A ROBUST BAYESIAN APPROACH TO MODELLING EPISTEMIC UNCERTAINTY IN COMMON-CAUSE FAILURE MODELS

MATTHIAS C. M. TROFFAES, GERO WALTER, AND DANA KELLY

ABSTRACT. In a standard Bayesian approach to the alpha-factor model for common-cause failure, a precise Dirichlet prior distribution models epistemic uncertainty in the alpha-factors. This Dirichlet prior is then updated with observed data to obtain a posterior distribution, which forms the basis for further inferences.

In this paper, we adapt the imprecise Dirichlet model of Walley to represent epistemic uncertainty in the alpha-factors. In this approach, epistemic uncertainty is expressed more cautiously via lower and upper expectations for each alpha-factor, along with a learning parameter which determines how quickly the model learns from observed data. For this application, we focus on elicitation of the learning parameter, and find that values in the range of 1 to 10 seem reasonable. The approach is compared with Kelly and Atwood's minimally informative Dirichlet prior for the alpha-factor model, which incorporated precise mean values for the alpha-factors, but which was otherwise quite diffuse.

Next, we explore the use of a set of Gamma priors to model epistemic uncertainty in the marginal failure rate, expressed via a lower and upper expectation for this rate, again along with a learning parameter. As zero counts are generally less of an issue here, we find that the choice of this learning parameter is less crucial.

Finally, we demonstrate how both epistemic uncertainty models can be combined to arrive at lower and upper expectations for all common-cause failure rates. Thereby, we effectively provide a full sensitivity analysis of common-cause failure rates, properly reflecting epistemic uncertainty of the analyst on all levels of the common-cause failure model.

1. INTRODUCTION

Common-cause failure has been recognized since the time of the Reactor Safety Study [12] as a dominant contributor to the unreliability of redundant systems. A number of models have been developed for common-cause failure over the time since the publication of the Reactor Safety Study, with perhaps the most widely used one being the Basic Parameter Model, at least in the U.S. [8].

The alpha-factor parametrisation of this model uses a multinomial distribution as its aleatory model for observed failures [8]. The conjugate prior to the multinomial model is the Dirichlet distribution. In the standard Bayesian approach, the analyst specifies the parameters of a precise Dirichlet distribution to model epistemic uncertainty in the alpha-factors, which are the parameters of the multinomial aleatory model. This Dirichlet prior is then updated with observed data to obtain a precise posterior distribution, also Dirichlet.

In this paper, we follow Troffaes et al. [11], and adapt the imprecise Dirichlet model of Walley [14] to represent epistemic uncertainty in the alpha-factors. In this approach the analyst specifies lower or upper expectations (or both) for each alpha-factor, along with a learning parameter, which determines how quickly the prior distribution learns from observed data. As the meaning of this learning parameter is less tangible, we discuss reasonable values for the learning parameter in this application.

Following Troffaes et al. [11], the approach is compared with that of Kelly and Atwood [7], which attempted to find a precise Dirichlet prior that was minimally informative [2], in the sense that it incorporated specified mean values for the alpha-factors, but which was otherwise quite diffuse. The numerical example from Kelly and Atwood [7] is addressed in the imprecise Dirichlet framework, which can be seen as an extension of their approach to the case where a precise mean for each alpha-factor cannot be specified.

Finally, we address the problem—not discussed in Troffaes et al. [11]—of inference about actual failure rates. These failure rates are rational functions of the alpha-factors and the marginal failure rate per component. Modelling failures as a Poisson process, we take a Gamma distribution as conjugate prior for the marginal failure rate. Similar to the procedure for the alpha-factors, we can model epistemic uncertainty on the marginal failure rate by considering lower and upper expected prior failure rates, along with a learning parameter that determines how quickly the prior distribution learns from observed data.

By combining our epistemic uncertainty models for both the alpha-factors and the marginal failure rate, we are able to perform a global sensitivity analysis on the common-cause failure rates. We provide an algorithm that calculates, up to reasonable precision, bounds on these failure rates. The resulting novel procedure is demonstrated on a simple electrical network reliability problem.

The paper is organized as follows. Section 2 reviews the basic parameter model and its reparametrisation as the alpha-factor model. Section 3 explores how the parameters of the alpha-factor model can be estimated, using Dirichlet and Gamma priors. Section 4 discusses the handling of epistemic uncertainty for the alpha-factors. Two ways to choose a Dirichlet prior (or sets of Dirichlet priors) starting from epistemic prior expectations of the alpha-factors are considered. Throughout, the main ideas are demonstrated on a numerical example. Section 5 shows how, similarly to the alpha-factor case, epistemic uncertainty can be expressed for the marginal failure rate. A set of conjugate Gamma priors is elicited by considering lower and upper expected prior marginal failure rates. Section 6 describes an algorithm that infers bounds on all common-cause failure rates based on our imprecise alpha-factor model and our imprecise marginal failure rate model. Section 7 demonstrates our methodology on a simple electrical network reliability problem. Section 8 ends the paper with some conclusions and thoughts for further research.

2. Common-Cause Failure Modelling

2.1. **The Basic Parameter Model.** Consider a system that consists of k components. Throughout, we make the following standard assumptions: (i) repair is immediate, and (ii) failures follow a Poisson process.

For simplicity, we assume that all k components are exchangeable, in the sense that they have identical failure rates. More precisely, we assume that all events involving *exactly j* components failing have the same failure rate, which we denote by q_j. This model is called the *basic parameter model* [8], and we write $\boldsymbol{q}$ for $(q_1, \ldots, q_k)$.

For example, if we have three components, A, B, and C, then the rate at which we see only A failing is equal to the rate at which we see only B failing, and is also equal to the rate at which we see only C failing; this failure rate is q_1. Moreover, the rate at which we observe only A and B jointly failing is equal to the rate at which we observe only B and C jointly failing, and also equal to the rate at which we observe only A and C jointly failing; this failure rate is q_2. The rate at which we see all three components jointly failing is q_3.

In case of k identical components without common-cause failure modes, thus each failing independently at rate λ, we would have[1]

$$(1) \qquad q_1 = \lambda \qquad \text{and} \qquad q_j = 0 \text{ for } j \geq 2.$$

The fact that we allow arbitrary values for the q_j reflects the lack of independence, and whence, our modelling of common-cause failures. At this point, it is worth noting that we do not actually write down a statistical model for all possible common-cause failure modes—we could do so if this information was available, and in fact, this could render the basic parameter model obsolete, and allow for more detailed inferences. In essence, the basic parameter model allows us to statistically model lack of independence between component failures, without further detail as to where dependencies arise from: all failure modes are lumped together, so to speak.

It is useful to note that it is possible, and sometimes necessary, to relax the exchangeability assumption to accommodate specific asymmetric cases. For example, when components are in different state of health, single failures would clearly not have identical failure rates. A preliminary investigation of the asymmetric case, for two components, can be found in Troffaes and Blake [10]. In a nutshell, the analysis is very similar, but the formulas become somewhat more complicated because there are more variables to keep track of. Therefore, we stick to the exchangeable case here.

Clearly, to answer typical reliability questions, such as for instance "what is the probability that two or more components fail in the next month?", we need q. In practice, the following three issues commonly arise. First, q is rarely measured directly, as failure data is often collected only per component. Secondly, when direct data about joint failures is available, typically, this data is sparse, because events involving more than two components failing simultaneously are usually quite rare. Thirdly, there are usually two distinct sources of failure data, one usually very large data set related to failure per component, and one usually much smaller data set related to joint failures. For these reasons, it is sensible to reparametrise the model in terms of parameters that can be more easily estimated, as follows.

2.2. **The Alpha-Factor Model.** The alpha-factor parametrisation of the basic parameter model [8] starts out with considering the total failure rate of a component q_t, which could involve failure of any number of components, that is, this is the rate obtained by looking at just one component, ignoring everything else. Clearly,

$$(2) \qquad q_t = \sum_{j=1}^{k} \binom{k-1}{j-1} q_j.$$

For example, again consider a three component system, A, B, and C. The rate at which A fails is then the rate at which only A fails (q_1), plus the rate at which A and B, or A and C fail ($2q_2$), plus the rate at which all three components fail (q_3).

Next, the alpha-factor model introduces α_j—the so-called alpha-factor—which denotes the probability of *exactly* j of the k components failing given that failure occurs; in terms of relative frequency, α_j is the fraction of failures that involve *exactly* j failed components. We write $\boldsymbol{\alpha}$ for $(\alpha_1, \ldots, \alpha_k)$. Clearly,

$$(3) \qquad \alpha_j = \frac{\binom{k}{j} q_j}{\sum_{\ell=1}^{k} \binom{k}{\ell} q_\ell}.$$

For example, again consider A, B, and C. Then the rate at which exactly one component fails is $3q_1$ (as we have three single components, each of which failing with rate q_1), the rate at which

---

[1]This is due to our Poisson assumption, and the assumption of immediate repair: independent Poisson processes never generate events simultaneously when we observe failure times precisely.

exactly two components fail is $3q_2$ (as we have three combinations of two components, each combination failing with rate q_2), and the rate at which all components fail is q_3. Translating these rates into fractions, we arrive precisely at Eq. (3).

It can be shown that [8, Table C-1, p. C-5]:[2]

$$(4) \qquad q_j = \frac{1}{\binom{k-1}{j-1}} \frac{j\alpha_j}{\sum_{\ell=1}^{k} \ell\alpha_\ell} q_t.$$

Eqs. (2), (3), and (4) establish a one-to-one link between the so-called basic parameter model ($\boldsymbol{q}$) and the alpha-factor model (q_t, $\boldsymbol{\alpha}$). The benefit of the alpha-factor model over the basic parameter model lies in its distinction between the total failure rate of a component q_t, for which we generally have a lot of information, and common-cause failures modelled by $\boldsymbol{\alpha}$, for which we generally have very little information.

One of the goals of this paper is to perform a sensitivity analysis, in the sense of robust Bayes [3, 4, 13], over $\boldsymbol{\alpha}$, and to measure its effects on q_j. Because the q_j are proportional to q_t, in fact, it turns out to take only very little additional effort to perform a sensitivity analysis over $\boldsymbol{\alpha}$ and q_t jointly. So, although in many cases of practical interest, we will know q_t quite well, interestingly, we do not need to assume that we know much at all about q_t.

3. Parameter Estimation

3.1. **Dirichlet Prior for Alpha-Factors.** Suppose that we have observed a sequence of N failure events, where we have counted the number of components involved with each failure event, say n_j of the N observed failure events involved *exactly* j failed components. We write $\boldsymbol{n}$ for $(n_1, \ldots, n_k)$. In terms of the alpha-factors, the likelihood for $\boldsymbol{n}$ has a very simple form:

$$(5) \qquad \Pr(\boldsymbol{n}|\boldsymbol{\alpha}) = \prod_{j=1}^{k} \alpha_j^{n_j},$$

which is a multinomial distribution with parameter $\boldsymbol{\alpha}$.

As mentioned already, typically, for $j \geq 2$, the n_j are very low, with zero being quite common for larger j. In such cases, standard techniques such as maximum likelihood for estimating the alpha-factors fail to produce sensible inferences. For any inference to be reasonably possible, it has been recognized [8] that we have to rely on epistemic information, that is, information which is not just described by the data.

A standard way to include epistemic information in the model is through specification of a Dirichlet prior for the alpha-factors [8]:

$$(6) \qquad f(\boldsymbol{\alpha}|s, \boldsymbol{t}) \propto \prod_{j=1}^{k} \alpha_j^{st_j - 1}$$

which is a conjugate prior for the multinomial likelihood specified in Eq. (5). In Eq. (6), we use Walley's [13, §7.7.3, p. 395] $(s, \boldsymbol{t})$ notation for the hyperparameters. Here, $s > 0$ and $\boldsymbol{t} \in \Delta$, where Δ is the $(k-1)$-dimensional unit simplex:

$$(7) \qquad \Delta = \left\{ (t_1, \ldots, t_k) \colon t_1 \geq 0, \ldots, t_k \geq 0, \sum_{j=1}^{k} t_j = 1 \right\}$$

---

[2]Hint: consider $\sum_{j=1}^{k} j\alpha_j$.

An interpretation for these parameters will be given shortly. First, let us calculate the posterior density for $\boldsymbol{\alpha}$:

$$(8) \qquad f(\boldsymbol{\alpha}|\boldsymbol{n}, s, \boldsymbol{t}) \propto \prod_{j=1}^{k} \alpha_j^{st_j + n_j - 1}.$$

Of typical interest is for instance the posterior expectation of the probability α_j of observing j of the k components failing due to a common cause given that failure occurs:

$$(9) \qquad E(\alpha_j|\boldsymbol{n}, s, \boldsymbol{t}) = \int_\Delta \alpha_j f(\boldsymbol{\alpha}|\boldsymbol{n}, s, \boldsymbol{t}) \, \mathrm{d}\boldsymbol{\alpha} = \frac{n_j + st_j}{N + s} = \frac{N}{N + s}\frac{n_j}{N} + \frac{s}{N + s}t_j$$

where $N = \sum_{j=1}^{k} n_j$ is the total number of observations.

Eq. (9) provides the usual well-known interpretation for the hyperparameters s and $\boldsymbol{t}$:

- If $N = 0$, then $E(\alpha_j|s, \boldsymbol{t}) = t_j$, so t_j is the prior expected chance of observing j of the k components failing due to a common cause, given that failure occurs.
- $E(\alpha_j|\boldsymbol{n}, s, \boldsymbol{t})$ is a weighted average of t_j and n_j/N (the proportion of j-component failures in the N observations), with weights s and N, respectively. The parameter s thus determines how much data is required for the posterior to start moving away from the prior. If $N \ll s$ then the prior will weigh more; if $N = s$, then prior and data will weigh equally; and if $N \gg s$, then the data will weigh more. In particular, $E(\alpha_j|\boldsymbol{n}, s, \boldsymbol{t}) = t_j$ if $N = 0$ (as already mentioned), and $E(\alpha_j|\boldsymbol{n}, s, \boldsymbol{t}) \to \frac{n_j}{N}$ as $N \to \infty$.

For inference about q_j, which we will discuss in Section 6, we will also need, for natural numbers $p_1, \ldots, p_k$, with $P := \sum_{j=1}^{k} p_j$:

$$(10) \qquad E\left(\prod_{j=1}^{k} \alpha_j^{p_j}|\boldsymbol{n}, s, \boldsymbol{t}\right) = \frac{\prod_{j=1}^{k}(n_j + st_j)_{p_j}}{(N + s)_P}.$$

where $(x)_n$, for $n \in \mathbb{N}_0$, denotes the raising factorial, also known as Pochhammer's symbol [1, 6.1.22, p. 256]:

$$(11) \qquad (x)_n := \frac{\Gamma(x + n)}{\Gamma(x)} = (x + n - 1)(x + n - 2)\ldots(x + 1)x.$$

By linearity of expectation, Eq. (10) allows us to calculate the expectation of an arbitrary polynomial in $\boldsymbol{\alpha}$.

3.2. **Per Component Failure Rate.** Now we turn to the estimation of q_t, the total failure rate per component. As mentioned at the start of Section 2, we assume that failures follow a Poisson process. Suppose we observe M failures of our component over a time interval of length T. If M is sufficiently large, then a reasonable point estimate for q_t would be M/T.

Often, that will be enough. However, in case M is not terribly large, we can easily propose a conjugate prior for q_t. Specifically, the likelihood for M, given T, is:

$$(12) \qquad \Pr(M|q_t, T) = \frac{(q_t T)^M e^{-q_t T}}{M!}$$

which is simply a Poisson distribution with parameter $q_t T$.

A standard way to include epistemic information in the model is through specification of a Gamma prior [5, 9]:[3]

$$(13) \qquad f(q_t|u, v) \propto q_t^{uv-1} e^{-q_t u},$$

---

[3]We use a non-standard parametrisation to allow easier interpretation of the hyperparameters.

which is a conjugate prior for the Poisson likelihood specified in Eq. (12). The posterior density for q_t is:

$$(14) \qquad f(q_t|M,T,u,v) \propto q_t^{uv+M-1} e^{-q_t(u+T)}$$

Of typical interest is the posterior expectation of q_t:

$$E(q_t|M,T,u,v) = \int_{u,v} f(q_t|M,T,u,v)\,\mathrm{d}u\,\mathrm{d}v = \frac{M+uv}{u+T}$$

$$(15) \qquad\qquad = \frac{T}{u+T}\frac{M}{T} + \frac{u}{u+T}v$$

Eq. (15) provides a straightforward interpretation for the hyperparameters u and v, which mimicks our discussion concerning the Dirichlet prior:[4]

- If $T=0$, then $E(q_t|u,v) = v$, so v is the prior expected failure rate.
- $E(q_t|M,T,u,v)$ is a weighted average of v and M/T (the empirical observed failure rate), with weights u and T, respectively. The parameter u thus determines for how long we need to observe the process until the posterior starts to move away from the prior. If $T \ll u$ then the prior will weigh more; if $T = u$, then prior and data will weigh equally; and if $T \gg u$, then the data will weigh more. In particular, $E(q_t|M,T,u,v) = v$ if $T = 0$ (as already mentioned), and $E(q_t|M,T,u,v) \to \frac{M}{T}$ as $T \to \infty$.

4. Handling Epistemic Uncertainty in Alpha-Factors

Crucial to reliable inference in the alpha-factor model is proper modelling of epistemic uncertainty about failures, which is in the above approach expressed through the $(s, \boldsymbol{t})$ parameters. We focus on two methods for elicitation of these parameters, and the inferences that result from them.

Throughout, we will use the following example, which is taken from Kelly and Atwood [7]. Consider a system with four redundant components ($k = 4$). The probability of j out of k failures, given that failure has happened, was denoted by α_j. We assume that the analyst's prior expectation $\mu_{\text{spec},j}$ for each α_j is:

$$(16) \qquad \mu_{\text{spec},1} = 0.950 \qquad \mu_{\text{spec},2} = 0.030 \qquad \mu_{\text{spec},3} = 0.015 \qquad \mu_{\text{spec},4} = 0.005$$

We have 36 observations, in which 35 showed one component failing, and 1 showed two components failing:

$$n_1 = 35 \qquad n_2 = 1 \qquad n_3 = 0 \qquad n_4 = 0$$

4.1. **Constrained Non-Informative Prior.** Atwood [2] studied priors for the binomial model which maximise entropy (and whence, are 'non-informative') whilst constraining the mean to a specific value. Although these priors are not conjugate, Atwood [2] showed that they can be well approximated by Beta distributions, which are conjugate. Kelly and Atwood [7] applied this approach to the multinomial model with conjugate Dirichlet priors, by choosing a constrained non-informative prior for the marginals of the Dirichlet—which are Beta. This leads to an over-specified system of equalities, which can be solved via least-squares optimisation.

For the problem we are interested in, $\mu_{\text{spec},1}$ is close to 1. In this case, the solution of the least-squares problem turns out to be close to:

$$t_j = \mu_{\text{spec},j} \text{ for all } j \in \{1,\dots,k\}$$

$$(17) \qquad\qquad s = \frac{1}{2(1-\mu_{\text{spec},1})}$$

---

[4]In fact, we arrive at similar interpretations because both priors are members of the canonical exponential family [5, 9].

For our example, this means that $s = 10$ [7, p. 400, §3]. An obvious calculation reveals that, under this prior [7, p. 401, §3.1]:

$$E(\alpha_1|\boldsymbol{n}, s, \boldsymbol{t}) = \frac{35 + 9.5}{36 + 10} = 0.967 \qquad E(\alpha_2|\boldsymbol{n}, s, \boldsymbol{t}) = \frac{1 + 0.3}{36 + 10} = 0.028$$

$$E(\alpha_3|\boldsymbol{n}, s, \boldsymbol{t}) = \frac{0 + 0.15}{36 + 10} = 0.003 \qquad E(\alpha_4|\boldsymbol{n}, s, \boldsymbol{t}) = \frac{0 + 0.05}{36 + 10} = 0.001$$

Kelly and Atwood [7, p. 402, §4] compare these results against a large number of other choices of priors, and note that the posterior resulting from Eq. (17) seems too strongly influenced by the prior, particularly in the presence of zero counts. For instance, the uniform prior is a Dirichlet distribution with hyperparameters $t_j = 0.25$ and $s = 4$, which gives:

$$E(\alpha_1|\boldsymbol{n}, s, \boldsymbol{t}) = \frac{35 + 1}{36 + 4} = 0.9 \qquad E(\alpha_2|\boldsymbol{n}, s, \boldsymbol{t}) = \frac{1 + 1}{36 + 4} = 0.05$$

$$E(\alpha_3|\boldsymbol{n}, s, \boldsymbol{t}) = \frac{0 + 1}{36 + 4} = 0.025 \qquad E(\alpha_4|\boldsymbol{n}, s, \boldsymbol{t}) = \frac{0 + 1}{36 + 4} = 0.025$$

Jeffrey's prior is again a Dirichlet distribution with hyperparameters $t_j = 0.125$ and $s = 4$, which gives:

$$E(\alpha_1|\boldsymbol{n}, s, \boldsymbol{t}) = \frac{35 + 0.5}{36 + 4} = 0.8875 \qquad E(\alpha_2|\boldsymbol{n}, s, \boldsymbol{t}) = \frac{1 + 0.5}{36 + 4} = 0.0375$$

$$E(\alpha_3|\boldsymbol{n}, s, \boldsymbol{t}) = \frac{0 + 0.5}{36 + 4} = 0.0125 \qquad E(\alpha_4|\boldsymbol{n}, s, \boldsymbol{t}) = \frac{0 + 0.5}{36 + 4} = 0.0125$$

The degree of variation in the posterior under different priors is evidently somewhat alarming. In the next section, we aim to robustify the model by using sets of priors from the start.

4.2. Imprecise Dirichlet Model.

4.2.1. *Near-Ignorance Model.* In case no prior information is available, Walley proposes as a so-called *near-ignorance prior* a set of Dirichlet priors, with hyperparameters constrained to the set:

$$\mathcal{H} = \{(s, \boldsymbol{t}) : \boldsymbol{t} \in \Delta\}$$

for some fixed value of s, which determines the learning speed of the model [13, p. 218, §5.3.2] [14, p. 9, §2.3].

4.2.2. *General Model.* When prior information is available, more generally, we may assume that we can specify a subset $\mathcal{H}$ of $(0, +\infty) \times \Delta$. Following Walley's suggestions [13, p. 224, §5.4.3] [14, p. 32, §6], we take

$$(18) \qquad \mathcal{H} = \left\{(s, \boldsymbol{t}) : s \in [\underline{s}, \overline{s}], \, \boldsymbol{t} \in \Delta, \, t_j \in [\underline{t}_j, \overline{t}_j]\right\}$$

where the analyst has to specify the bounds $[\underline{t}_j, \overline{t}_j]$ for each $j \in \{1, \ldots, k\}$, and $[\underline{s}, \overline{s}]$.

The posterior lower and upper expectations of α_j are:

$$(19) \qquad \underline{E}(\alpha_j|\boldsymbol{n}, \mathcal{H}) = \min\left\{\frac{n_j + \underline{s}t_j}{N + \underline{s}}, \frac{n_j + \overline{s}t_j}{N + \overline{s}}\right\} = \begin{cases} \frac{n_j + \underline{s}t_j}{N + \underline{s}} & \text{if } \underline{t}_j \geq n_j/N \\ \frac{n_j + \overline{s}t_j}{N + \overline{s}} & \text{if } \underline{t}_j \leq n_j/N \end{cases}$$

$$(20) \qquad \overline{E}(\alpha_j|\boldsymbol{n}, \mathcal{H}) = \max\left\{\frac{n_j + \underline{s}\overline{t}_j}{N + \underline{s}}, \frac{n_j + \overline{s}\overline{t}_j}{N + \overline{s}}\right\} = \begin{cases} \frac{n_j + \overline{s}\overline{t}_j}{N + \overline{s}} & \text{if } \overline{t}_j \geq n_j/N \\ \frac{n_j + \underline{s}\overline{t}_j}{N + \underline{s}} & \text{if } \overline{t}_j \leq n_j/N \end{cases}$$

For the model to be of any use, we must be able to elicit the bounds. The interval $[\underline{t}_j, \overline{t}_j]$ simply represents bounds on the prior expectation of the chance α_j.

Fixed Learning Parameter. Typically, the learning parameter s is taken to be 2 (not without controversy; see insightful discussions in [14]). One might therefore be tempted to using the same prior expectations t_j for the α_j as above (Eq. (16)), with $s = 2$, resulting in the following posterior expectations:

$$E(\alpha_1|\boldsymbol{n}, s, \boldsymbol{t}) = \frac{35 + 1.9}{36 + 2} = 0.971 \qquad E(\alpha_2|\boldsymbol{n}, s, \boldsymbol{t}) = \frac{1 + 0.06}{36 + 2} = 0.028$$

$$E(\alpha_3|\boldsymbol{n}, s, \boldsymbol{t}) = \frac{0 + 0.03}{36 + 2} = 0.0007 \qquad E(\alpha_4|\boldsymbol{n}, s, \boldsymbol{t}) = \frac{0 + 0.01}{36 + 2} = 0.0002$$

Whence, for this example, it is obvious that $s = 2$ is an excessively poor choice: the posterior expectations in case of zero counts are pulled way too much towards zero. One might suspect that this is partly due to the strong prior information, that is, the knowledge of t_j. However, even if we interpret the given probabilities as bounds, say:

$$\text{(21a)} \qquad\qquad [\underline{t}_1, \bar{t}_1] = [0.950, 1]$$

$$\text{(21b)} \qquad\qquad [\underline{t}_2, \bar{t}_2] = [0, 0.030]$$

$$\text{(21c)} \qquad\qquad [\underline{t}_3, \bar{t}_3] = [0, 0.015]$$

$$\text{(21d)} \qquad\qquad [\underline{t}_4, \bar{t}_4] = [0, 0.005]$$

we still find:

$$\text{(22a)} \qquad\qquad [\underline{E}(\alpha_1|\boldsymbol{n}, \mathcal{H}), \overline{E}(\alpha_1|\boldsymbol{n}, \mathcal{H})] = [0.971, 0.974]$$

$$\text{(22b)} \qquad\qquad [\underline{E}(\alpha_2|\boldsymbol{n}, \mathcal{H}), \overline{E}(\alpha_2|\boldsymbol{n}, \mathcal{H})] = [0.026, 0.028]$$

$$\text{(22c)} \qquad\qquad [\underline{E}(\alpha_3|\boldsymbol{n}, \mathcal{H}), \overline{E}(\alpha_3|\boldsymbol{n}, \mathcal{H})] = [0, 0.0007]$$

$$\text{(22d)} \qquad\qquad [\underline{E}(\alpha_4|\boldsymbol{n}, \mathcal{H}), \overline{E}(\alpha_4|\boldsymbol{n}, \mathcal{H})] = [0, 0.0002]$$

Clearly, only the posterior inferences about α_1 (and perhaps also α_2) seem reasonable. We conclude that *the imprecise Dirichlet model with $s = 2$ learns too fast from the data in case of zero counts.*

On the one hand, when counts are sufficiently far from zero, the posterior probability with $s = 2$, and perhaps even $s = 1$ or $s = 0$, seem appropriate. For zero counts, however, a larger value of s seems mandatory. Therefore, it seems logical to pick an interval for s.

A further argument for choosing an interval for s, in case of an informative set of priors, is provided by Walley [13, p. 225, §5.4.4]: a larger value of $\bar{s}$ ensures that the posterior does not move away too fast from the prior, which is particularly important for zero counts, and the difference between $\underline{s}$ and $\bar{s}$ effectively results in greater posterior imprecision if $n_j/N \notin [\underline{t}_j, \bar{t}_j]$.

To see this, note that, if $\underline{t}_j \le n_j/N \le \bar{t}_j$, it follows from Eqs. (19) and (20) that both lower and upper posterior expectation are calculated using $\bar{s}$. When $n_j/N \le \underline{t}_j$ (or $\bar{t}_j \le n_j/N$), the lower (upper) posterior expectation is calculated using $\underline{s}$ instead, which is nearer to n_j/N due to the lower weight $\underline{s}$ for the prior bound $\underline{t}_j$ ($\bar{t}_j$). The increased imprecision reflects the conflict between the prior assignment $[\underline{t}_j, \bar{t}_j]$ and the observed fraction n_j/N, and this is referred to as *prior-data conflict* (also see [15]).

Interval for Learning Parameter. We follow Good [6, p. 19] (as suggested by Walley [13, Note 5.4.1, p. 524]), and reason about posterior expectations of hypothetical data to elicit $\underline{s}$ and $\bar{s}$; also see [13, p. 219, §5.3.3] for further discussion on elicitation on s—our approach is similar, but simpler for the case under study. We assume that $\bar{t}_1 = 1$ and $\underline{t}_j = 0$ for all $j \ge 2$.

The upper probability of multiple ($j \geq 2$) failed components in trial $m + 1$, given one ($j = 1$) failed component in all of the first m trials, is

$$\overline{E}(\alpha_j | n_1 = m, N = m, \mathcal{H}) = \frac{\overline{s}\overline{t}_j}{m + \overline{s}}$$

(Note: there is no prior-data conflict in this case.) Whence, for the above probability to reduce to $\overline{t}_j/2$ (i.e., to reduce the prior upper probability by half), we need that $m = \overline{s}$. In other words, $\overline{s}$ *is the number of one-component failures required to reduce the upper probabilities of multi-components failure by half.*

Conversely, the lower probability of one ($j = 1$) failed component in trial $m + 1$, given only multiple ($j \geq 2$) failed components in the first m trials, is

$$\underline{E}(\alpha_1 | n_1 = 0, N = m, \mathcal{H}) = \frac{\underline{s}\underline{t}_1}{m + \underline{s}}$$

(Note: there is strong prior-data conflict in this case.) In other words, $\underline{s}$ *is the number of multi-component failures required to reduce the lower probability of one-component failure by half.* Note that, in this case, a few alternative interpretations present themselves. First, for $j \geq 2$,

$$\overline{E}(\alpha_j | n_j = m, N = m, \mathcal{H}) = \frac{m + \underline{s}\overline{t}_j}{m + \underline{s}}$$

In other words, $\underline{s}$ *is also the number of j-component failures required to increase the upper probability of j components failing to* $(1 + \overline{t}_j)/2$ *(generally, this will be close to $1/2$, provided that $\overline{t}_j$ is close to zero). Secondly, for $j \geq 2$,*

$$\underline{E}(\alpha_j | n_j = m, N = m, \mathcal{H}) = \frac{m}{m + \overline{s}}$$

so $\overline{s}$ *is also the number of multi-component failures required to increase the lower probability of multi-component failures to a half.*

Any of these counts seem well suited for elicitation, and are easy to interpret. As a guideline, we suggest the following easily remembered rules:

- $\overline{s}$ is the number of one-component failures required to reduce the upper probabilities of multi-component failures by half, and
- $\underline{s}$ is the number of multi-component failures required to reduce the lower probability of one-component failures by half.

Taking the above interpretation, the difference between $\overline{s}$ and $\underline{s}$ reflects the fact that the rate at which we reduce upper probabilities is less than the rate at which we reduce lower probabilities, and thus reflects a level of caution in our model.

Coming back to our example, reasonable values are $\underline{s} = 1$ (if we immediately observe multi-component failures, we might be quite keen to reduce our lower probability for one-component failure) and $\overline{s} = 10$ (we are happy to halve our upper probabilities of multi-component failures after observing 10 one-component failures). With these values, when taking for t_j the values given in Eq. (16), we find the following posterior lower and upper expectations of α_j:

(23a) $$[\underline{E}(\alpha_1 | \boldsymbol{n}, \mathcal{H}), \overline{E}(\alpha_1 | \boldsymbol{n}, \mathcal{H})] = [0.967, 0.972]$$

(23b) $$[\underline{E}(\alpha_2 | \boldsymbol{n}, \mathcal{H}), \overline{E}(\alpha_2 | \boldsymbol{n}, \mathcal{H})] = [0.0278, 0.0283]$$

(23c) $$[\underline{E}(\alpha_3 | \boldsymbol{n}, \mathcal{H}), \overline{E}(\alpha_3 | \boldsymbol{n}, \mathcal{H})] = [0.00041, 0.00326]$$

(23d) $$[\underline{E}(\alpha_4 | \boldsymbol{n}, \mathcal{H}), \overline{E}(\alpha_4 | \boldsymbol{n}, \mathcal{H})] = [0.00014, 0.00109]$$

These bounds indeed reflect caution in inferences where zero counts have occurred ($j = 3$ and $j = 4$), with upper expectations considerably larger as compared to the model with fixed s, while still giving a reasonable expectation interval for the probability of one-component failure.

If we desire to specify our initial bounds for t_j more conservatively, as in Eqs. (21), we find similar results:

$$\text{(24a)} \qquad [\underline{E}(\alpha_1|\boldsymbol{n},\mathcal{H}), \overline{E}(\alpha_1|\boldsymbol{n},\mathcal{H})] = [0.967, 0.978]$$

$$\text{(24b)} \qquad [\underline{E}(\alpha_2|\boldsymbol{n},\mathcal{H}), \overline{E}(\alpha_2|\boldsymbol{n},\mathcal{H})] = [0.0270, 0.0283]$$

$$\text{(24c)} \qquad [\underline{E}(\alpha_3|\boldsymbol{n},\mathcal{H}), \overline{E}(\alpha_3|\boldsymbol{n},\mathcal{H})] = [0, 0.00326]$$

$$\text{(24d)} \qquad [\underline{E}(\alpha_4|\boldsymbol{n},\mathcal{H}), \overline{E}(\alpha_4|\boldsymbol{n},\mathcal{H})] = [0, 0.00109]$$

5. Handling Epistemic Uncertainty in Marginal Failure Rate

Before we can consider inferences on the common-cause failure rates q_j, we will briefly explain how we express epistemic uncertainty on the marginal failure rate q_t. As seen in Section 3.2, we will use conjugate Gamma priors with hyperparameters u and v, where v is the prior failure rate parameter, and u determines the learning speed. Similarly to the alpha-factor case, we can express vague prior information on q_t by considering sets of priors, which are generated by sets of hyperparameters, i.e., we specify a parameter set $\mathcal{J} \subseteq (0,\infty) \times (0,\infty)$. Unlike Section 4.2.1, here $\mathcal{J} = \{u\} \times (0,\infty)$, for some fixed value of u, does not lead to a practically useful near-ignorant set of priors, as then $\overline{E}(q_t|M,T,\mathcal{J}) = \infty$ for any M and T. In practice, it should not be a big issue to find bounds $[\underline{v}, \overline{v}]$ for the prior expected marginal failure rate.

Similarly to Eqs. (19) and (20), when $\mathcal{J} = [\underline{u}, \overline{u}] \times [\underline{v}, \overline{v}]$, the posterior lower and upper expectations of q_t are

$$\text{(25)} \qquad \underline{E}(q_t|M,T,\mathcal{J}) = \min\left\{\frac{M+\underline{u}v}{T+\underline{u}}, \frac{M+\overline{u}\underline{v}}{T+\overline{u}}\right\} = \begin{cases} \frac{M+\underline{u}v}{T+\underline{u}} & \text{if } \underline{v} \geq M/T \\ \frac{M+\overline{u}\underline{v}}{T+\overline{u}} & \text{if } \underline{v} \leq M/T \end{cases}$$

$$\text{(26)} \qquad \overline{E}(q_t|M,T,\mathcal{J}) = \max\left\{\frac{M+\underline{u}\overline{v}}{T+\underline{u}}, \frac{M+\overline{u}v}{T+\overline{u}}\right\} = \begin{cases} \frac{M+\overline{u}v}{T+\overline{u}} & \text{if } \overline{v} \geq M/T \\ \frac{M+\underline{u}\overline{v}}{T+\underline{v}} & \text{if } \overline{v} \leq M/T \end{cases}$$

To elicit bounds for the learning parameter u, similar considerations as in Section 4.2.2 can be made. Assuming $\underline{v} = 0$, the posterior lower expectation for q_t is

$$\text{(27)} \qquad \underline{E}(q_t|M,T,\mathcal{J}) = \frac{M}{T+\overline{u}}$$

(Note: there is no prior-data conflict in this case.) Whence, $\overline{u}$ is the amount of time needed to observe the process until we raise the lower expectation of q_t from 0 to half of the observed failure rate M/T.

Conversely, assuming $\underline{v} > 0$, and no failures at all during time T, the posterior lower expectation for q_t is

$$\text{(28)} \qquad \underline{E}(q_t|M=0,T,\mathcal{J}) = \frac{\underline{u}v}{T+\underline{u}} = \frac{v}{\frac{T}{\underline{u}}+1}$$

(Note: prior-data conflict is present in this case.) Whence, $\underline{u}$ is the time needed to observe the process—without any failures—until $\underline{v}$ is reduced by half.

Contrary to the situation in Section 4, zero counts are much less of a concern when estimating the marginal failure rate. Whence, for sake of simplicity, it might therefore suffice to consider parameter sets of the form

$$\text{(29)} \qquad \mathcal{J} = \{u\} \times [\underline{v}, \overline{v}]$$

only. Both Eqs. (27) and (28) can then serve to determine $u = \underline{u} = \overline{u}$.

A numerical example will be given in Section 7.

x	$\frac{1}{1+x}$	$1-x$	$1-x+x^2$
0.0	1.0	1.0	1.0
0.1	0.91	0.9	0.91
0.2	0.83	0.8	0.84
0.3	0.77	0.7	0.79
0.4	0.71	0.6	0.76
0.5	0.67	0.5	0.75
0.6	0.63	0.4	0.76
0.7	0.59	0.3	0.79
0.8	0.56	0.2	0.84
0.9	0.53	0.1	0.91

TABLE 1. Accuracy of first and second order Taylor approximations.

6. INFERENCE ON FAILURE RATES

6.1. **Expected Failure Rates.** For inference on the failure rates q_j, we will now combine our models for alpha-factors and marginal failure rate by using Eq. (4). The problem in doing this is that there is, as far as we know, no immediate closed expression for the posterior expectation of q_j, because Eq. (4) is a rational function of $\boldsymbol{\alpha}$. However, naively, we can approximate it using Taylor expansion. Specifically,

$$(30) \qquad q_j = \frac{1}{\binom{k-1}{j-1}} \frac{j\alpha_j}{\sum_{\ell=1}^{k} \ell\alpha_\ell} q_t$$

$$(31) \qquad = \frac{1}{\binom{k-1}{j-1}} \frac{j\alpha_j}{\sum_{\ell=1}^{k} (\alpha_\ell + (\ell-1)\alpha_\ell)} q_t$$

$$(32) \qquad = \frac{1}{\binom{k-1}{j-1}} \frac{j\alpha_j}{1 + \sum_{\ell=2}^{k} (\ell-1)\alpha_\ell} q_t$$

and, as long as $\sum_{\ell=2}^{k} (\ell-1)\alpha_\ell < 1$—this is always true if $k \leq 2$; for larger k, it is usually true because α_ℓ is usually very small for $\ell \geq 3$—we can use the Taylor expansion $1/(1+x) = 1 - x + x^2 - x^3 + \dots$ (valid for $|x| < 1$), to arrive at:

$$(33) \qquad = \frac{1}{\binom{k-1}{j-1}} j\alpha_j \left[1 - \sum_{\ell=2}^{k} (\ell-1)\alpha_\ell + \left\{ \sum_{\ell=2}^{k} (\ell-1)\alpha_\ell \right\}^2 - \dots \right] q_t$$

The posterior expectation of Eq. (33) can now be evaluated, using Eqs. (10) and (15), under the usual assumption that q_t is independent of the alpha-factors.

To get a better idea of accuracy, Table 1 tabulates first and second order approximations. For example, second order approximation remains fairly accurate for $\sum_{\ell=2}^{k} (\ell-1)\alpha_\ell < 0.5$, and first order approximation for $\sum_{\ell=2}^{k} (\ell-1)\alpha_\ell < 0.3$.

An obvious issue with Taylor approximation is that the domain of integration includes values for $\boldsymbol{\alpha}$ where the Taylor series does not converge. However, it is easy to see that, for any $x \geq 0$ (not just those for which $|x| < 1$):

$$(34) \qquad 0 \leq (1 - x + x^2 - \dots + (-x)^p) - \frac{1}{1+x} \leq x^{p+1} \qquad \text{for even } p, \text{ and}$$

$$(35) \qquad 0 \leq \frac{1}{1+x} - (1 - x + x^2 - \dots + (-x)^p) \leq x^{p+1} \qquad \text{for odd } p.$$

Therefore, for any non-negative random variables x and y:

$$(36) \qquad 0 \leq E[y(1 - x + x^2 - \cdots + (-x)^p)] - E\left(\frac{y}{1+x}\right) \leq E(yx^{p+1}) \qquad \text{for even } p, \text{ and}$$

$$(37) \qquad 0 \leq E\left(\frac{y}{1+x}\right) - E[y(1 - x + x^2 - \cdots + (-x)^p)] \leq E(yx^{p+1}) \qquad \text{for odd } p.$$

So as long as the expectation of yx^{p+1} is small enough, taking the expectation over the Taylor expansion, of order p, will provide a reasonable approximation.

As an example, for the special but important case of $k = 2$, we derive expressions for the posterior expectation of q_1 and q_2, under second order approximation, along with error term:

$$(38) \qquad E(q_1|\boldsymbol{n}, s, \boldsymbol{t}; M, T, u, v)$$

$$(39) \qquad \approx E\left(\alpha_1 \left[1 - \alpha_2 + \alpha_2^2\right]\middle|\boldsymbol{n}, s, \boldsymbol{t}\right) E(q_t|M, T, u, v)$$

$$(40) \qquad = \left(\frac{n_1 + st_1}{N + s} - \frac{(n_1 + st_1)_1(n_2 + st_2)_1}{(N+s)_2} + \frac{(n_1 + st_1)_1(n_2 + st_2)_2}{(N+s)_3}\right)\frac{M + uv}{u + T}$$

$$(41) \qquad = \frac{n_1 + st_1}{N + s}\left(1 - \frac{n_2 + st_2}{N + s + 1}\left(1 - \frac{n_2 + st_2 + 1}{N + s + 2}\right)\right)\frac{M + uv}{u + T}$$

up to an absolute expected error less than

$$(42) \quad E\left(\alpha_1 \alpha_2^3\middle|\boldsymbol{n}, s, \boldsymbol{t}\right) E(q_t|M, T, u, v)$$

$$= \frac{(n_1 + st_1)(n_2 + st_2)(n_2 + st_2 + 1)(n_2 + st_2 + 2)}{(N+s)(N+s+1)(N+s+2)(N+s+3)}\frac{M + uv}{u + T}$$

and similarly,

$$(43) \qquad E(q_2|\boldsymbol{n}, s, \boldsymbol{t}; M, T, u, v)$$

$$(44) \qquad \approx E\left(2\alpha_2 \left[1 - \alpha_2 + \alpha_2^2\right]\middle|\boldsymbol{n}, s, \boldsymbol{t}\right) E(q_t|M, T, u, v)$$

$$(45) \qquad = 2\left(\frac{n_2 + st_2}{N + s} - \frac{(n_2 + st_2)_2}{(N+s)_2} + \frac{(n_2 + st_2)_3}{(N+s)_3}\right)\frac{M + uv}{u + T}$$

$$(46) \qquad = 2\frac{n_2 + st_2}{N + s}\left(1 - \frac{n_2 + st_2 + 1}{N + s + 1}\left(1 - \frac{n_2 + st_2 + 2}{N + s + 2}\right)\right)\frac{M + uv}{u + T}$$

up to an absolute expected error less than

$$(47) \quad E\left(2\alpha_2 \alpha_2^3\middle|\boldsymbol{n}, s, \boldsymbol{t}\right) E(q_t|M, T, u, v)$$

$$= 2\frac{(n_2 + st_2)(n_2 + st_2 + 1)(n_2 + st_2 + 2)(n_2 + st_2 + 3)}{(N+s)(N+s+1)(N+s+2)(N+s+3)}\frac{M + uv}{u + T}$$

6.2. **Sensitivity Analysis.** As mentioned in Sections 4 and 5, due to epistemic uncertainty, generally, an analyst specifies bounds for the hyperparameters s, $\boldsymbol{t}$, u, and v. The parameter sets are, as before, denoted by $\mathcal{H}$ and $\mathcal{J}$. As we assumed q_t and $\boldsymbol{\alpha}$ to be independent (see Section 6.1), we can seperate the analysis into two simpler problems. We first calculate lower and upper bounds on the expectation of the terms depending on α, based on the results from Sections 4 and 6.1. Independently, we calculate lower and upper bounds on the expectation of q_t as we did in Section 5. These bounds uniquely determine $\underline{E}(q_j|\boldsymbol{n}, M, T, \mathcal{H}, \mathcal{J})$ and $\overline{E}(q_j|\boldsymbol{n}, M, T, \mathcal{H}, \mathcal{J})$, as follows.

For convience of notation, define

$$(48) \qquad\qquad g_j(\boldsymbol{\alpha}) := \frac{1}{\binom{k-1}{j-1}}\frac{j\alpha_j}{\sum_{\ell=1}^{k}\ell\alpha_\ell}.$$

Clearly, by Eq. (4),

$$(49) \qquad q_j = g_j(\boldsymbol{\alpha})q_t$$

so,

$$(50) \qquad \underline{E}(q_j|\boldsymbol{n}, M, T, \mathcal{H}, \mathcal{J}) = \underline{E}(g_j(\boldsymbol{\alpha})|\boldsymbol{n}, \mathcal{H})\underline{E}(q_t|M, T, \mathcal{J})$$

where

$$(51) \qquad \underline{E}(g_j(\boldsymbol{\alpha})|\boldsymbol{n}, \mathcal{H}) = \min_{(s,\boldsymbol{t})\in\mathcal{H}} E(g_j(\boldsymbol{\alpha})|\boldsymbol{n}, s, \boldsymbol{t})$$

$$(52) \qquad \underline{E}(q_t|M, T, \mathcal{J}) = \min_{(u,v)\in\mathcal{J}} E(q_t|M, T, u, v)$$

Similar expressions for the upper expectation hold as well by simply replacing min by max at all instances. When using Taylor approximation, an upper bound on the error term follows readily as well (see example in Section 7).

As seen in Section 5, the optimisation problem in Eq. (52) (and its counterpart of the upper bound) can be done exactly, using Eqs. (25) and (26). In contrast, the optimisation problem in Eq. (51) (and its counterpart of the upper bound) is not so obvious, and we have to rely on standard numerical algorithms for non-linear optimisation. However, the particular form of $\mathcal{H}$ we assumed in Section 4.2.2 (see Eq. (18)) makes this optimisation problem fairly easily solvable by computer.

7. Example

To conclude the paper, we demonstrate our methodology on a simple electrical network reliability problem. Numbers are fictional, yet are representative of a typical network in the North-East of England.

A group of customers is supplied with power from two identical distribution lines. Supply is lost when both lines fail. Nationwide statistics show typical failure rate of similar distribution lines to be within $\pm 50\%$ of 0.35 per year. Nationwide statistics also show the typical fraction of double failures to be between 10% and 20%. On the actual system under study, over the last 12 years, 11 failures were observed, 3 of which were double failures.

A typical quantity one would be interested in is $q_2 = g_2(\boldsymbol{\alpha})q_t$, the rate of double failures, as this is also the rate at which customers lose power.

For the lower and upper expectation of q_t, we take $u = 3$ (in years)—this means that we need about 3 years of data before we start moving away from our prior. For v, we take $[0.175, 0.525]$, that is all values within $\pm 50\%$ of the nationwide average 0.35. We find:[5]

$$(53) \qquad \underline{E}(q_t|M, T, \mathcal{J}) = \frac{14 + 3 \times 0.175}{24 + 3} = 0.538$$

$$(54) \qquad \overline{E}(q_t|M, T, \mathcal{J}) = \frac{14 + 3 \times 0.525}{24 + 3} = 0.577$$

For the lower and upper expectation of $g_2(\boldsymbol{\alpha})$, we take $[1, 4]$ for s,[6] $[0.8, 0.9]$ for t_1, and $[0.1, 0.2]$ for t_2.

Our choice of $\overline{s} = 4$ means that after observing four single failures (and no double failures), we are prepared to reduce the prior upper fraction of double failures (0.2) by half. Our choice of $\underline{s} = 1$ means that after observing one double failure (and no single failures), we are prepared to reduce the prior lower fraction of single failures (0.8) by half.

---

[5]We have two lines, each observed 12 years, with 8 single failures on either line, and 3 failures occurring in both lines; whence, marginally, we observed $8 + 3 \times 2 = 14$ failures of a distribution line over a total timespan of 24 years.

[6]In this simple example, we have no zero counts, so we can do with a lower upper bound for s.

Then, using the 4th order Taylor approximation of $g_2(\boldsymbol{\alpha})$ as explained in Section 6.1,

$$\underline{E}(g_2(\boldsymbol{\alpha})|\boldsymbol{n}, \mathcal{H}) = \min_{(s,\boldsymbol{t})\in\mathcal{H}} E(g_2(\boldsymbol{\alpha})|\boldsymbol{n}, s, \boldsymbol{t}) \tag{55}$$

$$\approx \min_{(s,\boldsymbol{t})\in\mathcal{H}} 2\frac{n_2 + st_2}{N + s}\left(1 - \frac{n_2 + st_2 + 1}{N + s + 1}\left(1 - \frac{n_2 + st_2 + 2}{N + s + 2}\right.\right.$$

$$\left.\left.1 - \frac{n_2 + st_2 + 3}{N + s + 3}\left(1 - \frac{n_2 + st_2 + 4}{N + s + 4}\right)\right)\right) \tag{56}$$

$$= 0.360 \tag{57}$$

where $N = 11$, and $n_2 = 3$. A similar expression holds for $\overline{E}(g_2(\boldsymbol{\alpha})|\boldsymbol{n}, \mathcal{H})$; simply replace min by max:

$$\overline{E}(g_2(\boldsymbol{\alpha})|\boldsymbol{n}, \mathcal{H}) = 0.410 \tag{58}$$

Both expressions for lower and upper expecation are accurate up to the following absolute error:

$$\max_{(s,\boldsymbol{t})\in\mathcal{H}} 2\prod_{p=0}^{5}\frac{n_2 + st_2 + p}{N + s + p} = 0.006 \tag{59}$$

Concluding,

$$0.190 = (0.360 - 0.006) \times 0.538 \tag{60}$$

$$\leq \underline{E}(q_2|\boldsymbol{n}, M, T, \mathcal{H}, \mathcal{J}) \leq \overline{E}(q_2|\boldsymbol{n}, M, T, \mathcal{H}, \mathcal{J})$$

$$\leq (0.410 + 0.006) \times 0.577 = 0.240$$

or in other words, double failures occur at an expected rate that lies between 0.19 and 0.24 per year.

A similar analysis for q_1 yields:

$$0.318 = (0.595 - 0.003) \times 0.538 \tag{61}$$

$$\leq \underline{E}(q_1|\boldsymbol{n}, M, T, \mathcal{H}, \mathcal{J}) \leq \overline{E}(q_1|\boldsymbol{n}, M, T, \mathcal{H}, \mathcal{J})$$

$$\leq (0.643 + 0.003) \times 0.577 = 0.373$$

or in other words, single failures occur at an expected rate that lies between 0.318 and 0.373 per year.

In this simple example with two redundant components, posterior imprecision for the single failure rate is similar to the posterior imprecision for the double failure rate. This is essentially a special feature of the two component case, because it must hold that $\alpha_1 + \alpha_2 = 1$ when $k = 2$. In case of larger k, the differences in posterior imprecision between common-cause failure rates will be considerably larger, as in the numerical examples of Section 4.2, where, for instance, in case of Eq. (24), $\overline{E}(\alpha_j|\boldsymbol{n}, \mathcal{H}) - \underline{E}(\alpha_j|\boldsymbol{n}, \mathcal{H})$ ranges from 0.001 to 0.011.

8. Conclusion

We studied elicitation of hyperparameters for inferences that arise in the alpha-factor representation of the basic parameter model. For the hyperparameters of the Dirichlet prior for the alpha-factors, we argued that bounds, rather than precise values, are desirable, due to inferences being strongly sensitive to the choice of prior distribution, particularly when faced with zero counts. We concluded that assigning an interval for the learning parameter is especially important. In doing so, we effectively adapted the imprecise Dirichlet model [14] to represent epistemic uncertainty in the alpha-factors.

For the marginal failure rate, the second part of the model, we proposed a set of Gamma priors with similar properties as the set of Dirichlet priors used for the alpha-factors. As zero counts are generally not an issue for this part of the model, it may suffice to consider a fixed learning parameter here.

We identified simple ways to elicit information about the hyperparameters, by reasoning on hypothetical data, rather than by maximum entropy arguments as done in an earlier study [2, 7] on the estimation of alpha-factors. Essentially, the analyst needs to specify how quickly he is willing to learn from various sorts of hypothetical data.

Taking everything together, we arrived at a powerful procedure for analysing the influence of epistemic uncertainty on all common-cause failure rates, the central quantities of interest in the basic parameter model. As there is no immediate closed-form solution for the expectation of these failure rates, we presented an approximation based on Taylor expansion, and quantified the error of the approximation at any order.

By allowing the analyst to specify bounds for all hyperparameters, along with clear interpretations of these bounds, we effectively provided an operational method for full sensitivity analysis of common-cause failure rates, properly reflecting epistemic uncertainty of the analyst on all levels of the model. The procedure was illustrated by means of a simple electrical network example, demonstrating its feasability and usefulness.

In the paper, we chose the sets of hyperparameters to be of a very specific convex form (Eqs. (18) and (29)). This led to simple calculations (at least for this problem), and made elicitation fairly straightforward. Nevertheless, other shapes could still provide a better fit to any given epistemic information, and perhaps also have better updating properties. Such shapes may, however, be more difficult to elicit. More general shapes for sets of Beta priors are discussed in Walter et al. [16]. Already for Beta priors, elicitation of these shapes is non-trivial, and provides an interesting challenge. We leave a thorough study of such issues, for Dirichlet and Gamma priors, to future work.

Another aspect we neglected in this paper is the calculation of (imprecise) credible intervals—credible intervals are the Bayesian equivalent of confidence intervals, and allow us to judge the uncertainty due to sample variability. Marginal credible intervals on each component of $\boldsymbol{\alpha}$ separately are reasonably easy to obtain through the quantile function of the beta distribution [14, p. 18]. Concerning the actual quantity of interest, namely the vector $\boldsymbol{q}$, we note that the distribution of $\boldsymbol{q}$, $\boldsymbol{q}$ being a rational function of $\boldsymbol{\alpha}$ and q_t, is not easily tractable. A first approach could be to approximate the posterior distribution of $\boldsymbol{q}$ for fixed $(s, \boldsymbol{t}, u, v) \in \mathcal{H} \times \mathcal{J}$ by Monte Carlo, that is, by random sampling from $f(\boldsymbol{\alpha}|\boldsymbol{n}, s, \boldsymbol{t})$ and $f(q_t|M, T, u, v)$. We can then include epistemic uncertainty by considering a suitably constructed grid of $(s, \boldsymbol{t}, u, v)$ values in $\mathcal{H} \times \mathcal{J}$. Of course, this will be computationally quite expensive, so it would be interesting to investigate other approximation procedures.

References

[1] M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions*, 10th ed., National Bureau of Standards, Washington D.C., 1972.

[2] Corwin L. Atwood, *Constrained noninformative priors in risk assessment*, Reliability Engineering and System Safety **53** (1996), 37–46.

[3] James O. Berger, *The robust Bayesian viewpoint*, Robustness of Bayesian Analyses (J. B. Kadane, ed.), Elsevier Science, Amsterdam, 1984, pp. 63–144.

[4] James O. Berger, *Robust Bayesian analysis: sensitivity to the prior*, Journal of Statistical Planning and Inference **25** (1990), 303–328.

[5] Jose M. Bernado and Adrian F. M. Smith, *Bayesian theory*, John Wiley and Sons, 1994.

[6] I. J. Good, *The estimation of probabilities*, MIT Press, Cambridge (MA), 1965.

[7] Dana L. Kelly and Corwin L. Atwood, *Finding a minimally informative Dirichlet prior distribution using least squares*, Reliability Engineering and System Safety **96** (2011), no. 3, 398–402, `doi:10.1016/j.ress.2010.11.008`.

[8] A. Mosleh, K. N. Fleming, G. W. Parry, H. M. Paula, D. H. Worledge, and D. M. Rasmuson, *Procedures for treating common cause failures in safety and reliability studies: Procedural framework and examples*, Tech. Report NUREG/CR-4780, PLG Inc., Newport Beach, CA (USA), January 1988.

[9] Erik Quaeghebeur and Gert de Cooman, *Imprecise probability models for inference in exponential families*, ISIPTA'05: Proceedings of the Fourth International Symposium on Imprecise Probabilities and Their Applications (Pittsburgh, USA) (Fabio G. Cozman, Robert Nau, and Teddy Seidenfeld, eds.), July 2005, pp. 287–296.

[10] Matthias C. M. Troffaes and Simon Blake, *A robust data driven approach to quantifying common-cause failure in power networks*, Submitted for ISIPTA'13.

[11] Matthias C. M. Troffaes, Dana L. Kelly, and Gero Walter, *Imprecise Dirichlet model for common-cause failure*, Proceedings of PSAM 11 & ESREL 2012, June 2012.

[12] U.S. Nuclear Regulatory Commission, *Reactor safety study: an assessment of accident risk in U.S. commercial nuclear power plants*, 1975, NUREG-75/014 (WASH-1400).

[13] Peter Walley, *Statistical reasoning with imprecise probabilities*, Chapman and Hall, London, 1991.

[14] _____, *Inferences from multinomial data: Learning about a bag of marbles*, Journal of the Royal Statistical Society, Series B **58** (1996), no. 1, 3–34.

[15] Gero Walter and Thomas Augustin, *Imprecision and prior-data conflict in generalized Bayesian inference*, Journal of Statistical Theory and Practice **3** (2009), 255–271.

[16] Gero Walter, Thomas Augustin, and Frank P. A. Coolen, *On prior-data conflict in predictive Bernoulli inferences*, ISIPTA'11: Proceedings of the Seventh International Symposium on Imprecise Probability: Theories and Applications (Innsbruck) (F. Coolen, G. de Cooman, T. Fetz, and M. Oberguggenberger, eds.), SIPTA, 2011, pp. 391–400.

Durham University, UK
E-mail address: `matthias.troffaes@gmail.com`

LMU Munich, Germany
E-mail address: `Gero.Walter@stat.uni-muenchen.de`