Safety and Reliability – Bedford & van Gelder (eds)
© 2003 Swets & Zeitlinger, Lisse, ISBN 90 5809 551 7

A non-parametric two-stage Bayesian model using Dirichlet distribution

C. Bunea, R.M. Cooke & T.A. Mazzuchi
Delft University of Technology, Delft, Netherlands

G. Becker
RISA, Berlin, Germany

ABSTRACT: As an alternative to standard two-stage Bayesian models, a non-parametric or Dirichelet two-stage model is presented. The analytic solution of the model and its clear interpretation are its main advantages over the classical models. A number of case study are simulated in order to check the robustness of the model. Three data sets from German project – ZEDB are also used to compare the results of the Dirichelet model with the results for standard one-stage and two-stage Bayesian models.

1 INTRODUCTION

Within the context of a recent review of a two-stage Bayesian model for processing data at a population of German nuclear plants, a nonparametric or Dirichelet two stage model was developed. This model has some advantages relative to the standard two stage models: it is analytically solvable, no numerical integration need to be performed, and it allows an intuitive interpretation of the (hyper)parameters – so called "equivalent observations". We check the robustness of this model with a simple numerical example. Preliminary calculations show some sensitivity of the model with regard to the number of cells that characterized the prior distribution and their end points. For the purposes of comparison with classical Bayesian model [Vaurio, Hofer, Becker], the results for three data sets are presented. Considering its apparent advantages, the authors may recommend that the Dirichlet model deserves further development, to qualify it for practical use in data base analysis.

2 BAYESIAN TWO STAGE HIERARCHICAL MODELS

A two-stage model is really nothing more than a joint distribution [Cooke et al 2002]. To be useful, however, we must derive conditional distributions. Typically we want to use data from "other plants" to make predictions about a given plant. This is very attractive in cases where the data from the given plant is sparse.

By specifying the model assumptions one can derive the posterior distribution $P(\lambda_0|X_0, \dots X_n)$ for failure rate λ_0 at plant of interest 0, given X_i failures and T_i observation times at plant i, $i = 0, 1, \dots n$. First, we can identify the conditional independence assumptions in order to factor the joint distribution. The conditional independence assumptions met in the literature, with one possible exception [Hofer et al 1997][Hofer 1999], are stated below:

CI.1 Given Q, λ_i is independent of $\{X_j, \lambda_j\}_{j \neq i}$
CI.2 Given λ_i, X_i is independent of $\{Q, \lambda_j, X_j\}_{j \neq i}$,

where Q is the hyperparameter of the prior distribution from which the Poisson intensities $\lambda_1 \dots \lambda_n$ are drawn.

The expression "X_i is independent of $\{Q, \lambda_j, X_j\}_{j \neq i}$" entails that X_i is independent of Q, and X_i is independent of λ_j.

Giving the conditional independence assumptions, [Cooke et al 2002] derived the explicit form of the posterior distribution $P(\lambda_0|X_0, \dots X_n)$ for failure rate λ_0:

$$P(\lambda_0|X_0..X_n) \propto P(X_0|\lambda_0)$$

$$\int_Q P(\lambda_0|Q) \prod_{i=1}^{n} \int P(X_i|\lambda_i) P(\lambda_i|Q) d\lambda_i P(Q) dQ \qquad (1)$$

Assumptions must be made also regarding the fixed distribution types and the hyperprior Q. In the two stage Bayesian models considered here, the likelihood of the failure times from each plant i, $P(X_i, T_i|\lambda_i)$, given λ_i and given any information from other plants, are independent and follow a Poisson distribution with parameter λ_i.

Parametric two stage Bayesian models consider usually gamma or log normal as prior distribution $P(\lambda_i|Q)$. The second stage places a hyperprior distribution over the parameters of the prior gamma or log normal distribution.

Controversies arise over the choice of hyperprior. [Cooke et al 1995] showed that the noninformativeness is not a good criteria if it leads to improper distribution. Since improper hyperpriors do not always become proper after observation, they should be avoided if property cannot be demonstrated.

Another major criticism to parametric two stage Bayesian models is the non-analytically solution of the model. Numerical integration should be performed in order to obtain the posterior distribution for λ_0. [Cooke et al 2002] showed that the method of truncation seems to have a large influence on the posterior distribution of λ_0.

3 VAURIO MODEL

As an alternative to standard two-stage Bayesian models, [Vaurio, 1987] proposed an analytic empirical Bayes approach to the problem of assimilating data from other plants. A simple one-stage Bayesian model for one plant would use a Poisson likelihood with intensity λ and a $Gamma(\lambda|a, b)$ prior. Updating the prior with X_i failures in time T_i yields a $Gamma\ (\lambda|\alpha + X_i, \beta + T_i)$ posterior. Vaurio proposes to use data from the population of plants to choose the $Gamma(\lambda|\alpha, \beta)$ prior by moment fitting. Any other two moment prior could be used as well. Data from other plants are not used in updating, hence, this is a one-stage model.

We sketch Vaurio's model in the simple case that the observation times at all $n + 1$ plants are equal to T. The population mean and (unbiased) variance are estimated as:

$$m = (\sum_{i=0..n} X_i/T)/(n+1)$$

$$v = (\sum i = 0..n(X_i/T - m)^2)/n$$

A shifted variance estimate, which is positive when at least one of the $X_i > 0$, $i = 0, \dots n$; is defined as:

$$V = vm/nT$$

V and m are used to solve for the shape α and scale β of a gamma prior $G(\lambda_i|\alpha, \beta)$:

$$\alpha = m^2/V$$
$$\beta = m/V$$

Using the familiar gamma-Poisson one stage model, the posterior mean and variance for λ_i after observing X_i failures in time T, are:

$$E(\lambda_i|X_i, T) = (\alpha + X_i)/(\beta + T)$$
$$Var(\lambda_i|X_i, T) = (\alpha+X_i)/(\beta+T)^2 \tag{2}$$

The model is consistent, in the sense that as $X_i, T_i \to \infty$, with $X_i/T_i \to \lambda_i$, his model does entail that $E(\lambda_i|X_i, T_i) \to \lambda_i$. Elegance and simplicity are its main advantages. Disadvantages are that it cannot be applied if all $X_i = 0$, or if the population consists of only 2 plants. Further, numerical results in section 5 indicate that the model is non-conservative when the empirical failure rate at plant 0 is low and the empirical failure rates at other plants are high. A final criticism, which applies to most empirical Bayes models is that the data for the plant of interest is used twice, once to estimate the prior and once again in the Poisson likelihood. Thus, X_i occurs in (2) twice, once as X_i, and again in the estimate of α and β. This may contribute to the non-conservativism noted in section 5.

4 A NON-PARAMETRIC OR DIRICHLET MODEL

Given the problems with the approaches described above, we explore the possibility of a non-parametric Bayesian two stage model. Very roughly, in this model we select a number of points $L_0, L_1, \dots L_k$. The parameters of our prior distribution are probabilities $q = (q_1, \dots q_k)$, adding to unity, such that:

$$P(L_{i-1} < \lambda \le L_i|Q = q) = q_i; i = 1, ...k. \tag{3}$$

The mechanism for doing this is the Dirichlet distribution. In this section we set up the model in simple terms, and examine its assumptions.

4.1 *Prior parameters*

The prior distribution $P(\lambda|q)$ is characterized by

- a fixed number of points: $L_0 < L_1 < \dots L_k$
- a probability vector $q = (q_1, \dots q_k)$; $q_i \geqslant 0, \Sigma q_i = 1$
- a probability density $g(\lambda)$ defined on (L_0, L_k)

4.2 *Prior distribution*

The points $L_0, \dots L_k$ and the density $g(\lambda)$ are chosen by the user in a manner discussed below, and are not uncertain. There is no hyperprior over these. Letting $C_i = (L_i - 1, L_i)$, the prior may be written as

$$P(\lambda|q) = \sum_{i=1..k} 1_{\{\lambda \in C_i\}}(\lambda) q_i g(\lambda) / \int_{C_i} g(\lambda) d\lambda. \tag{4}$$

Here, $1_A(x)$ denotes the indicator function taking the value 1 if $x \in A$, and zero otherwise.

We shall take $g(\lambda)$ to be the uniform density, then $g(\lambda)$ is constant and we may write this as:

$$P(\lambda|q) = \sum_{i=1..k} 1_{\{\lambda \in C_i\}}(\lambda) q_i / |C_i| \quad (5)$$

Note that a gamma distribution for $g(\lambda)$ will provide also an analytical solution of the model.

4.3 *Hyperprior*

The hyperprior is a Dirichlet distribution characterized by parameters $a_1 \ldots a_k$; $a_i > 0$:

$$P(q) = \prod q_i^{(a_i - 1)} / D(a_1, \ldots a_k) \quad (6)$$

where $D(a_1, \ldots a_k)$ is the Dirichlet integral:

$$\int_{q_1..q_k} \prod q_i^{(a_i-1)} dq_1..dq_k = \frac{\Pi(a_i - 1)!}{(\sum a_i - 1)!} \quad (7)$$

where $q_1 + \cdots q_k = 1$.

For reasons explained below, the parameters are sometimes called "equivalent observations of $\lambda \in C_i$".

We will choose $a_i = 1$; $i = 1, \ldots k$. Then the $D(a_1, \ldots a_k) = 1/(k-1)!$, and may be absorbed into the normalization constant, so that the hyperprior becomes:

$$P(q) = 1 \quad (8)$$

4.4 *Updating the model*

4.4.1 *Hyperposterior*

Considering only plant i, after observing X_i failures at plant i, the hyperposterior becomes

$$\int P(X_i|\lambda_i) P(\lambda_i|q) d\lambda_i P(q)$$

$$= (1/X_i!) \int (\lambda_i T_i) X_i exp(-\lambda_i T_i) \sum_{h=1..k} 1_{C_h}(\lambda_i) q_h d\lambda_i$$

$$= (1/X_i!) \sum_{h=1..k} q_h \int_{C_h} (\lambda_i T_i)_i^X exp(-\lambda_i T_i) d\lambda_i \quad (9)$$

Since X_i is an integer, the integral in (9) can be evaluated explicitly. We find for $C_h = (L_{h-1}, L_h)$:

$$(1/X_i!) \int_{C_h} (\lambda_i T_i)_i^X exp(-\lambda_i T_i) d\lambda_i$$

$$= (1/T_i) exp(-L_{h-1} T_i) \sum_{j=0..X_i} (L_{h-1} T_i)^j / j!$$

$$-(1/T_i) exp(-L_h T_i) \sum_{j=0..X_i} (L_h T_i)^j / j! = A_{i,h} \quad (10)$$

4.4.2 *Posterior*

Substituting (10) into (9), and this into (1) we find:

$$P(\lambda_0|X_0, \ldots X_n) \propto$$

$$P(x_0|\lambda_0) \int \sum_{m=1..k} 1_{C_m}(\lambda_0) q_m \prod_{i=1..n} \sum_{h=1..k} A_{i,h} q_h dq \quad (11)$$

The integral is taken over the set $\{q = q_1, \ldots q_k | q_i \geq 0, \Sigma q_i = 1\}$. We can write the integrand as a sum of products.

Let us consider one such term. We may write this term as

$$A_{1,h(1)} A_{2,h(2)} \ldots A_{n,h(n)} q_1^{r_1} q_2^{r_2} \ldots q_k^{r_k} \quad (12)$$

Where, since there are $n+1$ plants it total,

$$r_1 + r_2 + ..r_k = n + 1$$

The terms $A_{1,h(1)} A_{2,h(2)} \ldots A_{n,h(n)}$ do not contain q, and we may evaluate this integral explicitly for each term, using the Dirichlet integral. We find

$$P(\lambda_0|X_0 \ldots X_n) \propto$$

$$P(\lambda_0|X_0) \sum A_{1,h(1)} A_{2,h(2)} \ldots A_{n,h(n)} \prod r_i! \quad (13)$$

Where the summation goes over all k^n terms, and $r = r_1, \ldots r_k$ is specific to each of the k^n terms. Note that (13) expresses $P(\lambda_0|X_0, \ldots X_n)$ completely in terms of $P(\lambda_0|X_0)$, $L_j, j = 0, \ldots k$; $X_i, T_i, i = 1, \ldots n$. In other words, this model is completely solvable analytically.

If n is, say 10, and k is, say 4, then we have $4^{10} = 10^6$ such terms, which is a feasible number.

4.5 *Equivalent observations*

The parameters $a_1, \ldots a_k$ of the hyperprior are sometimes called equivalent observations. Indeed, the Dirichlet distribution is the natural conjugate for the multinomial likelihood. Consider rolling a die with k faces M times, where the probability of seeing face i is q_i. The probability of seeing face "i" r_i times; $i = 1, \ldots k$ is

$$\prod q_i^{r_i} M! / (r_1!, \ldots r_k!) \quad (14)$$

The result of updating the Dirichlet prior

$$P(q) = \prod q_i^{(a_i-1)} / D(a_1, a_k)$$

with these observations is again a Dirichlet:

$$P(q|r_1..r_k) = \prod q_i^{(a_i+r_i-1)}/D(a+r_1..a_k+r_k) \quad (15)$$

This suggests that the parameters a_i of the original Dirichlet prior may be interpreted as if we started with the prior

P(q) 5 1/D(1, 1… 1)

and observed face "i" $a_i - 1$ times, $i = 1, k$. This yields a useful heuristic for interpreting the parameters of the hyperprior in the ordered Dirichlet two stage model. When we choose the Dirichlet hyperprior (7) with $a_i = 1$, we are adopting a (hyper)prior belief state in which we have not yet observed one λ falling in any cell $C_j, j = 1, \ldots k$.

4.6 *Choosing parameters*

The Dirichlet model requires the user to choose parameter values which cannot be updated. These are

1. The number k of cells $C_1, \ldots C_k$,
2. The points $L_0 \leq L_1 \leq L_2, \ldots \leq L_k$
3. The values $a_1, \ldots a_k$.
4. The density $g(\lambda)$.

We have already indicated that the parameters a_i are chosen equal to one, to reflect "no equivalent observations".

The number of cells should be chosen so as to guarantee that after a long period of observation at each of the $n + 1$ plants, the hyperprior is "forgotten". If we consider (15), we see that the hyperprior is forgotten when most of the terms r_i are greater than zero. This suggests that with, say 10 to 15 plants, the number of cells should not exceed, say, four.

Given that we will have k cells, we must choose the points L_j defining the cells $C_j = (L_j - 1, L_j)$. Since L is a measure of $\lambda_i = \lim_{T_i \to \infty} X_i/T_i$, the end points L_0 and L_k should be chosen such that the term $exp(-LT_i)\Sigma_{j=0 \ldots X_i}(LT_i)^j/j!$ in (9), with $L_0 \leq L \leq L_k$, covers all the possible values. This corresponds to calculate the hyperposterior distribution in one cell scenario for all $i = 0, 1 \ldots n$ plants. Hence, the lower and the upper bounds will be the minimum of L_0, respectively the maximum of L_k, found for every plant $i = 0, 1 \ldots n$. Figure 1 shows the hyperposterior distribution as a function of L in one cell scenario, for a plant with one failure over 10000 hours of observation. Using the assimptotic properties of term $exp(-LT_i)\Sigma_{j=0 \ldots X_i}(LT_i)^j/j!$, a natural band can be found as $L_0 = 10^{-6}$ and $L_k = 10^{-3}$.

The points L_j, $j \neq 0$ and $j \neq k$, should be chosen such that we expect, before performing observations,

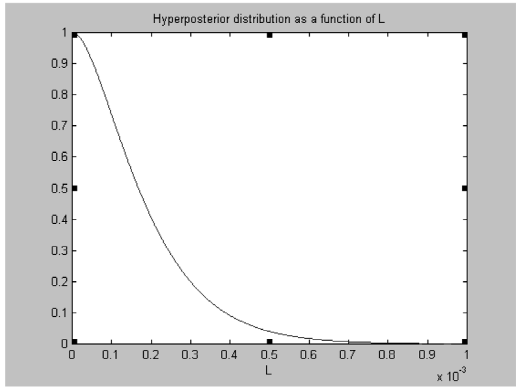

Figure 1. Hyperposterior distribution as a function of L in one cell scenario.

that the number of

$$\lambda_i = \lim_{T_i \to \infty} X_i/T_i$$

which, after a very long observation period at each plant, we will see, falling into each cell C_j is equal.

Finally we must choose the density $g(\lambda)$. We have chosen the uniform density as it is the least informative. Other choices could be made without sacrificing tractability.

4.7 *Summary of significant features*

We summarize the significant features of the Dirichlet model with the choices described above:

1. The model is solvable analytically, no numerical integration need be preformed.
2. The hyperprior has a clear intuitive interpretation.
3. The hyperprior is minimally informative in the sense of "no equivalent observations", and is proper.
4. The size and number of the cells C_j is chosen to insure that the hyperprior does not persist on observing $n + 1$ plants.

5 NUMERICAL EXAMPLE

We illustrate this model with a simple numerical example. These computations have all been performed on a spreadsheet, as no numerical integration is required. Nonetheless, for 7 plants the time required to compute the normalized posterior for λ_0 is about 10 minutes. We illustrate the model with 4 other plants, having between 1000 hours and 4000 hours operational time. Plant 0 is computed in three cases, namely with zero failures and 100, resp. 1000 resp. 10,000 operating hours. The

Table 1.

Plant	T	X	X/T	
1	1000	5	5.00E-3	
2	3000	20	6.67E-3	
3	3500	50	1.43E-2	
4	4000	100	2.50E-2	
0	100	0		
L0	L1	L2	L3	L4
1.00E-8	5.00E-4	5.00E-3	1.00E-2	1.00E-1
Ex w.o.	Ex w.	5%	50%	95%
4.08E-3	4.83E-3	7.57E-4	4.70E-3	9.46E-3
Prob	$(\lambda_i \in C_j)$			
Plant 0	0.04877	0.3447	0.23865	0.36783
Plant 1	1.42E-5	0.38402	0.54887	0.06709
Plant 2	−2.2E-16	0.08297	0.88174	0.03528
Plant 3	−2.2E-16	6.0E-11	0.00653	0.99347
Plant 4	3.3E-16	3.6E-15	−1.4E-15	1

Table 2.

Plant	T	X	X/T	
1	1000	5	5.000E-3	
2	3000	20	6.667E-3	
3	3500	50	1.429E-2	
4	4000	100	2.500E-2	
0	1000	0		
L0	L1	L2	L3	L4
1.00E-8	5.00E-4	5.00E-3	1.00E-2	1.00E-1
Ex w.o.	Ex w.	5%	50%	95%
1.10E-3	1.16E-3	8.67E-5	8.88E-4	3.05E-3
Prob	$(\lambda_i \in C_j)$			
Plant 0	0.39346	0.59979	0.00669	4.5E-5
Plant 1	1.42E-5	0.38402	0.54887	0.06709
Plant 2	−2.2E-16	0.08297	0.88174	0.03528
Plant 3	−2.2E-16	6.0E-11	0.00653	0.99347
Plant 4	3.3E-16	3.6E-15	−1.4E-15	1

results are shown in Tables 1, 2 and 3. Each table shows the operational data from the four other plants and the cells. This data is the same in each table. Also shown is the (unnormalized) probability that λ_i falls in cell C_j; this data differs for λ_0 in each table. Very small negative probabilities are caused by numerical errors in EXCEL. "Ex w." and "Ex w.o." denote updating with and without the data from plants 1 … 4. Data from other plants has the effect of raising the posterior expectation.

For plant 0 with 1000 operating hours, Table 2 shows that the updating with other plants now has less effect on the posterior expectation.

For 10,000 operating hours and still zero failures at plant 0, the posterior expectations with and without other plants are practically the same. Now, after 10,000 hours, the expectation at plant 0 is determined only by the data at plant 0, and the other plants have almost no effect. Note that the probability for λ_0 is concentrated in cell C_1.

Table 3.

Plant	T	X	X/T	
1	1000	5	5.000E-3	
2	3000	20	6.667E-3	
3	3500	50	1.429E-2	
4	4000	100	2.500E-2	
0	10000	0		
L0	L1	L2	L3	L4
1.00E-8	5.00E-4	5.00E-3	1.00E-2	1.00E-1
Ex w.o.	Ex w.	5%	50%	95%
1.23E-4	1.23E-4	8.67E-5	7.22E-5	2.84E-4
Prob	$(\lambda_i \in C_j)$			
Plant 0	0.99316	0.00674	1.9E-22	3.7E-44
Plant 1	1.42E-5	0.38402	0.54887	0.06709
Plant 2	−2.2E-16	0.08297	0.88174	0.03528
Plant 3	−2.2E-16	6.0E-11	0.00653	0.99347
Plant 4	3.3E-16	3.6E-15	−1.4E-15	1

Table 4.

Plant	T	X	X/T	
1	1000	5	5.000E-3	
2	3000	20	6.667E-3	
3	3500	50	1.429E-2	
4	4000	100	2.500E-2	
5	2000	15	7.500E-3	
6	2500	15	6.000E-3	
0	100	0		
L0	L1	L2	L3	L4
1.00E-8	5.00E-4	5.00E-3	1.00E-2	1.00E-1
Ex w.o.	Ex w.	5%	50%	95%
4.08E-3	5.49E-3	8.93E-4	5.75E-3	9.54E-3
Prob	$(\lambda_i \in C_j)$			
Plant 0	0.04877	0.3447	0.23865	0.36783
Plant 1	1.42E-5	0.38402	0.54887	0.06709
Plant 2	−2.2E-16	0.08297	0.88174	0.03528
Plant 3	−2.2E-16	6.0E-11	0.00653	0.99347
Plant 4	3.3E-16	3.6E-15	−1.4E-15	1
Plant 5	1.8E-14	0.04874	0.79475	0.15651
Plant 6	5.2E-13	0.19397	0.78374	0.02229

For purposes of comparison, Table 4 shows the results of updating with 6 other plants, when plant 0 has experience no failures in 100 operating hours. Plants 5 and 6 have empirical failure rates in the same order as the first 4 plants. We see that the results are a bit larger than those of Table 1.

Table 5 gives the results of the Dirichlet model for the Data set 2 of [Becker and Hofer 2001], and com-

Table 5.

Plant	T	X	X/T	
1	20000	1	5.00E-5	
2	2000	0	0	
3	4000	0	0	
4	6000	0	0	
5	10000	1	1.00E-4	
0	12000	2		
L0	L1	L2	L3	L4
1.00E-8	5.00E-6	5.00E-5	5.00E-4	1.00E-3
Ex w.o.	Ex w.	5%	50%	95%
2.32E-4	2.31E-4	7.45E-5	2.13E-4	4.32E-4
Gamma		1.22E-4	1.72E-4	3.45E-4
Lgnormal		2.54E-5	7.71E-5	2.06E-4
Vaurio		3.34E-5	1.14E-4	2.74E-4
Prob	$(\lambda_i \in C_j)$			
Plant 0	3.44E-5	0.02308	0.91492	0.06145
Plant 1	0.00468	0.25956	0.73526	0.0005
Plant 2	0.00993	0.08521	0.53696	0.23254
Plant 3	0.01976	0.16147	0.68339	0.11702
Plant 4	0.02949	0.22963	0.69103	0.04731
Plant 5	0.00121	0.08899	0.86937	0.03993

pares these with the results for standard Bayesian models [Cooke et al 2002]. We see that the results are of the same order, though a bit higher. The intervals (L_{i-1}, L_i) are chosen to bound the empirical rates (when failures are present).

Table 6 compares the Dirichlet results with Vaurio's estimator. To avoid numerical procedures, all observation times are equal. We see that there are significant differences. In general, Vaurio's estimate is closer to the empirical failure rate of plant 0. In those cases where plants 1..5 exhibit a high empirical failure rate, and the plant of interest, plant 0, has no failures, Vaurio's estimate is lower than the Dirichlet estimate by a factor 2. This feature is observed for short observation times (1000 hours) and long observation times (10,000 hours).

These results indicate that Vaurio's estimate displays a non-conservativism when the empirical failure rate at the plant of interest is much lower than the empirical failure rates of other plants in the population. The two-stage Dirichlet model (and presumably the other two stage models) are more sensitive to the empirical failure rates at 6 other plants.

Finally, we include the Dirichlet results for data sets 1 and 3 of [Becker and Hofer 2001]. The results are in the same order as those for classical Bayesian models [Cooke et al 2002], but tend to be a bit more conservative.

Figure 2 shows the density for λ_0 with and without updating from other plants. We see that the other plants have a (weak) tendency to lower the failure rate at plant 0.

Table 6.

Plant	T	X	X/T	
1	10000	0	0	
2	10000	0	0	
3	10000	0	0	
4	10000	0	0	
5	10000	0	0	
0	10000	5	0.0005	
L0	L1	L2	L3	L4
1.00E-8	5.00E-6	5.00E-5	5.00E-4	1.00E-3
	Ex w.	5%	50%	95%
Dir	5.03E-4	2.03E-4	4.16E-4	9.37E-4
Vaurio	4.87E-4			
Plant	T	X	X/T	
1	10000	5	0.0005	
2	10000	5	0.0005	
3	10000	5	0.0005	
4	10000	5	0.0005	
5	10000	5	0.0005	
0	10000	0	0.0005	
L0	L1	L2	L3	L4
1.00E-8	5.00E-6	5.00E-5	5.00E-4	1.00E-3
	Ex w.	5%	50%	95%
Dir	1.30E-4	1.81E-5	1.02E-4	3.17E-4
Vaurio	5.79E-5			
Plant	T	X	X/T	
1	1000	0	0	
2	1000	0	0	
3	1000	0	0	
4	1000	0	0	
5	1000	0	0	
0	1000	5	0.005	
L0	L1	L2	L3	L4
1E-8	5.00E-6	5.00E-5	5.00E-4	1.00E-2
	Ex w.	5%	50%	95%
Dir	5.61E-3	2.54E-3	5.46E-3	9.05E-3
Vaurio	4.87E-3			
Plant	T	X	X/T	
1	1000	5	0.005	
2	1000	5	0.005	
3	1000	5	0.005	
4	1000	5	0.005	
5	1000	5	0.005	
0	1000	0	0.005	
L0	L1	L2	L3	L4
1E-8	5.00E-6	5.00E-5	5.00E-4	1.00E-2
	Ex w.	5%	50%	95%
Dir	1.41E-3	4.39E-4	1.09E-3	3.39E-3
Vaurio	5.79E-4			

6 CONCLUSIONS

The Dirichlet model enjoys two advantages relative to the other models discussed here. First, it is analytically

Table 7. Results for datasets 1 and 3.

Dataset 1	5%	50%	95%
Dirichelet	1.12E-05	7.79E-05	2.34E-04
Gamma model	3.25E-05	6.99E-05	1.30E-04
Lognormal model	2.01E-05	5.91E-05	1.44E-04
Vaurio's model	1.83E-05	6.84E-05	1.73E-04
Dataset 3	5%	50%	95%
Dirichelet	1.45E-04	4.75E-04	9.25E-04
Gamma model	1.21E-04	2.58E-04	7.41E-04
Lognormal model	1.12E-05	6.23E-05	3.40E-04
Vaurio's model	2.84E-05	2.19E-04	7.64E-0.4

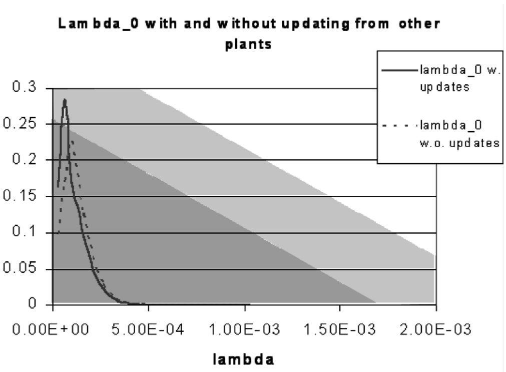

Figure 2. Dirichlet model, failure rates with and without updating from other plants, dataset 1.

solvable, no numerical integration need to be performed, and second, it allows an intuitive interpretation of the (hyper)parameters – so called "equivalent observation". It must be emphasized that this is a first implementation of this model. Additional testing should be performed before declaring the model fully operational. In particular, heuristics should be developed for choosing the number of cells C_j, and the values of their endpoints. Preliminary calculations show some sensitivity of the model in this regard.

REFERENCES

Becker G. & Hofer E. 2001. *Schtzung von Zuverlssigheits-Kennwerten aus der Evidenz veschiedener Anlagen*, GRS-A-2918.

Cooke R.M., Bunea C., Charitos T. & Mazzuchi T.A. 2002. *Mathematical Review of ZEDB Two-Stage Bayesian Models*, Report 02-45, Departement of Mathematics, Delft University of Technology.

Cooke R.M., Dorrepaal J. & Bedford T.J. 1995. *Review of SKI Data Processing Methodology* SKI Report.

Hofer E. 1999. On two-stage Bayesian modeling of initiating event frequencies and failure rates, *Reliability Engineering and System Safety*, vol. 66, pg. 97–99.

Hofer E., Hora S., Iman R. & Peschke J. 1997. On the solution approach for Bayesian modeling of initiating event frequencies and failure rates, *Risk Analysis*, vol. 17, no. 2, pg. 249–252.

Vaurio, J.K. 1987. On analytic empirical Bayes Estimation of failure rates, *Risk Analysis*, vol. 17, no.3, pg. 329–338.