Urn-Based Credit Risk Models for Portfolios of Dependent Risks

Riccardo Gusso Uwe Schmock Daniel Seiler

15 dicembre 2005

Introduction

- One of the main problems in credit portfolio management is the occurrence of many joint defaults of different counterparties over a fixed time horizon T.
- In the measurement of the expected credit loss of a portfolio it is then very important to take into account the dependence between individual risks.
- In this paper we present two models for several groups of firms that express dependence in terms of the *ratings* of the considered firms, in a monotone way.
- This means that defaults in the better ratings have a contagion effect in defaults in the worser ratings. Moreover there is also a dependence effect between defaults of firms in the same rating class.

Theory of Exchangeable Sequences

In the following, $(\Omega, \mathcal{F}, \mathbb{P})$ denotes a probability space with σ-algebra $\mathcal{F}$ and probability measure $\mathbb{P}$. The X_i's and $\mathbf{X}_i$'s denote respectively random variables and random vectors on this space.

The finite set $(X_1, X_2, \ldots, X_n)$ of random variables is said to be *exchangeable* if the joint distribution is invariant under all n-permutations:

$$(X_1, X_2, \ldots, X_n) \stackrel{d}{=} (X_{\pi(1)}, X_{\pi(2)}, \ldots, X_{\pi(n)}), \tag{1}$$

for every permutation π of $(1, 2, \ldots, n)$.

An infinite sequence of random variables $(X_n)_{n\geq 1}$ is said to be *exchangeable* if $(X_1, X_2, \ldots, X_n)$ is exchangeable for each $n \geq 2$.

Theory of Exchangeable Sequences

Theorem 1 (De Finetti Theorem). *Let $(\mathbf{X}_n)_{n\geq 1}$ be an exchangeable sequence of random vectors on $(\Omega, \mathcal{F}, \mathbb{P})$. Then there exists a sub-σ-algebra of $\mathcal{F}$ conditioned on which the $\mathbf{X}_n$ are independent and identically distributed.*

Corollary 1. *Let $(\mathbf{X}_n)_{n\geq 1}$ be an exchangeable sequence where the $\mathbf{X}_i$'s take only values in $\{\mathbf{e}_1, \dots, \mathbf{e}_d\} \subseteq \mathbb{R}^d$. Then there exists a random vector $(P_1, \dots, P_d)$ taking values in $\Delta^d = \{(p_1, \dots, p_d) \in [0,1]^d | \sum_{j=1}^d p_j = 1\}$ such that:*

1. *for every $\mathbf{l} \in \mathbb{N}_0^d$ such that $\sum_{j=1}^d l_j = n$ it is*

$$\mathbb{P}\left[\sum_{i=1}^n \mathbf{X}_i = \mathbf{l} \middle| P_1, \dots, P_d\right] \overset{\text{a.s.}}{=} \frac{n!}{(l_1!)(l_2!)\cdots(l_d!)} P_1^{l_1} P_2^{l_2} \cdots P_d^{l_d}$$

2. *for $1 \leq j \leq d$,*

$$P_j \overset{\text{a.s.}}{=} \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^n X_{i,j}.$$

Multidimensional Scheme for Defaults

- We are given k homogeneous groups of n_i , $1 \geq i \geq k$ companies, with credit ratings $r_1 \succ r_2 \succ \cdots \succ r_k$.
- Consider an urn which contains $b_j > 0$ balls of different colours for every $1 \leq j \leq k+1$.
- Draw randomly a ball from the urn: if its rating is r_j, and the ball has a colour from $1, \ldots, j$ then the firm defaults. If the colour is from $j+1, \ldots, k+1$ it does not default.
- Return the ball in the urn together with other $c > 0$ balls of the same colour.
- The random vector $\mathbf{X}_n = (X_{n,1}, \ldots X_{n,k+1})^\intercal \in \{0,1\}^{k+1}$ indicating the colour of the ball of the nth draw is defined in the following way:

$$X_{n,j} := \begin{cases} 1 & \text{if the } n\text{th ball drawn has colour j} \\ 0 & \text{otherwise.} \end{cases}$$

- Then the vector $\mathbf{P}_n = (P_{n,1}, \ldots, P_{n,k+1})^\intercal$ of the conditional probabilities that the nth ball drawn has colour $1 \leq j \leq k+1$ given the previous draws is:

$$P_{n,j} := \mathbb{P}[\mathbf{X}_n = \mathbf{e}_j | \mathbf{X}_1, \ldots, \mathbf{X}_{n-1}].$$

Multidimensional Scheme for Defaults

Proposition 1. *The sequence of random vectors* $(\mathbf{X}_n)_{n\geq 1}$ *is exchangeable: in fact for every* $n \geq 1$ *we have that*

$$\mathbb{P}[\mathbf{X}_1 = \mathbf{e}_{j_1}, \ldots, \mathbf{X}_n = \mathbf{e}_{j_n}] = \frac{\prod_{j=1}^{k+1} \prod_{i=0}^{l_{n,j}-1} (b_j + ic)}{b(b+c)\cdots(b+(n-1)c)} \tag{2}$$

where $\mathbf{l}_n = \sum_{i=1}^{n} \mathbf{e}_{j_i}$.

Proposition 2. *The sequence of random vectors* $(\mathbf{P}_n)_{n\geq 1}$ *is a convergent martingale w.r.t. the filtration* $\mathcal{F}_n = \sigma(\mathbf{X}_1, \ldots, \mathbf{X}_{n-1})$ $(\mathcal{F}_1 = \{\emptyset, \Omega\})$*. In fact for every* $1 \leq j \leq k+1$ *the sequence* $(P_{n,j})_{n\geq 1}$ *is a bounded martingale, and hence almost surely convergent to the limit random variable*

$$P_j \overset{a.s.}{=} \lim_{n\to\infty} P_{n,j} = \lim_{n\to\infty} \frac{b_j + c(\sum_{i=1}^{n-1} X_{i,j})}{b + c(n-1)} = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} X_{i,j}.$$

Multidimensional Scheme for Defaults

- Let N_j denotes the number of defaults within the group of companies with rating r_j, and $n = n_1 + \cdots + n_k$ be the total number of considered firms.
- To determine the joint distribution of the random vector $(N_1, N_2, \ldots, N_k)^\top$, since the sequence $(\mathbf{X}_n)_{n\geq 1}$ is exchangeable, it does not matter in which order we draw the ball for the companies, hence we can choose a special order to facilitate calculations; that is we first consider the firms with the best rating, than the next lower, and so on.
- We can use then corollary 1 to compute the distribution:

Multidimensional Scheme for Defaults

$$
\begin{aligned}
&\mathbb{P}[N_1 = l_1, \dots, N_k = l_k] \\
&= \mathbb{P}\left[\sum_{i=1}^{n_1} X_{i,1} = l_1, \dots, \sum_{i=n_1+\cdots+n_{k-1}+1}^{n_1+\cdots+n_k} (X_{i,1} + \cdots + X_{i,k}) = l_k\right] \\
&= \mathbb{E}\left[\mathbb{P}\left[\sum_{i=1}^{n_1} X_{i,1} = l_1, \sum_{i=n_1+\cdots+n_{k-1}+1}^{n_1+\cdots+n_k} (X_{i,1} + \cdots + X_{i,k}) = l_k \middle| P_1, \dots, P_{k+1}\right]\right] \\
&= \mathbb{E}\left[\mathbb{P}\left[\sum_{i=1}^{n_1} X_{i,1} = l_1 \middle| P_1, \dots, P_{k+1}\right] \mathbb{P}\left[\sum_{i=n_1+1}^{n_1+n_2} (X_{i,1} + X_{i,2}) = l_2 \middle| P_1, \dots, P_{k+1}\right]\right. \\
&\qquad \left.\cdots \mathbb{P}\left[\sum_{i=n_1+\cdots+n_{k-1}+1}^{n_1+\cdots+n_k} (X_{i,1} + \cdots + X_{i,k}) = l_k \middle| P_1, \dots, P_{k+1}\right]\right] \\
&= \mathbb{E}\left[\binom{n_1}{l_1} P_1^{l_1} (1-P_1)^{n_1-l_1} \binom{n_2}{l_2} (P_1+P_2)^{l_2} (1-P_1-P_2)^{n_2-l_2}\right. \\
&\qquad \left.\cdots \binom{n_k}{l_k} (P_1+\cdots+P_k)^{l_k} (1-P_1-\cdots-P_k)^{n_k-l_k}\right]
\end{aligned}
\tag{3}
$$

Multidimensional Scheme for Defaults

- We see then that this model belongs to class of Bernoulli mixture models:

  Definition 1. The random vector $\mathbf{Y}$ follows a *Bernoulli mixture model with factor vector* $\boldsymbol{\Psi}$ if there exists a random vector $\boldsymbol{\Psi} = (\Psi_1, \ldots, \Psi_p)$ and functions $Q_i : \mathbb{R} \to [0, 1]$ such that conditional on $\boldsymbol{\Psi}$ the vector $\mathbf{Y}$ is a vector of independent Bernoulli random variables with default probabilities $\mathbb{P}[Y_i = 1|\boldsymbol{\Psi} = \boldsymbol{\psi}] = Q_i(\boldsymbol{\psi})$.

- It follows immediately that for $\mathbf{y} = (y_1, \ldots, y_n)^\top \in \{0, 1\}^n$

$$\mathbb{P}[\mathbf{Y} = \mathbf{y}|\boldsymbol{\Psi} = \boldsymbol{\psi}] = \prod_{i=1}^{n} Q_i(\boldsymbol{\psi})^{y_i}(1 - Q_i(\boldsymbol{\psi}))^{1-y_i} \tag{4}$$

  and to obtain the unconditional joint probability we have to take the expectation w.r.t. the df of $\boldsymbol{\Psi}$.

Multidimensional Scheme for Defaults

- It is then necessary to compute the distribution of $(P_1, \ldots, P_{k+1})$. Using previous results we can compute generalized moments:

$$
\begin{aligned}
&\mathbb{E}[P_1^{l_1} P_2^{l_2} \cdots P_{k+1}^{l_{k+1}}] \\
&= \mathbb{E}\left[\mathbb{P}\left[\bigcap_{j=1}^{k+1}\left(\bigcap_{i=l_0+\cdots+l_{j-1}+1}^{l_0+\cdots+l_j}\{\mathbf{X}_i = \mathbf{e}_j\}\right)\Big| P_1, \ldots, P_{k+1}\right]\right] \\
&= \mathbb{P}\left[\left(\bigcap_{i=1}^{l_1}\{\mathbf{X}_i = \mathbf{e}_1\}\right)\bigcap\cdots\bigcap\left(\bigcap_{i=l_1+\cdots+l_k+1}^{l_1+\cdots+l_{k+1}}\{\mathbf{X}_i = \mathbf{e}_{k+1}\}\right)\right] \\
&= \frac{\prod_{j=1}^{k+1}\prod_{i=0}^{l_{n,j}-1}(b_j + ic)}{b(b+c)\cdots(b+(n-1)c)} = \frac{\Gamma(\frac{b}{c})}{\Gamma(\frac{b}{c}+n)}\prod_{j=1}^{k+1}\frac{\Gamma\left(\frac{b_j}{c}+l_j\right)}{\Gamma\left(\frac{b_j}{c}\right)} \\
&= \frac{\Gamma(\alpha)}{\Gamma(\alpha+n)}\prod_{j=1}^{k+1}\frac{\Gamma(\alpha_j+l_j)}{\Gamma(\alpha_j)} \qquad \alpha_j = \frac{b_j}{c}, \alpha = \sum_j \alpha_j
\end{aligned}
\tag{5}
$$

Multidimensional Scheme for Defaults

- Since the random vector $(P_1, \ldots, P_{k+1})$ has bounded support its moments determine the distribution function, so it has to be Dirichlet distributed $D_{k+1}(\alpha_1, \ldots, \alpha_{k+1})$ with density $f_{k+1}(p_1, \ldots, p_{k+1}) = \frac{\Gamma\left(\sum_{j=1}^{k+1} \alpha_j\right)}{\prod_{j=1}^{k+1} \Gamma(\alpha_j)} \prod_{j=1}^{k+1} p_j^{\alpha_j - 1}$, with the constraint $\sum_{j=1}^{k+1} p_j = 1$ and parameters $\alpha_1, \ldots, \alpha_{k+1} > 0$.
- We can then compute explicitely the joint distribution, using the fact that for $0 < a < 1$ and $s, t > 0$, and $m \in \mathbb{N}_0$

$$\begin{aligned} &\int_0^{1-a} p^{s-1}(a+p)^m(1-a-p)^{t-1} dp \\ &= \sum_{j=0}^{m} \binom{m}{j} a^{m-j}(1-a)^{s+j+t-1} B(s+j, t). \end{aligned} \tag{6}$$

where $B(x, y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$ and $\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} d\,t$.

Multidimensional Scheme for Defaults

$$
\begin{aligned}
&\mathbb{P}[N_1 = l_1, \ldots, N_k = l_k] \\
&= \mathbb{E}\Bigg[\binom{n_1}{l_1} P_1^{l_1}(1-P_1)^{n_1-l_1}\binom{n_2}{l_2}(P_1+P_2)^{l_2}(1-P_1-P_2)^{n_2-l_2}\cdots \\
&\quad \cdots \binom{n_k}{l_k}(P_1+\cdots+P_k)^{l_k}(1-P_1-\cdots-P_k)^{n_k-l_k}\Bigg] \\
&= \binom{n_1}{l_1}\cdots\binom{n_k}{l_k}\int_0^1\int_0^{1-p_1}\ldots\int_0^{1-p_1-\cdots-p_{k-1}} p_1^{l_1}(1-p_1)^{n_1-l_1} \\
&\quad \times\cdots(p_1+p_2+\ldots p_k)^{l_k}(1-p_1-\cdots-p_k)^{n_k-l_k} \\
&\quad \times \frac{\Gamma(\sum_{j=1}^{k+1}\alpha_j)}{\prod_{j=1}^{k+1}\Gamma(\alpha_j)} p_1^{\alpha_1-1}p_2^{\alpha_2-1}\cdots p_k^{\alpha_k-1}(1-p_1-\cdots-p_k)^{\alpha_{k+1}-1}dp_k dp_{k-1}\cdots dp_1
\end{aligned}
$$

Multidimensional Scheme for Defaults

$$
\begin{aligned}
&= \binom{n_1}{l_1} \cdots \binom{n_k}{l_k} \frac{\Gamma(\sum_{j=1}^{k+1} \alpha_j)}{\prod_{j=1}^{k+1} \Gamma(\alpha_j)} \sum_{j_1=0}^{l_k} \binom{l_k}{j_1} B(\alpha_k + j_1, \alpha_{k+1} + n_k - l_k) \\
&\times \sum_{j_2=0}^{l_{k-1}+l_k-j_1} \binom{l_{k-1} + l_k - j_1}{j_2} B(\alpha_{k-1} + j_2, \alpha_k + \alpha_{k+1} + n_{k-1} + n_k - l_{k-1} - l_k + j_1) \\
&\cdots \times \sum_{j_{k-1}=0}^{\substack{l_2+\cdots+l_k \\ -j_1-\cdots-j_{k-2}}} \binom{l_2 + \cdots + l_k - j_1 - \cdots - j_{k-2}}{j_{k-1}} \\
&\times B(\alpha_2 + j_{k-1}, \alpha_3 + \cdots + \alpha_{k+1} + n_2 + \cdots + n_k - l_2 - \cdots - l_k + j_1 + \cdots + j_{k-2}) \\
&\times B(\alpha_1 + l_1 + \cdots + l_k - j_1 - \cdots - j_{k-1}, \alpha_2 + \cdots + \alpha_{k+1} \\
&\qquad + n_1 + \cdots + n_k - l_1 - \cdots - l_k + j_1 + \cdots + j_{k-1}).
\end{aligned} \tag{7}
$$

Multidimensional Scheme for Defaults

It is possible to give a nice recursive representation of this joint probability:

Proposition 3. *Given $k \in \mathbb{N}$ rating classes with $n_j \in \mathbb{N}_0$ companies and $l_j \in \mathbb{N}$ defaults for $j \in \{1, \ldots, k\}$, and a Dirichlet distributed random vector $(P_1, \ldots, P_k)$ with parameters $\alpha_1, \ldots, \alpha_{k+1}$ we have $\mathbb{P}[N_1 = l_1, \ldots, N_k = l_k] = f(\alpha_{k+1}, k, 0, 0)$ where for $\beta > 0, n \in \mathbb{N}_0, l \in \{0, \ldots, n\}$ we define*

$$f(\beta, 1, l, n) = \binom{n_1}{l_1} \frac{B(\alpha_1 + l_1 + l, \beta + n_1 - l_1 + n - l)}{B(\alpha_1, \beta)} \tag{8}$$

and recursively for $j \in \{2, \ldots, k\}$

$$f(\beta, j, l, n) = \binom{n_j}{l_j} \sum_{i=0}^{l_j + l} \binom{l_j + l}{i} \frac{B(\alpha_j + i, \beta + n_j - l_j + n - l)}{B(\alpha_j, \beta)} \times f(\alpha_j + \beta, j - 1, l_j + l - i, n_j + n) \tag{9}$$

Iterative Scheme for Defaults

- Same notation as for the previous scheme.
- The number of defaults in the best rating group r_1 is determined with a (Pólya) unidimensional urn scheme.
- The number of defaults in the worser ratings are then determined by the number of firms that would have defaulted in the next better rating plus a certain part of the group that would have survived in the next better rating, and this additional part is determined again via unidimensional urn scheme.
- Formally: take independent beta distributions $\tilde{P}_1 \sim beta(\alpha_1, \beta_1), \ldots, \tilde{P}_k \sim beta(\alpha_k, \beta_k)$ and define the random default frequencies in the following way:
- $P_1 = \tilde{P}_1$
- $P_2 = P_1 + (1 - P_1)\tilde{P}_2$
- $\ldots$
- $P_k = P_{k-1} + (1 - P_{k-1})\tilde{P}_k$

Iterative Scheme for Defaults

- Using the (random) default frequencies we can compute the joint distribution of the number of defaults.
- For example for the case of only two rating groups we have:

$$\mathbb{P}[N_1 = l_1, N_2 = l_2] = \mathbb{E}[\mathbb{P}[N_1 = l_1, N_2 = l_2 | P_1, P_2]]$$

$$= \mathbb{E}\left[\binom{n_1}{l_1} P_1^{l_1} (1 - P_1)^{n_1 - l_1} \binom{n_2}{l_2} P_2^{l_2} (1 - P_2)^{n_2 - l_2}\right]$$

$$= \binom{n_1}{l_1}\binom{n_2}{l_2}$$

$$\times \mathbb{E}\left[\tilde{P}_1^{l_1} (1 - \tilde{P}_1)^{n_1 - l_1} (\tilde{P}_1 + (1 - \tilde{P}_1)\tilde{P}_2)^{l_2} ((1 - \tilde{P}_1)(1 - \tilde{P}_2))^{n_2 - l_2}\right]$$

- Using then binomial expansion and the fact that $\tilde{P}_j$'s are independent, we can obtain an explicit expression.

Iterative Scheme for Defaults

- In the general case we have:

$$\begin{aligned}
&\mathbb{P}[N_1 = l_1, \ldots, N_k = l_k] \\
&= \binom{n_1}{l_1} \cdots \binom{n_k}{l_k} \frac{1}{B(\alpha_1, \beta_1) \cdots B(\alpha_k, \beta_k)} \times \sum_{j_1=0}^{l_k} \binom{l_k}{j_1} B(\alpha_k + l_k - j_1, \beta_k + n_k - l_k) \\
&\cdots \times \sum_{j_i=0}^{l_{k+1-i}+j_{i-1}} \binom{l_{k+1-i} + j_{i-1}}{j_i} B(\alpha_{k+1-i} + l_{k+1-i} + j_{i-i} - j_i, \\
&\qquad\qquad \beta_{k+1-i} + n_k + \cdots + n_{k+1-i} - l_{k+1-i} - j_{i-1}) \\
&\cdots \times \sum_{j_{k-1}=0}^{l_2+j_{k-2}} \binom{l_2 + j_{k-2}}{j_{k-1}} B(\alpha_2 + l_2 + j_{k-2} - j_{k-1}, \beta_2 + n_k + \cdots + n_2 - l_2 - j_{k-2}) \\
&\qquad\qquad \times B(\alpha_1 + l_1 + j_{k-1}, \beta_1 + n_k + \cdots + n_1 - l_1 - j_{k-1})
\end{aligned} \tag{10}$$

Iterative Scheme for Defaults

- Introducing the new random variables $Q_1 = P_1$ and $Q_i = P_i - P_{i-1} = (1 - P_{i-1})\tilde{P}_i$ for $i = 2, \ldots, k$ it is possible to write the default frequencies as $P_i = Q_1 + \cdots + Q_i$.
- Then we have again a mixture representation:

$$\begin{aligned} &\mathbb{P}[N_1 = l_1, \ldots, N_k = l_k] \\ &= \mathbb{E}\Big[\binom{n_1}{l_1} Q_1^{l_1} (1 - Q_1)^{n_1 - l_1} \ldots \\ &\qquad \binom{n_k}{l_k} (Q_1 + \cdots + Q_k)^{l_k} (1 - Q_1 - \cdots - Q_k)^{n_k - l_k}\Big] \end{aligned} \tag{11}$$

- In this case $(Q_1, \ldots, Q_k)$ has a generalized Dirichlet distribution with density:

$$\begin{aligned} f(q_1, \ldots, q_k) &= \frac{1}{\prod_{i=1}^k B(\alpha_i, \beta_i)} (1 - \sum_{i=1}^k q_i)^{\beta_k - 1} \\ &\times \prod_{i=1}^k \left[q_i^{\alpha_i - 1} (1 - \sum_{j=0}^{i-1} q_j)^{\beta_{i-1} - (\alpha_i + \beta_i)} \right]. \end{aligned}$$

Iterative Scheme for Defaults

Again it is possible to give a recursive representation of this joint probability:

Proposition 4. *Given $k \in \mathbb{N}$ rating classes with $n_j \in \mathbb{N}_0$ companies and $l_j \in \mathbb{N}$ defaults for $j \in \{1, \dots, k\}$, and a Generalized Dirichlet distributed random vector $(Q_1, \dots, Q_k)$ with parameters $\alpha_1, \beta_1, \dots, \alpha_k, \beta_k$ we have $\mathbb{P}[N_1 = l_1, \dots, N_k = l_k] = f(k, 0, 0)$ where for $n \in \mathbb{N}_0$, $l \in \{0, \dots, n\}$ we define*

$$f(1, l, n) = \binom{n_1}{l_1} \frac{B(\alpha_1 + l_1 + l, \beta + n_1 - l_1 + n - l)}{B(\alpha_1, \beta_1)} \tag{12}$$

and recursively for $j \in \{2, \dots, k\}$

$$f(j, l, n) = \binom{n_j}{l_j} \sum_{i=0}^{l_j + l} \binom{l_j + l}{i} \frac{B(\alpha_j + l_j + l - i, \beta_j + n_j - l_j + n - l)}{B(\alpha_j, \beta_j)} \times f(j-1, i, n_j + n) \tag{13}$$

EM Algorithm

- The Expectation-Maximization algorithm is a tool for the iterative computation of maximum-likelihood estimates when the observations can be viewed as incomplete data.
- Sample spaces $\mathcal{X}$ and $\mathcal{Y}$, $\mathbf{x} \longrightarrow \mathbf{y}(\mathbf{x})$ mapping.
- Observed data $\mathbf{y} \in \mathcal{Y}$; corresponding unobserved $\mathbf{x} \in \mathcal{X}(\mathbf{y}) = \{\mathbf{x} \mid \mathbf{y}(\mathbf{x}) = y\}$.
- Postulate sampling densities $f(\mathbf{x}|\phi)$ and derive its corresponding $g(\mathbf{y}|\phi)$ through the relation:

$$g(\mathbf{y}|\phi) = \int_{\mathcal{X}(\mathbf{y})} f(\mathbf{x}|\phi) d\mathbf{x}. \tag{14}$$

- EM algorithm tries to find a value of ϕ which maximizes $g(\mathbf{y}|\phi)$ using the associated family $f(\mathbf{x}|\phi)$

EM Algorithm

- Assume $f(\mathbf{x}|\boldsymbol{\phi})$ has the form

$$f(\mathbf{x}|\boldsymbol{\phi}) = b(\mathbf{x}) \exp(\boldsymbol{\phi}^\intercal \mathbf{t}(\mathbf{x}))/a(\boldsymbol{\phi})$$

- Let $\boldsymbol{\phi}^{(p)}$ be the values of the parameters after p iterations; then $p+1$ values are computed in two steps:
- (E-step) estimate $\mathbf{t}^{(p)}$ by:

$$\mathbf{t}^{(p)} = \mathbb{E}[\mathbf{t}(\mathbf{x})|\mathbf{y}, \boldsymbol{\phi}^{(p)}]. \tag{15}$$

- (M-step) determine $\boldsymbol{\phi}^{(p+1)}$ solving the system of equations:

$$\mathbb{E}[\mathbf{t}(\mathbf{x})|\boldsymbol{\phi}] = \mathbf{t}^{(p)}. \tag{16}$$

- Why does the algorithm work?
- Denote with $L(\boldsymbol{\phi}) = \log g(\mathbf{y}|\boldsymbol{\phi})$, the log-likelihood.

- Introduction
- Theory of Exchangeable Sequences
- Multidimensional Scheme for Defaults
- Iterative Scheme for Defaults
- EM Algorithm
- EM Algorithm appl. to Mult. Scheme
- EM Algorithm appl. to Iter. Scheme
- Calibration of models
- Further improvements

EM Algorithm

- The conditional density of $\mathbf{x}$ given $\mathbf{y}$ and $\boldsymbol{\phi}$ is given in this case by:

$$k(\mathbf{x}|\mathbf{y},\boldsymbol{\phi}) = f(\mathbf{x}|\boldsymbol{\phi})/g(\mathbf{y}|\boldsymbol{\phi}) = b(\mathbf{x})\exp(\boldsymbol{\phi}^{\mathsf{T}}\mathbf{t}(\mathbf{x}))/a(\boldsymbol{\phi}|\mathbf{y}) \tag{17}$$

where

$$a(\boldsymbol{\phi}|\mathbf{y}) = \int_{\mathcal{X}(\mathbf{y})} b(\mathbf{x})\exp(\boldsymbol{\phi}^{\mathsf{T}}\mathbf{t}(\mathbf{x}))d\mathbf{x} \tag{18}$$

- We can then rewrite the log-likelihood as:

$$L(\boldsymbol{\phi}) = \log f(\mathbf{x}|\boldsymbol{\phi}) - \log k(\mathbf{x}|\mathbf{y},\boldsymbol{\phi}) = -\log a(\boldsymbol{\phi}) + \log a(\boldsymbol{\phi}|\mathbf{y}) \tag{19}$$

- Then, differentiating (19) we get:

$$\begin{aligned} \mathbf{D}L(\boldsymbol{\phi}) &= -\mathbf{D}\log a(\boldsymbol{\phi}) + \mathbf{D}\log a(\boldsymbol{\phi}|\mathbf{y}) \\ &= -\mathbb{E}[\mathbf{t}(\mathbf{x})|\boldsymbol{\phi}] + \mathbb{E}[\mathbf{t}(\mathbf{x})|\mathbf{y},\boldsymbol{\phi}] \end{aligned} \tag{20}$$

EM Algorithm

- We see then that if the algorithm converges to $\boldsymbol{\phi}^{(p)} = \boldsymbol{\phi}^{(p+1)} = \boldsymbol{\phi}^{*}$, we have then $\mathbb{E}[\mathbf{t}(\mathbf{x})|\boldsymbol{\phi}^{*}] = \mathbb{E}[\mathbf{t}(\mathbf{x})|\mathbf{y}, \boldsymbol{\phi}^{*}]$, so that $\mathbf{D}L(\boldsymbol{\phi}^{*}) = 0$.
- We apply now the algorithm to the two scheme examined.

EM Algorithm appl. to Mult. Scheme

- Observed data $\mathbf{y} = (\mathbf{y}_1, \ldots, \mathbf{y}_m)$, where $\mathbf{y}_i = (l_{i1}, \ldots, l_{ik})$ for $i = 1, \ldots, m$ and m is the total number of observations.
- Complete data $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_m)$, where $\mathbf{x}_i = (p_{i1}, \ldots, p_{ik}, l_{i1}, \ldots, l_{ik})$ with p_{ij} the $i - th$ (unknown) realisation of the random variable P_j
- Goal: MLE of the parameters of $g(\mathbf{y}|\boldsymbol{\alpha}) = \prod_{i=1}^m g_i(\mathbf{y}_i|\boldsymbol{\alpha})$ where g_i is given by equation (7).
- Sampling density: $f(\mathbf{x}|\boldsymbol{\alpha}) = \prod_{i=1}^m f_i(\mathbf{x}_i|\boldsymbol{\alpha})$ where

$$
\begin{aligned}
f_i(\mathbf{x}_i|\boldsymbol{\alpha}) = & \frac{\Gamma(\sum_{j=1}^{k+1} \alpha_j)}{\prod_{j=1}^{k+1} \Gamma(\alpha_j)} p_{i1}^{\alpha_1 - 1} \cdots p_{ik}^{\alpha_k - 1} (1 - p_{i1} - \cdots - p_{ik})^{\alpha_{k+1} - 1} \\
& \times \binom{n_{i1}}{l_{i1}} \cdots \binom{n_{ik}}{l_{ik}} p_{i1}^{l_{i1}} (1 - p_{i1})^{n_{i1} - l_{i1}} \\
& \times \cdots (p_{i1} + \cdots + p_{ik})^{l_{ik}} (1 - p_{i1} - \cdots - p_{ik})^{n_{ik} - l_{ik}}
\end{aligned}
\tag{21}
$$

EM Algorithm appl. to Mult. Scheme

- We can rewrite f as:

$$
\begin{aligned}
f(\mathbf{x}|\boldsymbol{\alpha}) &= \exp\left(\sum_{i=1}^{m} \log f_i(\mathbf{x}_i|\boldsymbol{\alpha})\right) \\
&= \left(\frac{\Gamma(\sum_{j=1}^{k+1}\alpha_j)}{\prod_{j=1}^{k+1}\Gamma(\alpha_j)}\right)^m \exp\left[\sum_{i=1}^{m} b_i(p_{i1},\dots,p_{im},l_{i1},\dots,l_{im})\right] \\
&\times \exp\left[(\alpha_1,\dots,\alpha_{k+1})^\intercal\left(\sum_{i=1}^{m}\log p_{i1},\dots,\sum_{i=1}^{m}\log p_{ik},\sum_{i=1}^{m}\log(1-p_{i1}-\dots-p_{ik})\right)\right]
\end{aligned}
$$

- The statistics of our interest are then $t_j(\mathbf{x}) = \sum_{i=1}^{m}\log p_{ij}$ for $j = 1,\dots,k$ and $t_{k+1}(\mathbf{x}) = \sum_{i=1}^{m}\log(1-p_{i1}-\dots-p_{ik})$.
- E-step: compute

$$
\begin{aligned}
\mathbb{E}[t_j(\mathbf{x})|\mathbf{y},\boldsymbol{\alpha}] &= \sum_{i=1}^{m}\mathbb{E}[\log p_{ij}|\mathbf{y},\boldsymbol{\alpha}] \qquad \text{for} \qquad j=1,\dots,k \\
\mathbb{E}[t_{k+1}(\mathbf{x})|\mathbf{y},\boldsymbol{\alpha}] &= \sum_{i=1}^{m}\mathbb{E}[\log(1-p_{i1}-\dots-p_{ik})|\mathbf{y},\boldsymbol{\alpha}]
\end{aligned} \tag{22}
$$

EM Algorithm appl. to Mult. Scheme

- We have then to compute:

$$\mathbb{E}[\log p_{ij}|\mathbf{y},\boldsymbol{\alpha}] = \\ = \frac{\int_0^1 \int_0^{1-p_{i1}} \cdots \int_0^{1-p_{i1}-\cdots-p_{i,k-1}} \log p_{ij} k_i(\mathbf{x}|\boldsymbol{\alpha}) dp_{ik} \cdots dp_{i1}}{\int_0^1 \int_0^{1-p_{i1}} \cdots \int_0^{1-p_{i1}-\cdots-p_{i,k-1}} k_i(\mathbf{x}|\boldsymbol{\alpha}) dp_{ik} \cdots dp_{i1}} \qquad (23)$$

where

$$k_i(\mathbf{x}|\boldsymbol{\alpha}) = p_{i1}^{\alpha_1 - 1} \cdots (1-p_{i1}-\cdots-p_{ik})^{\alpha_{k+1}-1} p_{i1}^{l_{i1}} \cdots (1-p_{i1}-\cdots-p_{ik})^{n_{ik}-l_{ik}}.$$

EM Algorithm appl. to Mult. Scheme

Key facts (analogous to (6)); defining $D(\alpha,\beta) := \frac{\Gamma'(\alpha)}{\Gamma(\alpha)} - \frac{\Gamma'(\alpha+\beta)}{\Gamma(\alpha+\beta)}$, we have:

$$\int_0^{1-a} \log p \, p^{s-1}(a+p)^m(1-a-p)^{t-1}dp$$
$$= \sum_{j=1}^{m} \binom{m}{j} (1-a)^{s+j+t-1} a^{m-j} B(s+j,t) \left[\log(1-a) + D(s+j,t)\right] \tag{24}$$

$$\int_0^{1-a} \log(1-a-p) \, p^{s-1}(a+p)^m(1-a-p)^{t-1}dp$$
$$= \sum_{j=1}^{m} \binom{m}{j} (1-a)^{s+j+t-1} a^{m-j} B(s+j,t) \left[\log(1-a) + D(t,s+j)\right] \tag{25}$$

EM Algorithm appl. to Mult. Scheme

As an example for $k = 3$ and $j = 1$ the numerator in (23) becomes:

$$C = \sum_{j_1=0}^{l_3} \binom{l_3}{j_1} B(\alpha_3 + j_1, \alpha_4 + n_3 - l_3)$$

$$\times \sum_{j_2=0}^{l_2+l_3-j_1} \binom{l_2 + l_3 - j_1}{j_2} B(\alpha_2 + j_2, \alpha_3 + \alpha_4 + n_2 + n_3 - l_2 - l_3 + j_1)$$

$$\times B(\alpha_1 + l_1 + l_2 + l_3 - j_1 - j_2, \alpha_2 + \alpha_3 + \alpha_4 + n_1 - l_1 + n_2 - l_2 + n_3 - l_3 + j_1 + j_2)$$

$$\times D(\alpha_1 + l_1 + l_2 + l_3 - j_1 - j_2, \alpha_2 + \alpha_3 + \alpha_4 + n_1 - l_1 + n_2 - l_2 + n_3 - l_3 + j_1 + j_2)$$

EM Algorithm appl. to Mult. Scheme

- M-step: compute the unconditional expectations $\mathbb{E}[t_j(\mathbf{x})|\boldsymbol{\alpha}]$.
- Since the statistics don't depend on the variables $(l_{11},\dots,l_{1k},\dots,l_{m1},\dots,l_{mk})$ we have for example that for $1\le j\le k$:

$$\begin{aligned}
\mathbb{E}[t_j(\mathbf{x})|\boldsymbol{\alpha}] &= \sum_{i=1}^{m}\mathbb{E}[\log p_{ij}|\boldsymbol{\alpha}] \\
&= \sum_{i=1}^{m}\int_0^{p_{i1}}\cdots\int_0^{1-p_{i1}-\cdots-p_{i,k-1}}\log p_{ij} \\
&\times\frac{\Gamma\left(\sum_{j=1}^{k+1}\alpha_j\right)}{\prod_{i=1}^{k+1}\Gamma(\alpha_j)}p_{i1}^{\alpha_1-1}\cdots p_{ik}^{\alpha_k-1}(1-p_{i1}-\cdots-p_{ik})^{\alpha_{k+1}}dp_{ik}\cdots dp_{i1} \\
&= mD(\alpha_j,\alpha_1+\cdots+\widehat{\alpha_j}+\cdots+\alpha_{k+1})
\end{aligned} \tag{26}$$

EM Algorithm appl. to Mult. Scheme

- Finally: suppose that $(\alpha_1^{(p)}, \ldots, \alpha_{k+1}^{(p)})$ are initial values of the parameters.
- Then new values $(\alpha_1^{(p+1)}, \ldots, \alpha_{k+1}^{(p+1)})$are given by the solutions of the following system:

$$\begin{cases} mD(\alpha_1^{(p+1)}, \alpha_2^{(p+1)} + \alpha_3^{(p+1)} + \cdots + \alpha_{k+1}^{(p+1)}) = f_1(\alpha_1^{(p)}, \ldots, \alpha_{k+1}^{(p)}) \\ mD(\alpha_2^{(p+1)}, \alpha_1^{(p+1)} + \alpha_3^{(p+1)} + \cdots + \alpha_{k+1}^{(p+1)}) = f_2(\alpha_1^{(p)}, \ldots, \alpha_{k+1}^{(p)}) \\ \qquad \ldots \qquad\qquad\qquad\qquad \ldots \\ mD(\alpha_{k+1}^{(p+1)}, \alpha_1^{(p+1)} + \alpha_2^{(p+1)} + \cdots + \alpha_k^{(p+1)}) = f_{k+1}(\alpha_1^{(p)}, \ldots, \alpha_{k+1}^{(p)}) \end{cases} \tag{27}$$

where $f_j(\alpha_1^{(p)}, \ldots, \alpha_{k+1}^{(p)})$ are the values of the conditional expectations.

EM Algorithm appl. to Iter. Scheme

- In an analogous way we obtain that for the iterative urn scheme the new values of the parameters $(\alpha_1^{(p+1)}, \beta_1^{(p+1)}, \ldots, \alpha_k^{(p+1)}, \beta_k^{(p+1)})$ are given by the solutions of the following system:

$$\begin{cases} mD(\alpha_1^{(p+1)}, \beta_1^{(p+1)}) = f_1(\alpha_1^{(p)}, \beta_1^{(p)}, \ldots, \alpha_k^{(p)}, \beta_k^{(p)}) \\ mD(\beta_1^{(p+1)}, \alpha_1^{(p+1)}) = f_2(\alpha_1^{(p)}, \beta_1^{(p)}, \ldots, \alpha_k^{(p)}, \beta_k^{(p)}) \\ \qquad \ldots \qquad\qquad\qquad \ldots \\ mD(\alpha_k^{(p+1)}, \beta_k^{(p+1)}) = f_{2k-1}(\alpha_1^{(p)}, \beta_1^{(p)}, \ldots, \alpha_k^{(p)}, \beta_k^{(p)}) \\ mD(\beta_k^{(p+1)}, \alpha_k^{(p+1)}) = f_{2k}(\alpha_1^{(p)}, \beta_1^{(p)}, \ldots, \alpha_k^{(p)}, \beta_k^{(p)}) \end{cases} \tag{28}$$

- Note that instead of solving a system of $2k$ equations in $2k$ unknowns we have only to solve k independent systems of two equations in two unknowns.

Calibration of models

- Data used: Standard & Poor's 2005 report.
- Considered only the six rating class from AA to CCC (AAA shows no default).
- Starting values of the parameters: moment estimated ones through underlying distributions.
- For example, for multidimensional scheme, since the default frequency of N_j is given by $P_1 + \cdots + P_j$, we can take, for n_j large, l_j/n_j as realization of $P_1 + \cdots + P_j$, $1 \leq j$, hence

$$\left(\frac{l_1}{n_1}, \frac{l_2}{n_2} - \frac{l_1}{n_1}, \ldots, \frac{l_k}{n_k} - \frac{l_{k-1}}{n_{k-1}}\right)$$

are realizations of $(P_1, \ldots, P_k) \sim D_{k+1}(\alpha_1, \ldots, \alpha_{k+1})$.

Calibration of models

- Implementation of the algorithm: first naive try using Mathematica 5.0®, only feasible for groups of three rating classes.
- Better results using C++ code; libraries for the computation of Gamma and related functions provided by GNU Scientific Library.
- For the solution of the system we used M. J. D. Powell hybrid method for nonlinear equations.
- Stopping criteria based on stability of log-likelihood and moments of underlying distributions.
- Improvememts on speed of computations for groups of 4 and 5 rating classes for iterative urn scheme.
- We have been able to calibrate also multidimensional urn scheme for groups of 4 and 5 rating classes.
- Sensibility with respect to starting values of the algorithm.
- Comparison with previous calibrations shows different values for multidimensional urn scheme.

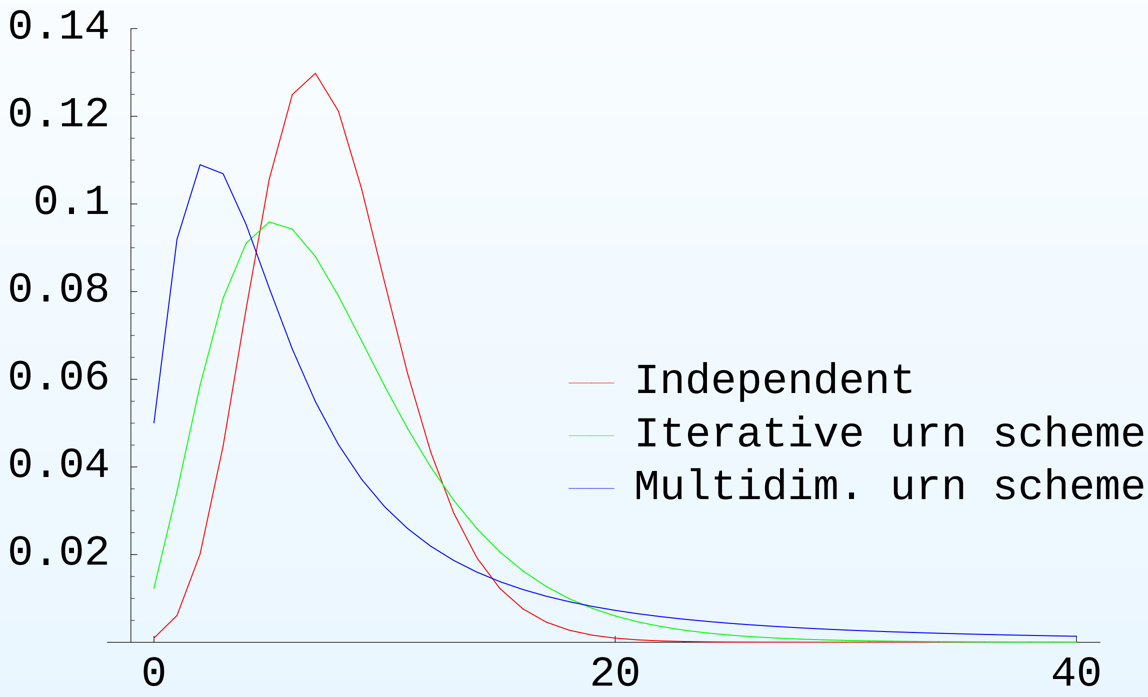

figure 1: The distribution of the total number of defaults in the group of 27808 AA-, 4832 A- and 753 BBB-rated firms, calculated with independent beta mixtures, the iterative urn scheme and the multidimensional urn scheme. Each rating group has an expectation of 2 defaults

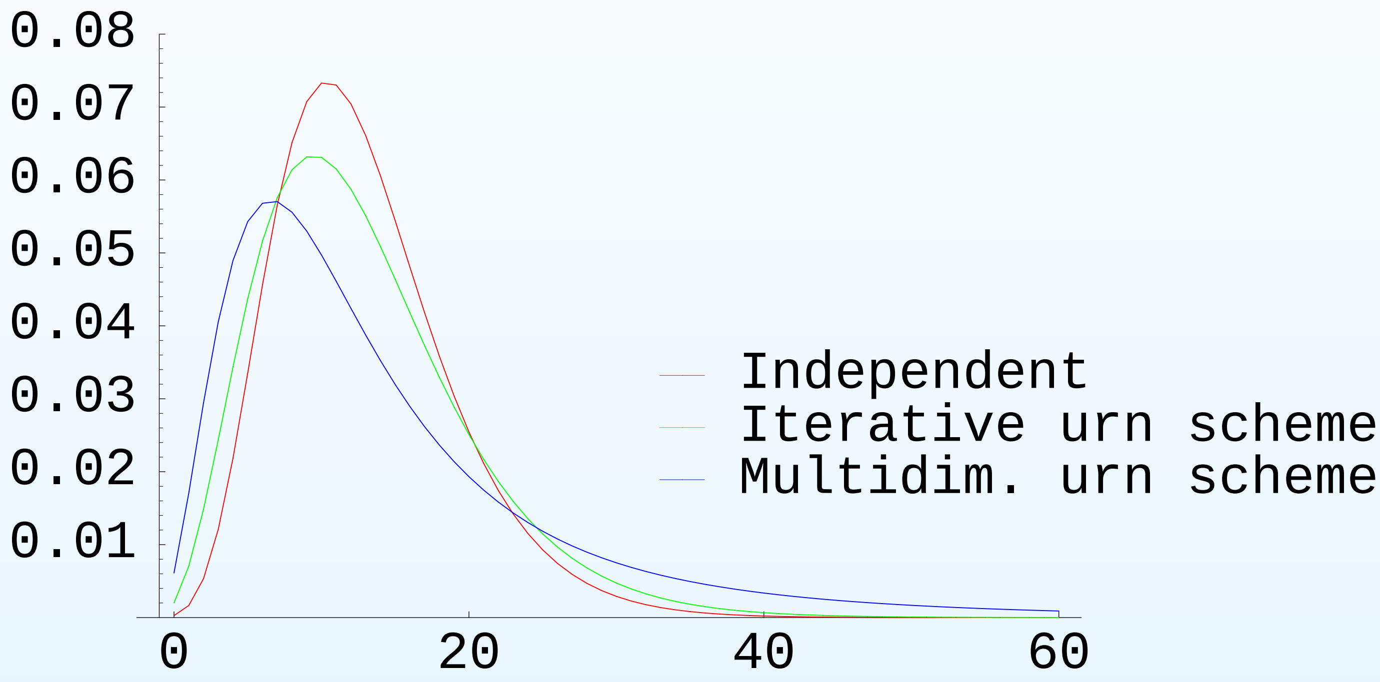

figure 2: The distribution of the total number of defaults in the group of 6103 A-, 1379 BBB- and 404 BB-rated firms, calculated with independent beta mixtures, the iterative urn scheme and the multidimensional urn scheme. Each rating group has an expectation of 5 defaults

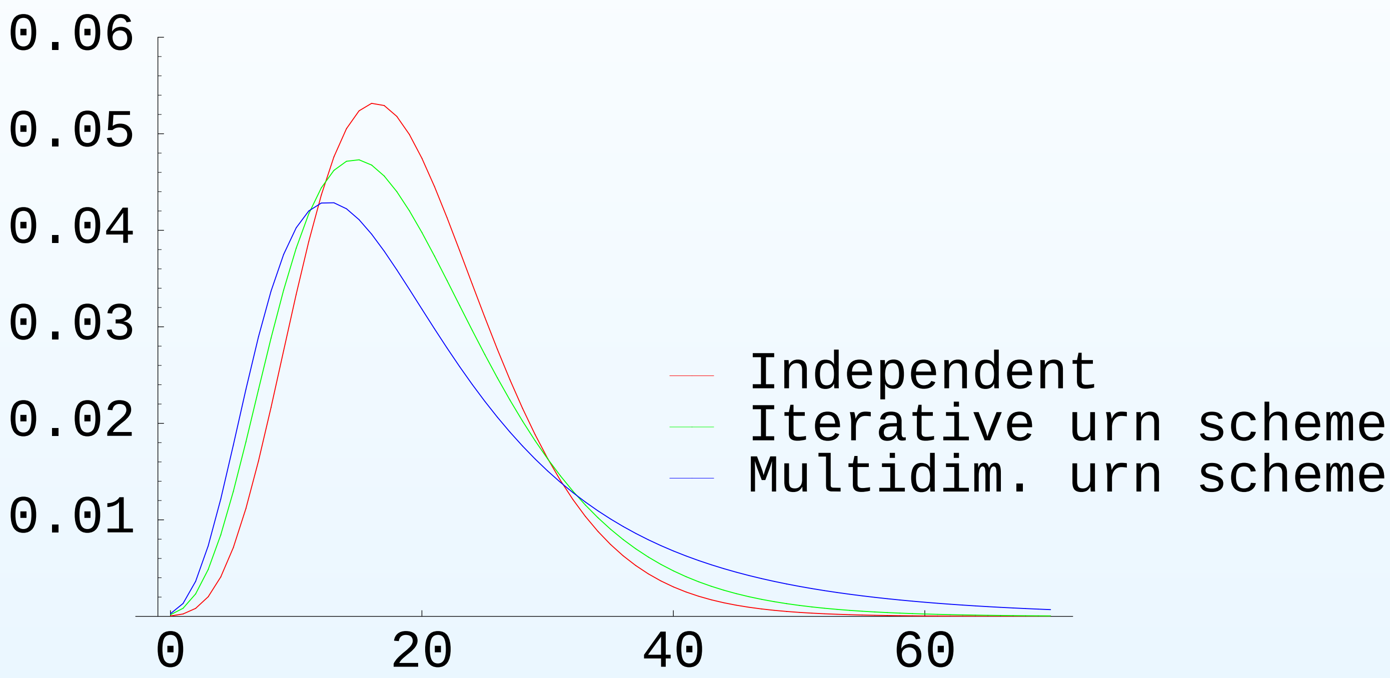

figure 3: The distribution of the total number of defaults in the group of 1639 BBB-, 525 BB- and 124 B-rated firms, calculated with independent beta mixtures, the iterative urn scheme and the multidimensional urn scheme. Each rating group has an expectation of 7 defaults.

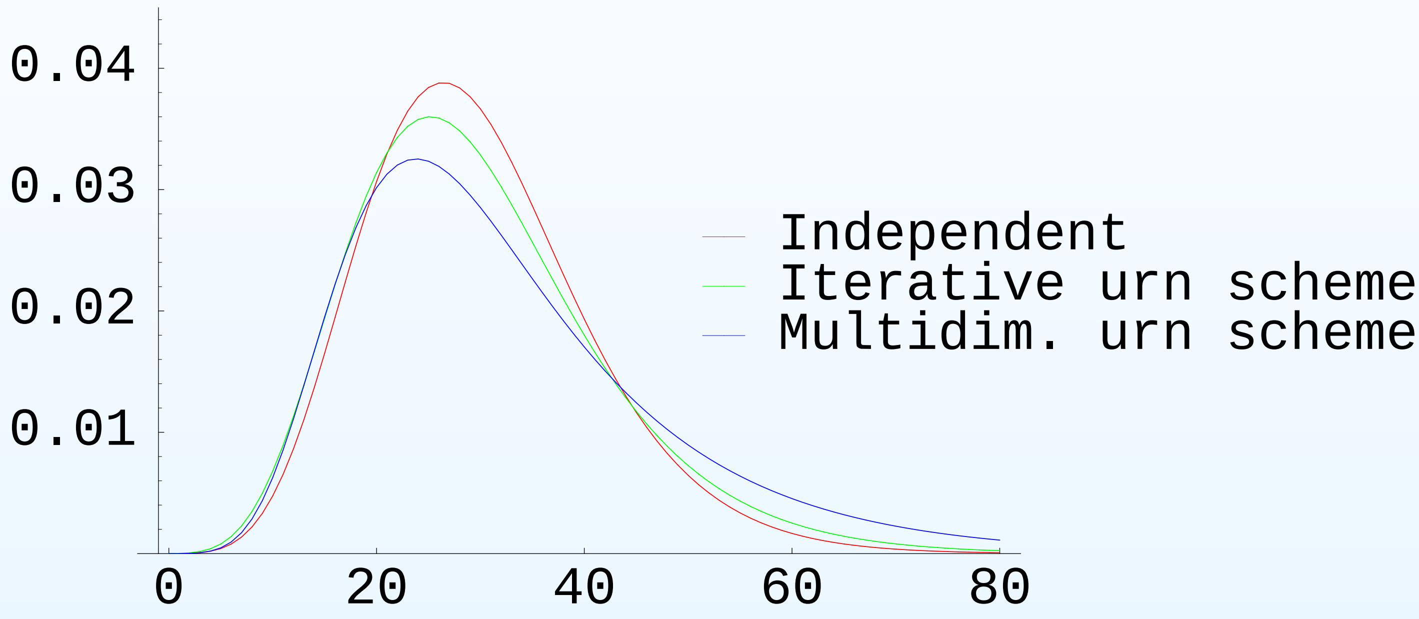

figure 4: The distribution of the total number of defaults in the group of 772 BB-, 178 B- and 44 CCC-rated firms, calculated with independent beta mixtures, the iterative urn scheme and the multidimensional urn scheme. Each rating group has an expectation of 10 defaults.

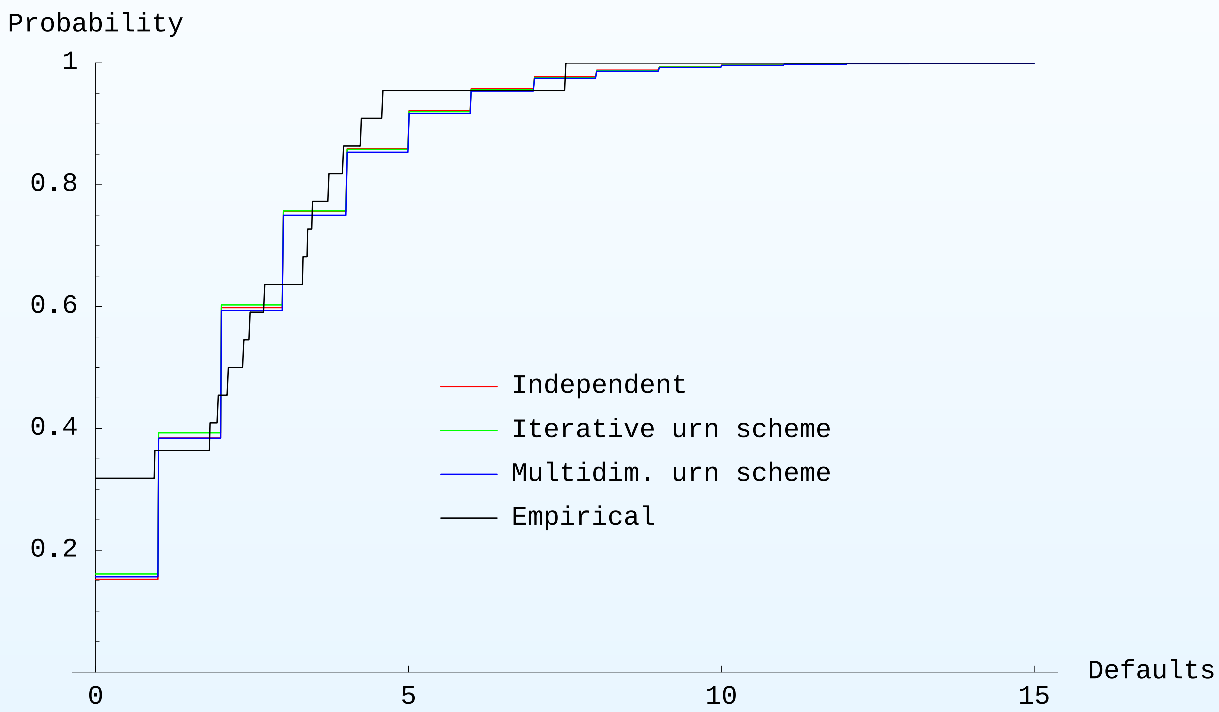

figure 5: Cumulative distribution of the total number of defaults in a group of 454 AA-, 794 A- and 594 BBB-rated firms (average portfolio). The probabilities are calculated with independent beta mixtures, the iterative urn scheme and the multidimensional urn scheme model, compared with the empirical from the scaled data.

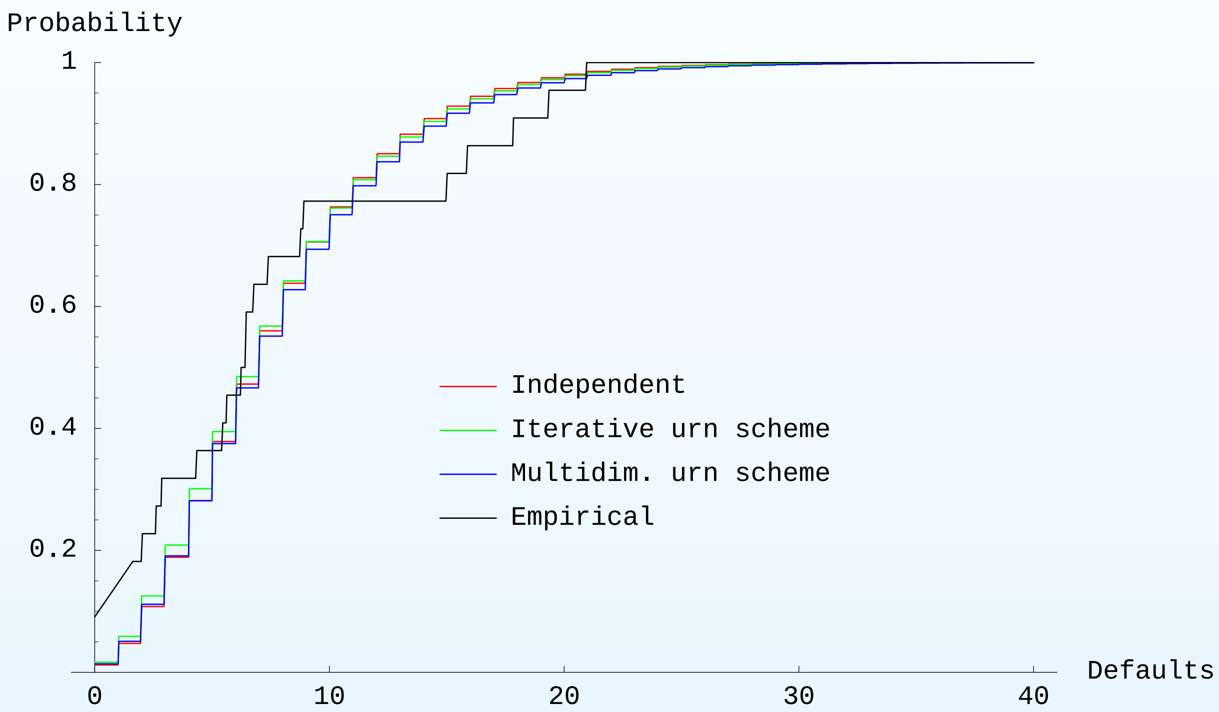

figure 6: Cumulative distribution of the total number of defaults in a group of 794 A-, 594 BBB- and 411 BB-rated firms (average portfolio). The probabilities are calculated with independent beta mixtures, the iterative urn scheme and the multidimensional urn scheme model, compared with the empirical from the scaled data.

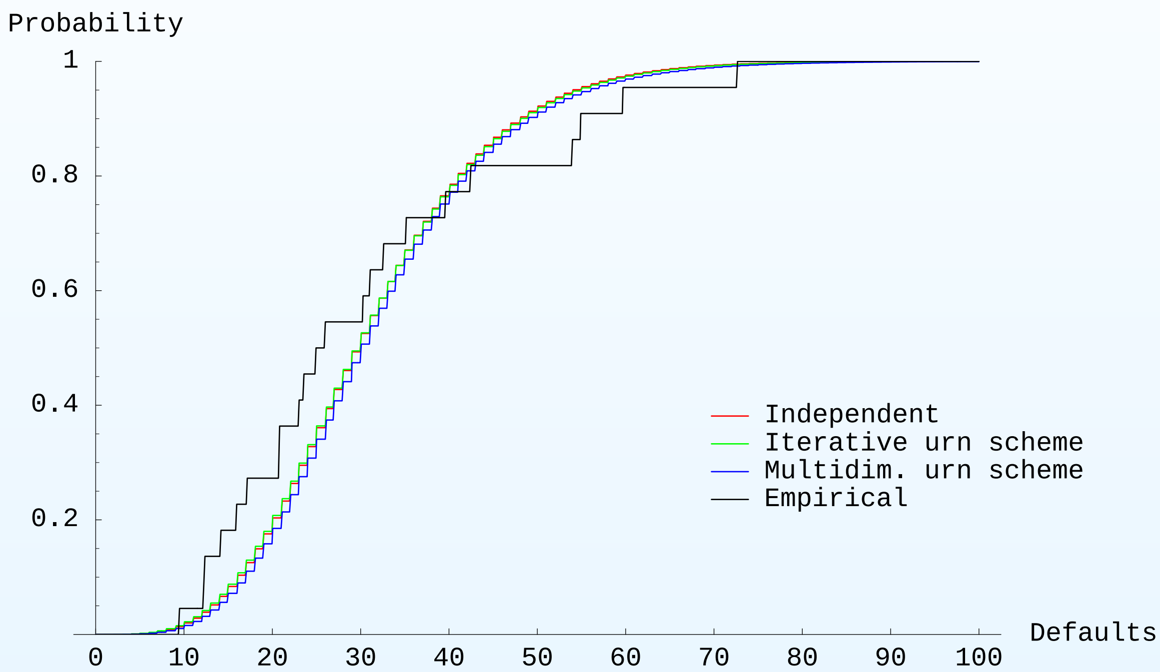

figure 7: Cumulative distribution of the total number of defaults in a group of 594 BBB-, 411 BB- and 427 B-rated firms (average portfolio). The probabilities are calculated with independent beta mixtures, the iterative urn scheme and the multidimensional urn scheme model, compared with the empirical from the scaled data.

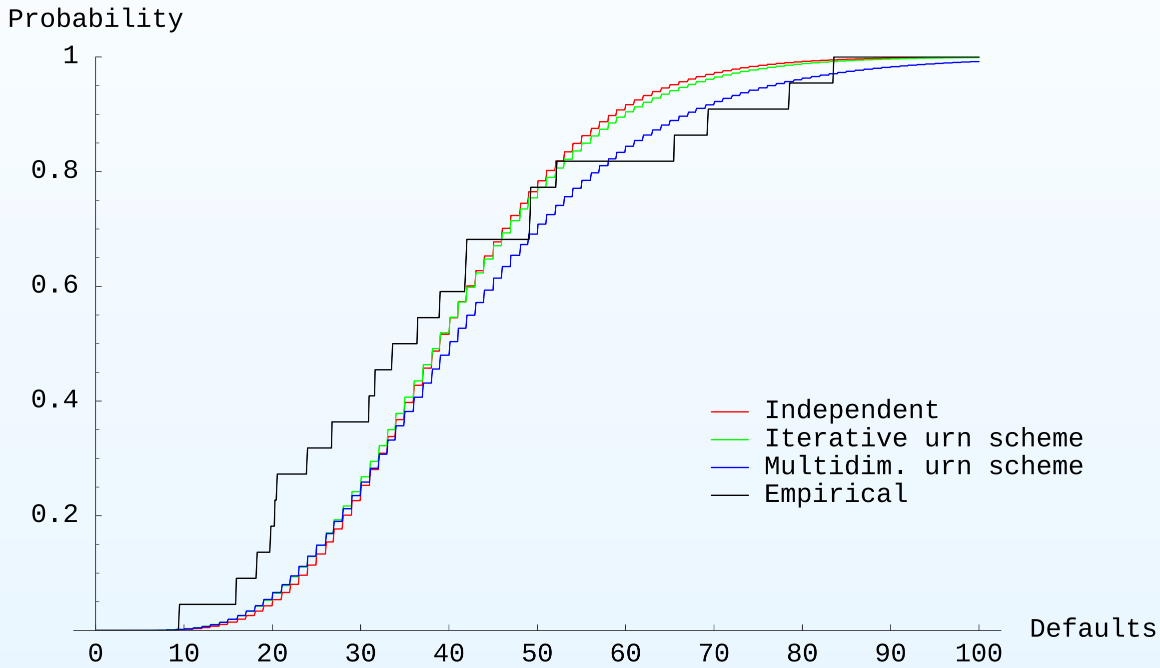

figure 8: Cumulative distribution of the total number of defaults in a group of 411 BB-, 427 B- and 49 C-rated firms (average portfolio). The probabilities are calculated with independent beta mixtures, the iterative urn scheme and the multidimensional urn scheme model, compared with the empirical from the scaled data.

Further improvements

- Biggest problem: huge amount of computations required.
- It implies time consuming numerical calculations both for the estimation of the parameters and for the plotting of the graphics.
- One chance: sum only the relevant terms in equations (7) and (10), that should correspond to sum only “in a neighbourhood” of the expected number of defaults given the parameters.
- Try to see if EM continues to give good results with such approximations.
- Another chance: try to find simpler analytical expression for (7) and (10).