Modelling Epistemic Uncertainty in Common-Cause Failure Models with Sets of Conjugate Priors

Matthias C. M. Troffaes[1] Gero Walter[2] Dana Kelly[3]

[1]Durham University, UK

[2]LMU, Munich, Germany

[3]Idaho National Laboratory, US

19 March 2013

Outline

Common-Cause Failures

Source: Wikimedia Commons, `http://commons.wikimedia.org/wiki/File:Fukushima_I_by_Digital_Globe.jpg`

Common-Cause Failures

Source: http://www.diakont.com/solutions/nuclear-energy/plant-systems/diesel-generator-control-systems/

Source: Wikimedia Commons, http://commons.wikimedia.org/wiki/File:Fukushima_I_by_Digital_Globe.jpg

Common-Cause Failures

- All 12 generators (for 6 reactors) at Fukushima Daiichi were not available due to flooding of machine rooms (Tsunami caused by Tōhoku earthquake)

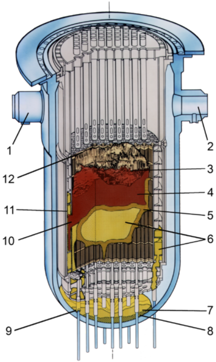
common-cause failure

simultaneous failure of several redundant components due to a common or shared root cause [3]

- Reliability of redundant systems
- Usually 2 – 4 emergency diesel generators per reactor
- Sufficient cooling of core if one generator works
- Redundant components may not fail independently: common-cause failure

Must include common-cause failures in overall system reliability analysis

Common-Cause Failure Modelling



Above: CDC, `http://phil.cdc.gov/phil/` ID 1194

Right: Wikimedia Commons,
`http://commons.wikimedia.org/wiki/File:Graphic_TMI-2_Core_End-State_Configuration.png`

Basic Parameter Model: Definition

Definition (Basic Parameter Model [5])

- immediate repair
- failures follow Poisson process
- system with k exchangeable components
- q_j: rate for failures involving *exact* j components ($j = 1, \dots, k$)
- $(q_1, \dots, q_k) =: \boldsymbol{q}$

$q_j \neq 0$ for $j \geq 2$: lack of independence for individual component failures

$\boldsymbol{q}$ is difficult to estimate directly:

- failure data often collected per component
- sparse data on joint failures

Alpha-Factor Model

Definition (Total Failure Rate)

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

total or *marginal failure rate:* failure rate obtained by looking just at single components

Definition (Alpha-Factors)

$$\alpha_j = \frac{\binom{k}{j} q_j}{\sum_{\ell=1}^{k} \binom{k}{\ell} q_\ell}. \qquad (2)$$

probability of j of the k components failing due to a common cause given that failure occurs

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

$$(\boldsymbol{q}) \Longleftrightarrow (q_t, \alpha_1, \dots, \alpha_k)$$

Data

observed per-component failure rates to estimate q_t

Data

common-cause failure counts to estimate $(\alpha_1, \dots, \alpha_k)$

Total Failure Rate: Data Model & Parameter Estimation

Poisson Process for Observed Per-Component Failures

$$\Pr(M|q_t, T) = \frac{(q_t T)^M e^{-q_t T}}{M!} \tag{4}$$

where

- total failure rate q_t
- number of per-component (i.e. marginal) failures M := total number of component failures occured (two-component failure = two failures, ...)
- time under risk T := sum of time elapsed for each of the components

The Good News

can estimate q_t directly from data, e.g. MLE:

$$\hat{q}_t = \frac{M}{T} \tag{5}$$

Alpha-Factors: Data Model

Multinomial Distribution for Common-Cause Failure Counts

$$\Pr(\boldsymbol{n}|\boldsymbol{\alpha}) = \prod_{j=1}^{k} \alpha_j^{n_j} \tag{6}$$

where

- alpha-factor α_j := probability of j of the k components failing due to a common cause given that failure occurs
- failure count n_j := corresponding number of failures observed
- $\boldsymbol{n}$ denotes $(n_1, \ldots, n_k)$ and $\boldsymbol{\alpha}$ denotes $(\alpha_1, \ldots, \alpha_k)$

Alpha-Factors: Parameter Estimation

The Good News

can estimate α directly from data, e.g. MLE:

$$\hat{\alpha}_j = \frac{n_j}{\sum_{j=1}^{n} n_j} \tag{7}$$

The Bad News

- typically, for $j \geq 2$, the n_j are very low
  with zero being quite common for larger j
- zero counts = flat likelihoods
  standard techniques such as MLE can struggle
  to produce sensible inferences for α

$\Longrightarrow$ **need to rely on epistemic information**

Dirichlet Prior

$\boldsymbol{\alpha}$ considered as uncertain parameter on which we put...

Definition (Dirichlet Distribution)

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

where $(s, \boldsymbol{t})$ are hyperparameters

$$s > 0 \qquad \boldsymbol{t} \in \Delta = \left\{ (t_1,\ldots,t_k) \colon t_1 \geq 0,\ldots,t_k \geq 0, \sum_{j=1}^{k} t_j = 1 \right\} \tag{9}$$

Interpretation

- $\boldsymbol{t}$ = prior expectation of $\boldsymbol{\alpha}$
- s = determines learning speed (see next slide)

Dirichlet Posterior

- posterior density for $\boldsymbol{\alpha}$ is again Dirichlet

$$f(\boldsymbol{\alpha}|\boldsymbol{n}, s, \boldsymbol{t}) \propto \prod_{j=1}^{k} \alpha_j^{st_j+n_j-1}. \tag{10}$$

- posterior expectation of α_j

$$E(\alpha_j|\boldsymbol{n}, s, \boldsymbol{t}) = \int_{\Delta} \alpha_j f(\boldsymbol{\alpha}|\boldsymbol{n}, s, \boldsymbol{t}) \, \mathrm{d}\boldsymbol{\alpha} = \frac{N}{N+s}\frac{n_j}{N} + \frac{s}{N+s} t_j \tag{11}$$

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

we shall focus on $E(\alpha_j|\boldsymbol{n}, s, \boldsymbol{t})$
(in a decision context, this expectation would typically end up in expressions for expected utility)

Example

(taken from [4])

Example

Consider a system with four redundant components ($k = 4$).
The analyst specifies the following prior expectation $\mu_{\text{spec},j}$ for each α_j:

$$\mu_{\text{spec},1} = 0.950 \quad \mu_{\text{spec},2} = 0.030 \quad \mu_{\text{spec},3} = 0.015 \quad \mu_{\text{spec},4} = 0.005 \tag{12}$$

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

$$n_1 = 35 \qquad n_2 = 1 \qquad n_3 = 0 \qquad n_4 = 0 \tag{13}$$

Non-Informative Priors

large variation in posterior under different non-informative priors

- with constrained maximum entropy prior (Kelly and Atwood [1, 4]):

$$E(\alpha_1|\boldsymbol{n}, s, \boldsymbol{t}) = 0.967 \qquad E(\alpha_2|\boldsymbol{n}, s, \boldsymbol{t}) = 0.028$$
$$E(\alpha_3|\boldsymbol{n}, s, \boldsymbol{t}) = 0.003 \qquad E(\alpha_4|\boldsymbol{n}, s, \boldsymbol{t}) = 0.001$$

- with uniform prior $t_j = 0.25$ and $s = 4$:

$$E(\alpha_1|\boldsymbol{n}, s, \boldsymbol{t}) = 0.9 \qquad E(\alpha_2|\boldsymbol{n}, s, \boldsymbol{t}) = 0.05$$
$$E(\alpha_3|\boldsymbol{n}, s, \boldsymbol{t}) = 0.025 \qquad E(\alpha_4|\boldsymbol{n}, s, \boldsymbol{t}) = 0.025$$

- with Jeffrey's prior $t_j = 0.25$ and $s = 2$:

$$E(\alpha_1|\boldsymbol{n}, s, \boldsymbol{t}) = 0.9342 \qquad E(\alpha_2|\boldsymbol{n}, s, \boldsymbol{t}) = 0.0395$$
$$E(\alpha_3|\boldsymbol{n}, s, \boldsymbol{t}) = 0.0132 \qquad E(\alpha_4|\boldsymbol{n}, s, \boldsymbol{t}) = 0.0132$$

Imprecise Dirichlet Model: Definition

- use a set of hyperparameters [7, 8]

$$\mathcal{H} = \Big\{(s, \boldsymbol{t}) \colon s \in [\underline{s}, \overline{s}],\ \boldsymbol{t} \in \Delta,\ t_j \in [\underline{t}_j, \overline{t}_j]\Big\} \tag{14}$$

  over which we do a sensitivity analysis (á la robust Bayes)

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

Imprecise Dirichlet Model: Elicitation

- $[\underline{t}_j, \overline{t}_j]$? cautious interpretation of prior specifications $\mu_{\text{spec},j}$:

$$[\underline{t}_1, \overline{t}_1] = [0.950, 1] \qquad [\underline{t}_2, \overline{t}_2] = [0, 0.030]$$
$$[\underline{t}_3, \overline{t}_3] = [0, 0.015] \qquad [\underline{t}_4, \overline{t}_4] = [0, 0.005]$$

- $[\underline{s}, \overline{s}]$? Good [2]:

  reason about posterior expectations of hypothetical data

$\overline{s}$ = number of one-component failures required
to reduce the upper probabilities of multi-components failure by half

$\underline{s}$ = number of multi-component failures required
to reduce the lower probability of one-component failure by half

Imprecise Dirichlet Model: Elicitation

reasonable values:

- $\underline{s} = 1$:
  immediate multi-component failure
  $\Longrightarrow$ keen to reduce lower probability for one-component failure
- $\overline{s} = 10$:
  after observing 10 one-component failures
  $\Longrightarrow$ halve upper probabilities of multi-component failures

there is a difference between $\overline{s}$ and $\underline{s}$
as the rate at which we reduce upper probabilities
is less than the rate at which we reduce lower probabilities
$\Longrightarrow$ reflects a level of caution

Imprecise Dirichlet Model: Inference

prior bounds + likelihood → posterior bounds

- with $t_j = \mu_{\text{spec},j}$:

j	$\underline{E}(\alpha_j \mid \mathbf{n}, \mathcal{H})$	$\overline{E}(\alpha_j \mid \mathbf{n}, \mathcal{H})$
1	0.967	0.972
2	0.0278	0.0283
3	0.00041	0.00326
4	0.00014	0.00109

- with bounds as earlier:

j	$\underline{E}(\alpha_j \mid \mathbf{n}, \mathcal{H})$	$\overline{E}(\alpha_j \mid \mathbf{n}, \mathcal{H})$
1	0.967	0.978
2	0.0270	0.0283
3	0	0.00326
4	0	0.00109

Gamma Prior and Posterior

q_t considered as uncertain parameter on which we put…

Definition (Gamma Distribution)

$$f(q_t|u,v) \propto q_t^{uv-1} e^{-q_t u} \tag{15}$$

where (u, v) are hyperparameters with $u > 0$ and $v > 0$.

Interpretation

- v = prior expectation of q_t
- u = determines learning speed (just like s in the IDM)

- posterior density for q_t is again Gamma

$$f(q_t|M,T,u,v) \propto q_t^{uv+M-1} e^{-q_t(u+T)}. \tag{16}$$

- posterior expectation of q_t

$$E(q_t|M,T,u,v) = \frac{T}{T+u}\frac{M}{T} + \frac{u}{T+u}v \tag{17}$$

Imprecise Gamma Model

use a set of hyperparameters:

$$\mathcal{J} = \Big\{(u, v) \colon u \in [\underline{u}, \overline{u}],\ v \in [\underline{v}, \overline{v}]\Big\} \tag{18}$$

- $[\underline{v}, \overline{v}]$? Bounds for prior expectation of q_t should be easy to find (choosing $\underline{v} = 0$ is possible)
- $[\underline{u}, \overline{u}]$? Similar reasoning as for the IDM leads to...

$\overline{u}$ = timespan for observing the process required to raise the lower expectation of q_t from 0 to half of observed failure rate $\frac{M}{T}$ ($\underline{v} = 0$ is assumed)

$\underline{u}$ = timespan for observing the process *without any failures* required to reduce the lower expectation of q_t by half ($\underline{v} > 0$ is assumed)

$\underline{u} = \overline{u}$ can be reasonable here, as zero counts are less of an issue

Inference on Common-Cause Failure Rates q_j

combine our models for $\boldsymbol{\alpha}$ and q_t by using Eq. (3):

$$q_j = g_j(\boldsymbol{\alpha}) q_t \qquad \text{where} \qquad g_j(\boldsymbol{\alpha}) = \frac{1}{\binom{k-1}{j-1}} \frac{j\alpha_j}{\sum_{\ell=1}^{k} \ell\alpha_\ell}$$

The Bad News

no closed expression for $E(g_j(\boldsymbol{\alpha}) | \dots)$ due to rational function of $\boldsymbol{\alpha}$

The Good News

naive approximation $\tilde{g}_j(\boldsymbol{\alpha})$ of $g_j(\boldsymbol{\alpha})$ by Taylor expansion works surprisingly well (absolute error term available)

$$E(q_j | \boldsymbol{n}, s, \boldsymbol{t}; M, T, u, v) \approx E\left(\tilde{g}_j(\boldsymbol{\alpha}) | \boldsymbol{n}, s, \boldsymbol{t}\right) E(q_t | M, T, u, v) \tag{19}$$

(q_t and $\boldsymbol{\alpha}$ are assumed to be independent)

Global Sensitivity Analysis

We can do a **global sensitivity analysis** for $E(q_j|\dots)$

$\Longrightarrow$ bounds for $E(q_j|\dots)$ taking into account approximation error and *epistemic uncertainty expressed through* $\mathcal{H}$ *and* $\mathcal{J}$:

$$\underline{E}(q_j|\boldsymbol{n}, M, T, \mathcal{H}, \mathcal{J}) \approx \underline{E}(\tilde{g}_j(\boldsymbol{\alpha})|\boldsymbol{n}, \mathcal{H})\underline{E}(q_t|M, T, \mathcal{J}) \qquad (20)$$

where

$$\underline{E}(\tilde{g}_j(\boldsymbol{\alpha})|\boldsymbol{n}, \mathcal{H}) = \min_{(s,\boldsymbol{t})\in\mathcal{H}} E(\tilde{g}_j(\boldsymbol{\alpha})|\boldsymbol{n}, s, \boldsymbol{t}) \quad \text{(by num. optimization)} \qquad (21)$$

$$\underline{E}(q_t|M, T, \mathcal{J}) = \min_{(u,v)\in\mathcal{J}} E(q_t|M, T, u, v) \quad \text{(by closed form solution)} \qquad (22)$$

Do the same for $\overline{E}(q_j|\boldsymbol{n}, M, T, \mathcal{H}, \mathcal{J})$ by replacing min with max

Conclusion

main messages:

- bounds, rather than precise values, are desirable
  due to inferences being strongly sensitive to the prior
  particularly when faced with zero counts.
- simple ways to elicit the parameters of the model
  by reasoning on hypothetical data
  rather than by maximum entropy arguments
- sets of hyperparameters allow a full sensitivity analysis
  reflecting epistemic uncertainty of the analyst
  on all levels of the model

stingy questions:

- hyperparameter sets have very specific form,
  do they fit to the epistemic information at hand?
  (other set shapes are currently investigated)
- can use of variance obliterate use of bounds on expectations?
  (operational interpretation of variance of an unknown parameter
  versus direct bounds on expectation of unknown parameter)
  (credible intervals do not save the example discussed)

References I

[1] C. L. Atwood.
Constrained noninformative priors in risk assessment.
Reliability Engineering and System Safety, 53:37–46, 1996.

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

[3] Arnljot Høyland and Marvin Rausand.
System reliability theory: models and statistical methods.
A Wiley interscience publication. Wiley, New York, NY, 1994.

[4] Dana Kelly and Corwin Atwood.
Finding a minimally informative Dirichlet prior distribution using least squares.
Reliability Engineering and System Safety, 96(3):398–402, 2011.

[5] 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.
Technical Report NUREG/CR-4780, PLG Inc., Newport Beach, CA (USA), January 1988.

[6] Matthias C. M. Troffaes, Gero Walter, and Dana Kelly.
A robust Bayesian approach to modelling epistemic uncertainty in common-cause failure models.
Submitted.

[7] Peter Walley.
Statistical Reasoning with Imprecise Probabilities.
Chapman and Hall, London, 1991.

[8] Peter Walley.
Inferences from multinomial data: Learning about a bag of marbles.
Journal of the Royal Statistical Society, Series B, 58(1):3–34, 1996.