Series on Quality, Reliability and Engineering Statistics Vol. 7

Mathematical and Statistical Methods in Reliability

editors

Bo H Lindqvist

Norwegian University of Science and Technology, Trondheim, Norway

Kjell A Doksum

University of Wisconsin, Madison, USA

World Scientific 2003
New Jersey • London • Singapore • Hong Kong

22

MODELLING HETEROGENEITY IN NUCLEAR POWER PLANT VALVE FAILURE DATA

Madhuchhanda Bhattacharjee and Elja Arjas
Rolf Nevanlinna Institute
University of Helsinki
Finland
E-mail: mab@rni.helsinki.fi, ela@rni.helsinki.fi

Urho Pulkkinen
VTT Industrial Systems
VTT, Finland
E-mail: urho.pulkkinen@vtt.fi

We develop an inferential procedure based on hierarchical Bayesian models which can take into account different kinds of heterogeneity in reliability/survival data. This is illustrated by an analysis of failure data from several motor operated closing valves at two nuclear power plants.

1. Introduction

Heterogeneity in data has been frequently encountered and incorporated in modelling and inference problems in areas like design of experiments, survival analysis, duration analysis, and many more. As far as reliability data are concerned this issue has been addressed only to a very limited extent from both these aspects. For example, burn-in problems have been considered for engineering systems, but possibly do not explain heterogeneous behaviour when longitudinal data are collected from such systems. Some relevant works in this context are by Pörn[8], Bedford et al.[2], Block and Savits [3], Lin and Singpurwalla[7], Hofer and Peshchke[5], Hougaard[6], etc.

This chapter develops inference procedures which can take into account different kinds of heterogeneity in reliability/survival data and enable inferences to be drawn not only on population parameters relating to the survival process, but also on underlying processes that lead to such observed heterogeneity between individuals.

We illustrate the proposed model by analysing a real life data set, wherein, although the observations consist of repeated failure times from similar systems, several factors can cause these systems to behave differently. In the current set-up, a large number of such variables can be suspected to be causing heterogeneity, but how exactly these affect the failure rate may be very difficult to establish.

2. Data

Our data pertain to several motor operated closing valves in different safety systems at two boiling water reactor plants in Finland, viz. TVO-I and TVO-II nuclear power units. Previous studies carried out on plants and their safety systems have shown significant variations in the number of valve failures and also indicated deviations in the reliability of some of the valves and their operating conditions (Pulkkinen and Simola[9]). The valves being parts of safety systems, such findings cast doubt about their performance when actually on demand in a real situation.

The data are based on 9' years operating experiences since the beginning of 1981, of 104 such motor operated valves, with 52 valves in each of the plants. These are parts of four different safety systems. Distribution of the valves according to Safety system and Plant is given in Table 1.

A motor operated valve, including the actuator, is an electromechanical system and consists of several parts, and hence can be quite complicated in structure. The valves have different design capacities, vary in diameter and actuator size and are manufactured by different manufacturers. The valves can experience different types of malfunctioning either when in operation or when pre-scheduled tests are carried out on them as a part of preventive maintenance (Simola and Laakso[10]).

Here we consider the failures of the type "External Leakage". An external leakage is recorded from a valve usually when there is a leakage from one of its subcomponents, such as a "Bonnet" or "Packing". Valves are continuously monitored for such failures and are rectified/repaired without delay. Here we assume that repair times are negligible.

Table 1. Distribution of valves according to safety system, plant and failure.

Safety System (≡ Valve Type)	Plant I		Plant II		All	
	Total	Failed	Total	Failed	Total	Failed
1. Shut-down cooling system	4	3	4	2	8	5
2. Containment vessel spray system	24	6	24	4	48	10
3. Core spray system	16	0	16	0	32	0
4. Auxiliary feed water system	8	1	8	0	16	1
All	52	10	52	6	104	16

As is expected, a large majority of systems did not fail during the 9 year follow up, as these are parts of safety systems and are built to perform successfully for long time periods. From the 104 valves 37 such external leakages were observed with the failures pertaining to 16 valves only.

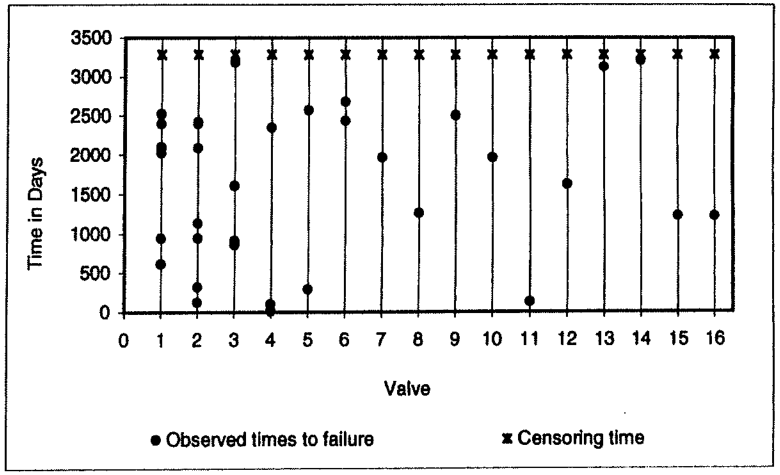

Fig. 1. Plots of cumulative times to failure of 20 valves.

The cumulative times to failures and the inter-failure times from these 16 valves are presented in Figure 1 and Table 2, respectively, with the valves being arranged according to the number of failures observed from them. The remaining 88 valves reporting no failures have not been included in this table, although each of them produces one censored time to first failure, censored at 3286-th day.

Table 2. Inter-event times of external leakages from the valves.

Valve Number	Inter-failure times in days								
	1-st	2-nd	3-rd	4-th	5-th	6-th	7-th	8-th	9-th
1	610	4	329	1081	63	17	295	126	761*
2	126	197	620	189	955	312	27	860*	
3	860	55	691	1575	105*				
4	10	9	85	2248	934*				
5	293	2274	719*						
6	2434	242	610*						
7	1963	1323*							
8	1262	2024*							
9	2501	785*							
10	1963	1323*							
11	132	3154*							
12	1623	1663*							
13	3127	159*							
14	3211	75*							
15	1225	2061*							
16	1222	2064*							

3. Preliminary Models:

In a preliminary analysis we decided to ignore the observed heterogeneity across valves and started with a simplistic model, assuming that the failures are governed by Poisson processes with common hazard/intensity parameter values for all 104 valves. We assumed a hierarchical model structure for the parameters. But this model predicted a median inter-failure time of 17 years for all valves. This, for an obvious reason, is not acceptable for valves which had experienced approximately 1 failure every year, e.g. for Valve 1 which had 8 failures in 9 years, or Valve 2 which had 7 failures in 9 years.

As a next step we relaxed the above model structure by assuming that the failure process of any valve is governed by a Poisson process but the processes are not necessarily identical in distribution for all valves. We assumed an appropriate hierarchical structure for these processes. This simple extension immediately captured the heterogeneous behavior of the valves and valve specific predicted inter-failure times now had median values ranging approximately from 2 years to 26 years. This is consistent with the observed behavior of the respective valves, as some of them had failed often whereas most others did not fail during the observed period.

We further tried an extension of this model by dropping the Poisson assumption. Instead, we considered a point process model where failure intensity for each valve is piecewise constant and changes randomly at its own failure epochs, the intensities being drawn independently from a Gamma distribution. This was done observing that also the inter-event times from any specific valve varied widely, and under the homogeneous Poisson process assumption they are assumed to be coming from the same exponential distribution. But even though this was a simple and intuitive extension it lost the valve specific prediction ability as due to the implicit identical assumption the model did not contain any valve specific features. The large number of long (censored or uncensored) inter-failure times dominated the prediction and the predicted median inter-arrival time was estimated to be approximately 14 years for all valves. Note that, by integrating out the piecewise constant parameters, this model reduces to modelling the valves by identical renewal processes with Pareto distributed inter-failure times.

These three preliminary models (M1, M2 and M3, say) can be summarized as follows:

A. Inter-failure times: For $i = 1, \ldots, 104$ and $j \geq 1$,

- M1. i.i.d. Homogeneous Poisson Process: $X_{i,j} \mid \theta \sim Exp(\theta)$
- M2. Non-i.i.d. homogeneous Poisson Process: $X_{i,j} \mid \theta_i \sim Exp(\theta_i)$
- M3. Point process with piecewise constant intensity: $X_{i,j} \mid \theta_{ij} \sim Exp(\theta_{ij})$

B. Hazard rate(s): For $i = 1, \ldots, 104$ and $j \geq 1$

- M1. i.i.d. Homogeneous Poisson Process: $\theta \mid \beta \sim Gamma(1, \beta)$,
- M2. Non-i.i.d. homogeneous Poisson Process: $\theta_i \mid \beta \sim Gamma(1, \beta)$,
- M3. Point process with piecewise constant intensity: $\theta_{ij} \mid \beta \sim Gamma(1, \beta)$,

C. Hyper parameter: (Under all three models M1-M3)

$$\beta \sim Gamma(5, 0.003)$$

Table 3. Preliminary analysis of valve failure data

Valve no.	Observed no. of failures	Predicted median failure time (apprx.) in years		
		i.i.d. HPP	Non-i.i.d. HPP	PP with piecewise const. intensity
1	8	17	2	14
2	7	17	2.3	14
10	1	17	10.5	14
17	0	17	25.5	14

Reasonably vague distributions were chosen for the hyper-parameters. A partial summary of the preliminary analysis is given in Table 3. From the preliminary analyzes we could see that, even in a situation in which there is considerable heterogeneity in the inter-failure times both between valves and within each valve, merely adding more latent variables may not be a good strategy for modelling. Here, in the third model, it resulted in a loss of the predictive performance.

4. Model and Analysis

It is apparent from the data that the valves exhibited varying behaviour during the period of study with respect to external leakages, with the majority experiencing no failures whereas some recording as many as 8. Some of the possible sources of heterogeneity have been described above.

The standard procedure is to express such heterogeneity in a mathematical form, by incorporating the variables (possibly) causing such deviations as "covariates". A popular method of incorporating the effects of a covariate on the probabilistic behaviour of a system is to use them while modelling the failure intensity/hazard rate of the system, possibly in a form known as the Cox regression model (Cox[4]). In the current set-up, we have a large number of such variables which can be suspected to be causing the valves to behave in different ways. Also, how exactly these affect the failure rate may be difficult to establish through any explicit mathematical formulation.

The inter-failure times of the valves are observed to be highly varying, with very small as well as very large values being observed from the same valve. No specific pattern like systematic deterioration, was observed. Several preliminary models for the valves, like i.i.d. and non-i.i.d. renewal processes, non-homogeneous Poisson processes with piecewise constant intensities, with the intensities changing/jumping at failure epochs, were attempted for initial analyses. Evident heterogeneity rejected models treating the valves in an i.i.d. manner. Also models without valve specific control on the models for the inter-failure times performed poorly from a prediction point of view. From these we noted that even if heterogeneity is allowed between valves, within valve heterogeneity, arising possibly due to repairs and other environmental factors, should be modelled carefully. An appropriate model should be able capture valve specific behaviour also.

Following Arjas and Bhattacharjee[1], we model the probabilistic behaviour of these systems in a hierarchical Bayesian model. The first layer of this model consists of the inter-failure times. As we have already pointed out above, due to evident heterogeneity in the valves the inter-failure times need to be modelled separately for each valve. Also for any fixed valve the inter-failure times exhibit a high amount of variation without any clear pattern over time in them. Hence for each valve we consider a renewal process with a Pareto inter-failure time distribution. This should be able to explain absence of time trend and to accommodate both very small and very large inter-failure times.

Next, we observe that there are at least two different patterns in the behaviour of the valves, with most not failing during the entire followup and a few failing even several times. This effectively means that each valve comes either from an, *a priori* unknown, group of "good" valves with possibly long failure times, or from a (smaller) group of "bad" valves with not so long inter-failure times. The two groups/classes are described by assigning to each valve a Bernoulli distributed latent variable identifying the group for that valve. Assuming that the inter-failure times are Pareto distributed, the two groups are characterised by drawing the shape parameters from two ordered Gamma distributions.

Our lack of further prior knowledge about these latent variables is expressed by assigning a Uniform distribution to the parameter of the Bernoulli distribution.

This model structure can be described briefly as follows:

A. Inter-failure times :	$X_{i,j} \mid \beta_0, \beta_1, C_i \sim Pareto(\beta_{C_i})$, where $i = 1, \ldots, 104$ and $j \geq 1$,
B. Latent variables :	$C_i \mid p \sim Bernoulli(p), i = 1, \ldots, 104$
C. Group-specific parameters :	$\beta_1 \sim Gamma(.,.)$ and $\beta_0 = \beta_1 + \eta_1$, where $\eta_1 \sim Gamma(.,.)$,
D. Mixing probability :	$p \sim Uniform(0,1)$.

Several values of the hyper-parameters were chosen in order to carry out sensitivity analysis. Parameters (5, 0.001) and (5, 0.005) were found to be appropriate choices for the distributions of η_1 and β_1 respectively, giving rise to reasonably uninformative priors for β's . The proportions of the "good" and "bad" valves in the population may be estimated from the posterior distribution of the mixing hyper-parameter, namely, p. The posterior distribution of C_i describes, for the i-th valve, the chances of coming from the "good" or the "bad" group, which may be highly useful for many purposes, like reliability assessment and maintenance.

The structure of the model makes it possible to evaluate posterior characteristics on population level (Table 4) and also at generic valve level (Table 5). As is expected the proportion of "bad" valves is only 7 %, consistent with the fact that these are parts of safety systems. The benefit of considering the quality of a valve is evident when the predicted lifetime of a generic valve is considered. With no added information on quality the predicted life time is more than 30 years, whereas with the information that the valve is of poor quality, the prediction is drastically reduced to about 1 year only.

Table 4. Results of analysis of valve failure data: Population level characteristics

Characteristic	Posterior expectation
Proportion of "bad" valves	0.07
Proportion of "good" valves	0.93
β Parameter for "bad" valves	494
β Parameter for "good" valves	13650

Table 5. Results of analysis of valve failure data: Generic-valve characteristics.

Characteristic	Posterior estimate
Predictive probability for a generic valve to be a "good" valve	0.93
Predictive median time to first failure for a generic valve	31.67 years
Predictive median time to first failure for a "bad" valve	1.27 years
Predictive median time to first failure for a "good" valve	37 years

For planning maintenance activities it would certainly be useful to have an estimate of the expected number of failures in future. Given the data for 9 consecutive years we predicted the expected number of failures from all valves during the 10-th year and during 10-th & 11-th year (see Table 6). Unsurprisingly, the expected number of failures is approximately proportional to the duration of prediction. A closer inspection reveals that it

is highly unlikely that more than many valves would experience more than one failure during the considered prediction periods.

Table 6. Results of analysis of valve failure data: Maintenance related characteristics.

Characteristic	Predictive expectation
No. of failures from all 104 valves during the 10-th year	4.00
No. of failures from all 104 valves during the 10 & 11-th year	7.93
No. of valves to fail at least once during 10-th year	3.47
No. of valves to fail ≥ 2 times during 10-th year	0.41
No. of valves to fail ≥ 3 times during 10-th year	0.09
⋮	⋮
No. of valves to fail at least once during 10 & 11-th year	6.32
No. of valves to fail ≥ 2 times during 10 & 11-th year	1.06
No. of valves to fail ≥ 3 times during 10 & 11-th year	0.37
⋮	⋮

If maintenance programs are designed for specific valves or valve types, predictions for individual valves are needed. Examples of such predictions are given in Tables 7 and 8. Note that for any valve chances of failing beyond the study period would depend not only on its quality but also on when it had failed last during the study period.

The concept of "good" and "bad" valves was extended within each category of valves (i.e. valves belonging to a specific safety system). It was observed that indeed the estimates of the population parameters for different safety systems differ significantly. The posterior estimates of some of the model parameters are presented in Table 8.

We observe from Table 8 that the parameter estimates for Valve-type /Safety-system 1 differ noticeably from the others. This is in agreement with the information provided by Table 1, showing that a relatively higher proportion of valves from the first safety system failed during the study. This difference may also be due to the fact that unlike the other three groups, valves in Safety system 1 are continuously in operation, and this probably makes them more failure prone than the others.

Also the number of predicted failures for the 10-th and 11-th years are seen to be varying across valve type, which is consistent with the observed

Table 7. Results of analysis of valve failure data: Unit level characteristics.

Characteristic	Posterior/predictive estimates for valve no.								
	1	2	3	4	5	6	7	8	9
Chances of being a "good" valve	0.00	0.00	0.02	0.01	0.72	0.73	0.97	0.96	0.97
Median future inter failure time	461	471	493	465	6760	7000	12770	12510	12910
Expected no. of failures during the 10-th year	0.31	0.29	0.54	0.28	0.10	0.11	0.03	0.03	0.03
Expected no. of failures during the 10 & 11-th year	0.64	0.60	0.97	0.56	0.21	0.22	0.06	0.06	0.07
Median (residual) time to next failure at the end of 9-th year	1248	1341	603	1468	7815	7907	14140	14450	13860

Characteristic	Posterior/predictive estimates for valve no.								
	10	11	12	13	14	15	16	17	
Chances of being a "good" valve	0.96	0.83	0.96	0.96	0.95	0.96	0.96	0.99	...
Median future inter failure time	12270	9119	12380	12330	12500	12480	12160	13080	...
Expected no. of failures during the 10-th year	0.03	0.04	0.03	0.04	0.05	0.03	0.03	0.02	...
Expected no. of failures during the 10 & 11-th year	0.06	0.08	0.06	0.09	0.09	0.06	0.06	0.04	...
Median (residual) time to next failure at the end of 9-th year	14170	13440	14100	12500	12550	14390	14270	16740	...

Table 8. Results of analysis of valve failure data: Valve type specific analysis

Characteristic	Posterior expectations for valve type			
	1	2	3	4
Proportion of "bad" valves	0.67	0.05	0.04	0.09
Proportion of "good" valves	0.33	0.95	0.96	0.91
β Parameter for "bad" valves	467	1273	1229	1193
β Parameter for "good" valves	5778	9559	12310	9230
Expected no. of failures during the 10-th year	1.93	1.56	0.81	0.54
Expected no. of failures during the 10 & 11-th year	3.80	3.10	1.60	1.08

data. Note that although the predicted number of failures for valve types 1 and 2 appear to be comparable, the number of valves in these two safety systems are quite different, being 8 and 48 respectively. Similar comments can be made about the other two types of valves (see Tables 1 and 8). Also observe that if predictions of future failures for valves in different safety systems were made using estimated population parameters from the previous model, the predictions would have been proportional to the number of valves in the specific safety system. The added information of valve type, therefore, refines the inference in this aspect.

If it is believed that the valves could be usefully divided into more than two categories, then this structure can be easily extended by introducing additional (η, β)'s and a Multinomial-Dirichlet combination for C_i's and p.

In a general situation, ordering the β parameters may be found as too restrictive, although it helps avoiding identifiability problems. Also it may not be easy to identify such orders in a multivariate parameter setup. In that case, instead, the "profiles" of the individuals may be used to fix/characterise the groups. For example, for the valve data, two valves may be identified to represent the two groups, namely "good" and "bad". This can easily be achieved by using predetermined values for C_i's of these two valves. The remaining valves may be modelled using the above mixture model. But note that now the estimation will be influenced by the profiles of the chosen "good" and "bad" valves which serve as fixed representatives of their categories.

5. Conclusions

The heterogeneity of failure behaviour of safety related components, such as valves in our case study, may have important implications for reliability analysis of safety systems. If such heterogeneity is not identified and taken into account, the decisions made to maintain or to enhance safety can be non-optimal or even erroneous. This non-optimality is more serious if the safety related decisions are made on the basis of failure histories of the components.

This work demonstrates how even rather simplistic models could describe the heterogeneous behaviour successfully. We were able to make an assessment of the quality (i.e. whether the valve is "good" or "bad") of individual valves, and we were also able to estimate the proportions of "good" and "bad" valves in the population. Furthermore, we could analyse how additional information of the valve quality helps in predicting the future failure behaviour of a considered valve. These predictions can be made both for individual valves and for the valve populations as a whole. Predictions of this kind can be used in developing maintenance programs and in predicting the cost of future maintenance. Another use of the results, in addition to qualitative failure analyses, can be the identification of weaknesses in earlier maintenance programs.

References

1. E. Arjas and M. Bhattacharjee, Modelling heterogeneity/selection and causal influence inreliability, *IISA-JSM 2000-2001 Conference, New Delhi, India*, (2001).
2. T. Bedford, R. M. Cooke, J. Dorrepaal, Mathematical review of swedish Bayesian methodology for nuclear plant reliability data bases, *The Practice of Bayesian Analysis. Eds, French, S, Smith J.Q.*, (1997).
3. H. W. Block and T. H. Savits, Burn-in, *Statistical Science*, **12**, 1-13 (1997).
4. D. R. Cox, Regression models and life tables, *Journal of the Royal Statistical Society, Series B*, **34**, 187-220 (1972).
5. E. Hofer, J. Peschke, Bayesian modelling of failure rates and initiating event frequencies, *Proceedings of the European Conference on Safety and Reliability (ESREL), Garching, Germany 13-17-September*, (1999).
6. P. Hougaard, *Analysis of Multivariate Survival Data*, (Springer) (2000).
7. N. J. Lynn and N. D. Singpurwalla, Comment: "Burn-in" makes us feel good, *Statistical Science*, **12**, 13-19 (1997).
8. K. Pörn, On empirical Bayesian inference applied to Poisson probability models, *Linköping studies in science and technology. Dissertations 234*, (1990).

9. U. Pulkkinen and K. Simola, Bayesian Models and Ageing Indicators for Analysing Random Changes in Failure Occurrence, *Reliability Engineering & System Safety*, **68**, 255-268 (2000).
10. K. Simola and K. Laakso, Analysis of Failure and Maintenance Experiences of Motor Operated Valves in a Finnish Nuclear Power Plant, *VTT Research Notes 1322* (1992).