mosiman distributions, which do not allow the correlation signs to vary within a row, the AD family includes distributions that do so. This is clear from Figure 3.5. Thus, the AD family of distributions is an improvement over the Dirichlet and Connor-Mosiman distributions in this regard.

One of the assumptions made in constructing the adaptive Dirichlet distribution is that all the ratios are independent. In chapter 5, the need for adaptive Dirichlet distributions with dependent ratios is demonstrated via several examples. In addition, we will show that the adaptive Dirichlet with dependent ratios is an improvement over the adaptive Dirichlet with independent ratios in the sense that it allows more general sign structures for the correlation matrix. In particular, this new class will relax the constraints on the correlation structure of fractions in the central rectangular region of Figure 3.5, where $i \leq k < j$; i.e., it will allow the correlation signs to vary within each row of this region. Finally, we will propose some feasible approaches for constructing distributions in this new class.

In section 3.2 we showed that the adaptive Dirichlet is a special case of the multiple Dickey distribution. Thus, the multiple Dickey family of distributions is clearly quite general. This also answers the question posed by Smith (1994) about whether multiple Dickey distributions can admit positive covariance.

3.4 Proposed extensions

In chapters 5 and 6 we propose several extensions to the multiple Dickey family of distributions. First, we will consider the case of bifurcation topologies, but will relax the assumption that the ratios y_i must be independent, giving rise to adaptive Dirichlet distributions with dependent ratios. Next, we will consider multi-furcation topologies, and will relax the assumption that the vectors of ratios must all be Dirichlet distributed. Finally, we will use the multiple Dickey distribution to construct a new family of distributions on the positive orthant. The motivation for these extensions will also be discussed.

$$
\begin{array}{ccccc|ccccc|c}
2 & 3 & \cdots & k-1 & k & k+1 & \cdots & n-1 & n & n+1 & \\
\hline
s_1 & s_1 & \cdots & s_1 & s_1 & - & \cdots & - & - & - & 1 \\
 & s_2 & \cdots & s_2 & s_2 & - & \cdots & - & - & - & 2 \\
 & & & s_3 & s_3 & - & \cdots & - & - & - & 3 \\
 & & & & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots \\
 & & & & s_{k-1} & - & \cdots & - & - & - & k-1 \\
 & & & & & - & \cdots & - & - & - & k \\
\hline
 & & & & & & & s_{k+1} & s_{k+1} & s_{k+1} & k+1 \\
 & & & & & & & & \vdots & \vdots & \vdots \\
 & & & & & & & & s_{n-1} & s_{n-1} & n-1 \\
 & & & & & & & & & s_n & n
\end{array}
$$

Figure 3.5. Correlation sign structure of the adaptive Dirichlet type B distribution resulting from a double-cascaded bifurcation topology, where $s_i \in \{-,+\}$ for $i = 1,\ldots,k-1,k+1,\ldots,n$.