# Chapter 9

## Simulation results

This chapter documents a simulation study comparing the performance of the adaptive Dirichlet distribution with and without dependent ratios at fitting data generated using a given correlation structure. The results of this simulation will help to determine whether the dependent ratios model in fact performs better than the independent ratios model when we know that the data were generated using dependence.

In our simulation, we consider four fractions ($x_1$, $x_2$, $x_3$, and $x_4$), where of course $x_1+ x_2+ x_3+ x_4=1$, with the topology given by figure 9.1.

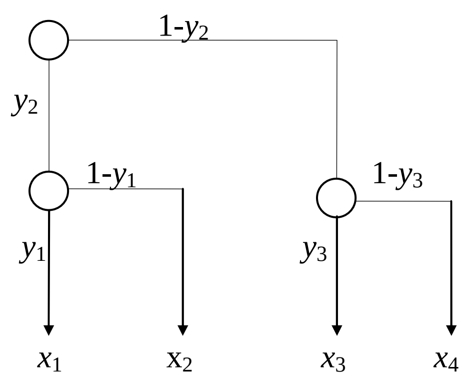


Figure 9.1 Double-cascaded bifurcation topology of four fractions (taken from Krzysztofowicz and Reese, 1991).

We have

$x_1 = y_1 y_2$,

$x_2 = (1 - y_1) y_2$,

$x_3 = (1 - y_2) y_3$, and (9.1)

$x_4 = (1 - y_2)(1 - y_3)$.

We consider four models: the independent ratios model; and models where $y_1$ and $y_2$, $y_1$ and $y_3$, or $y_2$ and $y_3$ are dependent, respectively. To create dependence between the ratios, we generate vectors of four independent beta random variates, $z_1$, $z_2$, $z_3$, and $z_4$. For each simulation run, a sequence of 54 such vectors is generated, because this is the average number of years of the river data analyzed in sections 8.4-8.5. Then, for example, in the model where $y_1$ and $y_3$ are positively correlated, we can take $y_1 = z_1 z_4$, $y_3 = z_3 z_4$, and $y_2 = z_2$. Similarly, for the model where $y_1$ and $y_3$ are negatively correlated, we can take $y_1 = 1 - z_1 z_4$, $y_3 = z_3 z_4$, and $y_2 = z_2$ (although for reasons of time and space the case where the ratios are negatively correlated is not actually considered in this chapter). Via equations (9.1) we can then compute the fractions $x_i$, $i$=1,2,3,4, from the ratios $y_i$, $i$=1,2,3. In order to systematically explore the space of possible correlation structures, we attempted to select the parameters of the beta distributions for the $z_i$ in the simulation runs in such a way that each possible correlation sign structure known to be achievable either with $\text{cov}(y_1,y_3) > 0$ or in the independent case is represented by at least one simulation run. We also considered four cases with $\text{cov}(y_1,y_2) > 0$, and four cases with $\text{cov}(y_2,y_3) > 0$.

Note, however, that this simulation should be regarded as purely exploratory, and does not constitute a thorough evaluation of the small-sample properties of our distribution and fitting method. Such an investigation would need to include multiple simulation runs for each set of parameters (to provide statistically significant results), and also data sets shorter and longer than 54 vectors (to investigate the effects of sample size on performance).

In chapter 5 we showed that the correlation sign structures obtainable from Figure 9.1 when $y_2$ and $(y_1, y_3)$ are independent and $\text{cov}(y_1,y_3) > 0$ are of the form

$$\begin{array}{ccc} \pm & \pm & - \\ & - & \pm \\ & & \pm \end{array}$$

We were able to construct examples for 15 of these 16 cases (but were unsure whether the 16th case is possible, as discussed in chapter 5). These 15 sign structures are reproduced in Table 9-1, along with an indication of which simulated data set (from Appendix B) corresponds to each sign structure.

In the case where $y_3$ and $(y_1, y_2)$ are independent and $\mathrm{cov}(y_1,y_2) > 0$, it can be shown that the correlation sign structures obtainable from Figure 9.1 are restricted to some subset of the following forms:

$$\begin{array}{ccc} + & - & - \\ & - & \pm \\ & & \pm \end{array} \qquad \begin{array}{ccc} + & - & - \\ & \pm & - \\ & & \pm \end{array} \qquad \begin{array}{ccc} - & - & - \\ & \pm & \pm \\ & & \pm \end{array}$$

Table 9-2 gives four of these sign structures, along with an indication of which simulated data set (from Appendix C) corresponds to each sign structure.

Similarly, in the case where $y_1$ and $(y_2, y_3)$ are independent and $\mathrm{cov}(y_2,y_3) > 0$, the correlation sign structures obtainable from Figure 9.1 are restricted to some subset of the following forms:

$$\begin{array}{ccc} + & + & - \\ & + & - \\ & & - \end{array} \qquad \begin{array}{ccc} - & + & \pm \\ & + & \pm \\ & & - \end{array} \qquad \begin{array}{ccc} \pm & - & \pm \\ & - & \pm \\ & & \pm \end{array}$$

Again, Table 9-3 gives four of these sign structures, along with an indication of which simulated data set (from Appendix D) corresponds to each one.

Table 9-1

Correlation sign structures included in Appendix B for the case when $y_2$ and $(y_1, y_3)$ are independent and $\mathrm{cov}(y_1,y_3) > 0$

$$
\overbrace{\begin{array}{ccc} + & - & - \\ & - & - \\ & & + \end{array}}^{\text{Table B−1}}
\quad
\overbrace{\begin{array}{ccc} + & + & - \\ & - & - \\ & & - \end{array}}^{\text{Table B−2}}
\quad
\overbrace{\begin{array}{ccc} + & - & - \\ & - & - \\ & & - \end{array}}^{\text{Table B−3}}
\quad
\overbrace{\begin{array}{ccc} - & + & - \\ & - & - \\ & & + \end{array}}^{\text{Table B−4}}
\quad
\overbrace{\begin{array}{ccc} + & + & - \\ & - & - \\ & & + \end{array}}^{\text{Table B−5}}
\quad
\overbrace{\begin{array}{ccc} - & + & - \\ & - & + \\ & & - \end{array}}^{\text{Table B−6}}
$$

$$
\overbrace{\begin{array}{ccc} - & + & - \\ & - & - \\ & & - \end{array}}^{\text{Table B−7}}
\quad
\overbrace{\begin{array}{ccc} - & - & - \\ & - & - \\ & & + \end{array}}^{\text{Table B−8}}
\quad
\overbrace{\begin{array}{ccc} - & - & - \\ & - & - \\ & & - \end{array}}^{\text{Table B−9}}
\quad
\overbrace{\begin{array}{ccc} - & - & - \\ & - & + \\ & & - \end{array}}^{\text{Table B−10}}
\quad
\overbrace{\begin{array}{ccc} + & - & - \\ & - & + \\ & & + \end{array}}^{\text{Table B−11}}
\quad
\overbrace{\begin{array}{ccc} - & + & - \\ & - & + \\ & & + \end{array}}^{\text{Table B−12}}
$$

$$
\overbrace{\begin{array}{ccc} + & + & - \\ & - & + \\ & & - \end{array}}^{\text{Table B−13}}
\quad
\overbrace{\begin{array}{ccc} - & - & - \\ & - & + \\ & & + \end{array}}^{\text{Table B-14}}
\quad
\overbrace{\begin{array}{ccc} + & - & - \\ & - & + \\ & & - \end{array}}^{\text{Table B-15}}
$$

Table 9-2

Correlation sign structures included in Appendix C for the case when $y_3$ and $(y_1, y_2)$ are independent and $\mathrm{cov}(y_1,y_2) > 0$

$$
\overbrace{\begin{array}{ccc} + & - & - \\ & - & - \\ & & - \end{array}}^{\textit{Table}\text{ C−1}}
\quad
\overbrace{\begin{array}{ccc} - & - & - \\ & - & - \\ & & + \end{array}}^{\textit{Table}\text{ C−2}}
\quad
\overbrace{\begin{array}{ccc} - & - & - \\ & - & - \\ & & - \end{array}}^{\textit{Table}\text{ C−3}}
\quad
\overbrace{\begin{array}{ccc} - & - & - \\ & + & + \\ & & + \end{array}}^{\textit{Table}\text{ C−4}}
$$

Table 9-3

Correlation sign structures included in Appendix D for the case when $y_1$ and $(y_2, y_3)$ are independent and cov($y_2$,$y_3$) > 0

| *Table* D–1 | | | *Table* D–2 | | | *Table* D–3 | | | *Table* D–4 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| − | + | − | + | + | − | + | − | − | + | − | − |
| | + | − | | + | − | | − | − | | − | − |
| | | − | | | − | | | + | | | − |

Finally, in the case where all of the ratios $y_i$ are independent, there are only four possible sign structures. Those are summarized in Table 9-4, and the simulated data sets corresponding to these sign structures are given in Appendix E.

Table 9-4

Correlation sign structures obtainable from Figure 9.1 when the $y_i$, $i$=1,2,3, are independent

| *Table* E–1 | | | *Table* E–2 | | | *Table* E–3 | | | *Table* E–4 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| − | − | − | + | − | − | − | − | − | + | − | − |
| | − | − | | − | − | | − | − | | − | − |
| | | − | | | − | | | + | | | + |