



Guidelines for Analysis of Data Related to Ageing of Nuclear Power Plant Components and Systems

**EC JRC Network on Use of Probabilistic Safety Assessments (PSA)
for Evaluation of Aging Effects to the Safety of Energy Facilities.
Task 4.**

EUR 23954 EN - 2009

The Institute for Energy provides scientific and technical support for the conception, development, implementation and monitoring of community policies related to energy. Special emphasis is given to the security of energy supply and to sustainable and safe energy production.

European Commission
Joint Research Centre
Institute for Energy

Contact information

Address: PO Box 2, NL-1755 ZG Petten
E-mail: andrei.rodionov@jrc.nl
Tel.: 0224 56 54 57
Fax: 0224 56 56 21

<http://ie.jrc.ec.europa.eu>
<http://www.jrc.ec.europa.eu>

Legal Notice

Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use which might be made of this publication.

A great deal of additional information on the European Union is available on the Internet. It can be accessed through the Europa server
<http://europa.eu/>

EUR 23954 EN
ISSN 1018-5593

Luxembourg: Office for Official Publications of the European Communities

© European Communities, 2009

Reproduction is authorised provided the source is acknowledged

Printed in Netherlands

Guidelines for Analysis of Data Related to Ageing of Nuclear Power Plant Components and Systems

Andrei Rodionov

Dana Kelly

Jens Uwe-Klügel

February 2009

Preface

This guideline is intended to provide practical methods for practitioners to use in analyzing component and system reliability data, with a focus on detection and modeling of ageing. The emphasis is on frequentist and Bayesian approaches, implemented with MS EXCEL and the open-source software package WinBUGS. The methods described in this document can be implemented with other software packages.

Contents

1	Introduction	1
1.1	Data Analysis Objectives	1
1.2	Overview and Flow of the Guidelines	1
2	Data Collection and Processing.....	4
2.1	Goals and objectives	4
2.2	Data specification	4
2.3	Collection and processing.....	6
2.3.1	Checking available data sources.....	6
2.3.2	Data processing : output formats.....	10
3	Statistical Techniques for Analysis of Reliability Data	13
3.1	Trend assessment	13
3.1.1	General provisions.....	13
3.1.2	Visual examination	13
3.1.3	Nonparametric tests	19
3.1.4	Parametric tests.....	24
3.2	Reliability parameters specification and elaboration	32
3.2.1	Generic process for PSA time-dependent parameter estimation	32
3.2.2	Specification of unavailability type and attributes associated with basic events	34
3.2.3	Reliability models and parameters definition	38
3.2.4	Unavailability as a function of unit age.....	40
4	Review of Bayesian Inference	43
4.1	Introduction.....	43
5	Parameter Estimation and Model Validation	45
5.1	Preface to Chapter 5.....	45

- 5.2 Inference for a Single Parameter 46
 - 5.2.1 Binomial Distribution for Failures on Demand 47
 - 5.2.2 Poisson Distribution for Initiating Events or Failures in Time 51
 - 5.2.3 Exponential Distribution for Random Durations..... 54
- 5.3 Checking Model Assumptions 57
 - 5.3.1 Binomial distribution..... 57
 - 5.3.2 Poisson Distribution..... 60
- 5.4 Time-Trend Models for p and λ 63
 - 5.4.2 Time-Dependent Model for λ in the Poisson Distribution 68
- 5.5 Population Variability Models..... 73
 - 5.5.1 Population Variability Model for p in the Binomial Distribution 73
 - 5.5.2 Population Variability Model for λ in the Poisson Distribution 79
- 5.6 More Complex Models for Time to Failure – Nonrepairable Failure 84
 - 5.6.1 Alternatives to Exponential Distribution for Random Durations 84
 - 5.6.2 Choosing Among Alternative Distributions – Deviance Information Criterion (DIC) 93
- 5.7 Modeling Failure with Repair 95
 - 5.7.1 Repair Same As New – Renewal Process 95
 - 5.7.2 Repair Same As Old – Nonhomogeneous Poisson Process (NHPP)..... 97
 - 5.7.3 Impact of Assumption Regarding Repair..... 106
 - 5.7.4 Incorporating Results into PSA..... 109
- 5.8 Treatment of Censoring..... 111
 - 5.8.1 Censored Data for Random Durations 112
 - 5.8.2 Left-Censoring of Failure Times for NHPP 119
- 5.9 Case Studies 120
 - 5.9.1 Case #3 120
 - 5.9.2 Case #7.1 123
 - 5.9.3 Case DEF..... 124
 - 5.9.4 Case Study for Random Durations..... 126

6	References.....	129
	Appendix A – Probability Fundamentals	A-1
	Appendix B – WinBUGS	B-1
	Appendix C – Supporting Mathematics.....	C-1

Tables

•	Table 1 Additional data sources.....	9
•	Table 2 Output format for sample specification.....	10
•	Table 3 Output format for individual times to failure data	11
•	Table 4 Output format for binned data	12
•	Table 5 Estimation of cumulated hazard values.....	14
•	Table 6 Point estimations and side-by-side confidence intervals for binned data	15
•	Table 7 Calculation of cumulative number of failures.....	17
•	Table 8 Calculation of inter-arrival times	18
•	Table 9 Lower (A_L) and upper (A_U) inversion limits	21
•	Table 10 Binned data for large sample of repairable NPP components	22
•	Table 11 Operating states / failure modes specification.....	37
•	Table 12 Parameter estimates for logistic model for time trend in valve leakage.....	67
•	Table 13 Data for Case #3.....	121
•	Table 14 Data for Case 7.1	123
•	Table 15 Results for Case 7.1.....	124
•	Table 16 Date for Case DEF.....	124

Figures

- Figure 1 Workflow for the investigation of time dependent trends in component failure data and the implementation of the results in a plant specific PSA..... 2
- Figure 2 Potential data sources for age-dependent reliability analysis..... 7
- Figure 3 Nomenclature of PSA Component Reliability Data..... 8
- Figure 4 Availability and accessibility of data for different types of reliability models..... 10
- Figure 5 Cumulated hazard plot for data from Example 3 Times to failure data for non-repairable components 15
- Figure 6 Side-by-side confidence intervals (a) and point estimations (a, b) of failure rates per bin 16
- Figure 7 Cumulative failure plot for repairable components 18
- Figure 8 Inter-arrival times plot for individual component. 19
- Figure 9 Calculation of inversions number 23
- Figure 10 Processed data for goodness of fit test..... 26
- Figure 11 EXEL SOLVER dialog window 27
- Figure 12 Presentation of goodness of fit test results. Basis case 28
- Figure 13 Presentation of goodness of fit test results. Early failures excluded..... 29
- Figure 14 Presentation of “ageing” parameters lower confidence limit 31
- Figure 15 EXEL SOLVER dialog window 32
- Figure 16 Generic process of elaboration of age-dependent input data. 34
- Figure 17 Step-wise approximation of time-dependent failure intensity function..... 41
- Figure 18 Unit and component ages correlation. 42
- Figure 19 DAG for binomial inference with lognormal prior 50
- Figure 21 DAG for Poisson inference with lognormal prior 53
- Figure 22 DAG for exponential inference with lognormal prior..... 56
- Figure 23 Side-by-side plot of 95% credible intervals for p in each of the 9 years. Dots indicate posterior mean, and red line is the average of the posterior means. 59

- Figure 24 Side-by-side interval plot illustrating varying λ over time 62
- Figure 25 DAG model for inference with time trend in p 64
- Figure 26. Plot of values of first 1,000 values for “a” parameter, illustrating convergence..... 66
- Figure 27. Plot of values of first 1,000 values for “b” parameter, illustrating convergence..... 66
- Figure 28. BGR diagnostic for “a” parameter, indicating convergence..... 66
- Figure 29. BGR diagnostic for “b” parameter, indicating convergence..... 67
- Figure 30. Posterior distribution for “b” parameter. Values below zero are very unlikely, suggesting an increasing trend in p over time. 67
- Figure 31 Plot of 95% interval for p in each year based on a trend model in which p is increasing with time. 68
- Figure 32 DAG model for inference with time trend in λ 69
- Figure 33. Plot of values of first 1,000 values for “a” parameter, illustrating convergence..... 71
- Figure 34. Plot of values of first 1,000 values for “b” parameter, illustrating convergence..... 71
- Figure 35. BGR diagnostic for “a” parameter, indicating convergence..... 71
- Figure 36. BGR diagnostic for “b” parameter, indicating convergence..... 72
- Figure 37 Posterior distribution for “b” parameter. Values above zero are very likely, suggesting an increasing trend in λ over time. 72
- Figure 38 Plot of 95% interval for λ in each year based on a trend model in which λ is increasing with time
73
- Figure 39 DAG model for hierarchical Bayesian inference in binomial distribution 75
- Figure 40 Plot of first 1,000 values of α , indicating convergence. 77
- Figure 41 BGR diagnostic for α , indicating convergence. 77
- Figure 42 Plot of first 1,000 values of β , indicating convergence..... 78
- Figure 43 BGR diagnostic for β , indicating convergence. 78
- Figure 44 Plot of first 1,000 values of $p.avg$, indicating convergence..... 78
- Figure 45 BGR diagnostic for $p.avg$, indicating convergence..... 79
- Figure 46 DAG model for hierarchical Bayesian inference in Poisson distribution..... 80
- Figure 47 Plot of first 1,000 values of CV , indicating convergence. 82
- Figure 48 BGR plot for CV , indicating convergence..... 82
- Figure 49 Plot of first 1,000 values of mean parameter, indicating convergence. 83

- Figure 50 BGR plot for mean parameter, indicating convergence. 83
- Figure 51 Plot of first 1,000 values of λ_{avg} , indicating convergence. 83
- Figure 52. BGR plot for λ_{avg} , indicating convergence. 84
- Figure 53 DAG model for gamma-distributed random durations 86
- Figure 54 Posterior distribution for α indicates that values above 1 are likely 88
- Figure 55 DAG model for Weibull-distributed random durations 89
- Figure 56 Posterior distribution of α for Example 7. 90
- Figure 57 DAG model for lognormally-distributed random durations 91
- Figure 58 Posterior distribution of μ in Example 8. 92
- Figure 59 Posterior distribution of σ in Example 8. 93
- Figure 60 Cumulative hazard plot for Example 9, suggesting increasing failure rate with operating time 97
- Figure 61 Cumulative failure plot for Example 10, suggesting increasing ROCOF over time .99
- Figure 62 Cumulative failure plot for data in Example 9, showing lack of trend in slope over calendar time 100
- Figure 63 Cumulative failure plot for 1,000 simulated failure times from renewal process with decreasing failure rate 101
- Figure 64 Cumulative hazard plot for 1,000 simulated failure times from renewal process with decreasing failure rate 102
- Figure 65 Cumulative failure plot for 1,000 simulated failure times from renewal process with increasing failure rate 102
- Figure 66 Cumulative hazard plot for 1,000 simulated failure times from renewal process with increasing failure rate 103
- Figure 67 DAG for modeling failure with repair under same-as-old repair assumption 104
- Figure 68 Cumulative failure plot for 1,000 times simulated from power-law process with shape parameter of 2, illustrating increasing ROCOF. 107
- Figure 69 Histogram of times between failures for simulated failure times from power-law process with increasing ROCOF 108
- Figure 70 DAG model for Type I censoring in the exponential distribution 112
- Figure 71 DAG model for interval censoring in the exponential distribution 116
- Figure 72. Posterior distribution for shape parameter in Weibull distribution. 117
- Figure 73 DAG model for interval-censoring in the Weibull distribution 118

GUIDELINES FOR ANALYSIS OF DATA RELATED TO AGING OF NUCLEAR POWER PLANT COMPONENTS AND SYSTEMS

- Figure 74 Plot of 95% credible intervals shows no significant time trend in λ 121
- Figure 75 Marginal posterior density for aging coefficient in linear model 122
- Figure 76 Marginal posterior density for aging coefficient in loglinear (exponential) model... 122
- Figure 77 Marginal posterior density for shape parameter in power-law model..... 123
- Figure 78 Plot of 95% credible intervals suggests possible increasing time trend in λ 124
- Figure 79 Caterpillar plot for p suggests slight increasing trend over time (demands) 125
- Figure 80 Cumulative failure plot for example times 127
- Figure 81 Cumulative hazard plot, suggesting decreasing hazard rate 128

WinBUGS Scripts

- Script 1. WinBUGS script for Bayesian inference with binomial likelihood function and lognormal prior. 50
- Script 2 WinBUGS script for Bayesian inference with binomial likelihood function and logistic-normal prior 51
- Script 3 WinBUGS script for Bayesian inference with Poisson likelihood function and lognormal prior 53
- Script 4 WinBUGS script for Bayesian inference with exponential likelihood function and lognormal prior 56
- Script 5 WinBUGS script to generate caterpillar plots for p with multiple data sources (block data entry) 58
- Script 6. WinBUGS script for testing assumption that p in binomial distribution does not vary over time. 60
- Script 7 WinBUGS script to test for poolability of Poisson data..... 63
- Script 8. WinBUGS script for modeling a time trend in p 65
- Script 9. WinBUGS script for modeling time trend in λ 70
- Script 10. WinBUGS script for analyzing population variability in p 76
- Script 11 WinBUGS script for analyzing population variability in λ with Poisson likelihood. 82
- Script 12 WinBUGS script for Bayesian inference of random durations using gamma likelihood. 87
- Script 13 WinBUGS script for Bayesian analog of Cramer-von Mises test of exponential distribution for failure times..... 88
- Script 14 WinBUGS script for Bayesian inference of random durations using Weibull likelihood. 90
- Script 15. WinBUGS script for Bayesian inference of random durations using lognormal likelihood. 92
- Script 16 WinBUGS script for analyzing data under same-as-old repair assumption (power-law process) 105
- Script 17 WinBUGS script for predicting next failure time in NHPP and estimating failure probability 110
- Script 18 WinBUGS script for estimating failure probability for NHPP using cumulative distribution function 111
- Script 19 WinBUGS script for modeling censored random durations.(Type I censoring)..... 113
- Script 20 WinBUGS script for Example 12 – Type II censoring of failure times..... 115

GUIDELINES FOR ANALYSIS OF DATA RELATED TO AGING OF NUCLEAR POWER PLANT COMPONENTS AND SYSTEMS

- Script 21 WinBUGS script for Example 12, interval-censored failure times with exponential likelihood.
..... 116
- Script 22. WinBUGS script for Example 12, interval-censored failure times with Weibull likelihood.
..... 119
- Script 23 WinBUGS script to analyze data from NHPP (repair same as old) with left censoring.. 120

Examples

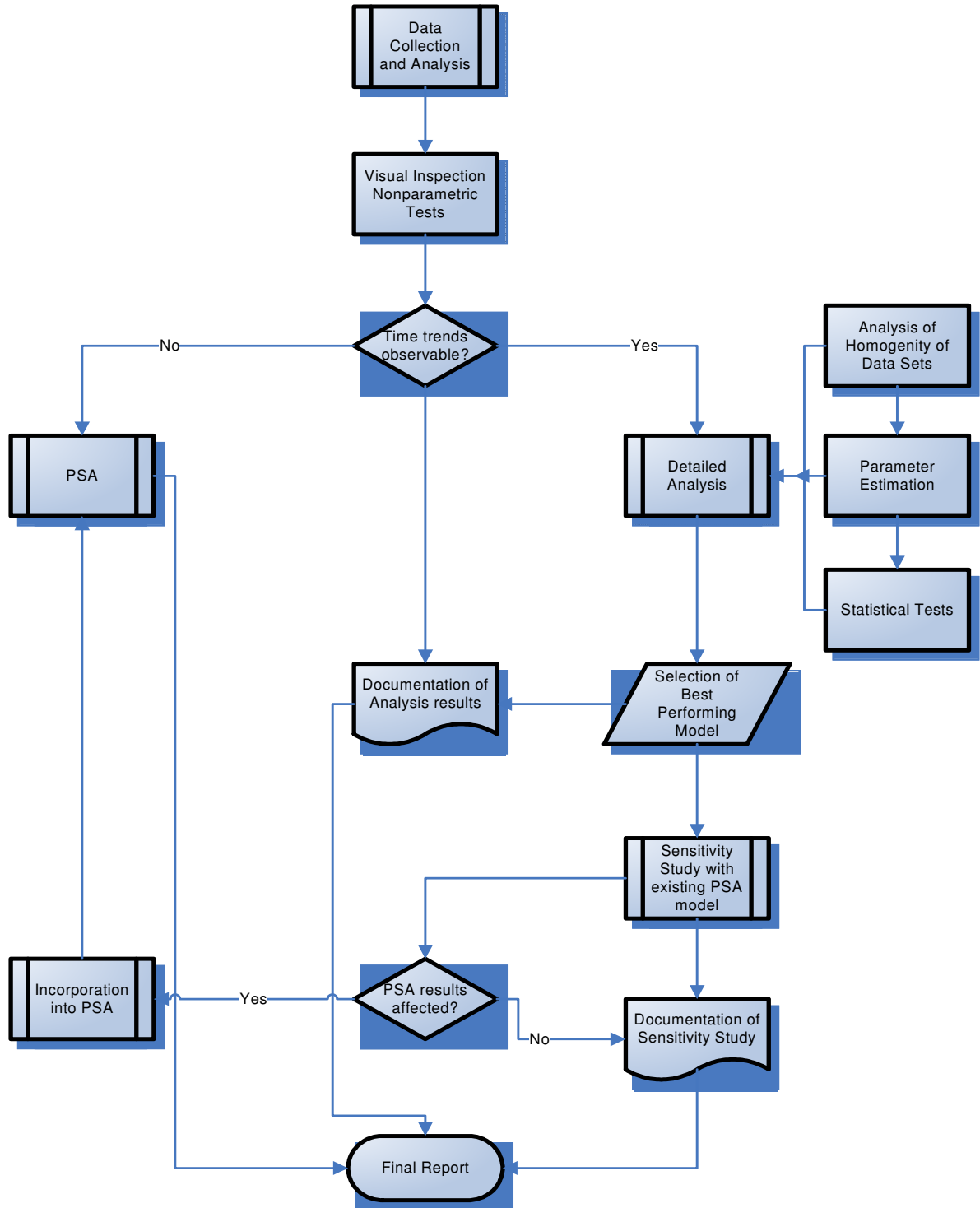
- Example 1 Calculation of individual times to failure..... 5
- Example 2 Calculation of failure rates per bin 6
- Example 3 Times to failure data for non-repairable components..... 14
- Example 4 Binned data for non-repairable components 15
- Example 5 Failure times of individual repairable component 18
- Example 6 Binned data for large sample of repairable NPP components 22
- Example 7 External leakage of HPSI valves..... 35
- Example 8 Relief valve fails to open (binomial model) and lognormal prior. 49
- Example 10 Cooling pump fails to operate (Poisson model) and lognormal prior..... 52
- Example 11 Circulating pump fails to operate (exponential model) and lognormal prior..... 55
- Example 12 Check for time trend in relief valve leakage probability in cooling water system 57
- Example 13 Check for constant component failure rate..... 61
- Example 14 Times between failure for analysis with gamma distribution 85
- Example 15 Inference for Weibull distribution as aleatory model for circulating pump failure times in ATCS system..... 89
- Example 16 Inference for lognormal distribution as aleatory model for failure times..... 91
- Example 17 Failure times from a renewal process 96
- Example 18 Failure times from a non-renewal process 98
- Example 19 Modeling failure times with Type II censoring..... 113
- Example 20 Modeling interval-censored failure times 115
- Example 21 Example 10 revisited with left censoring at 100 days 119

1.1 Data Analysis Objectives

This document, Guidelines for Analysis of Data Related to Aging of Nuclear Power Plant Components and Systems, is intended to provide guidelines for the collection and evaluation of data related to the aging of components and systems in operating nuclear power plants. It is aimed at applied scientists and engineers rather than statisticians and thus provides a hands-on approach to a variety of data assessment methods, tools, and techniques.

1.2 Overview and Flow of the Guidelines

Figure 1 shows the general work flow for the investigation of time dependent trends in failure data of active components starting from data collection and analysis up to a later incorporation of the results into a plant-specific PSA.



• Figure 1 Workflow for the investigation of time dependent trends in component failure data and the implementation of the results in a plant specific PSA

The first step of the procedure consists of systematic **data collection and analysis**. Traditional methods applied in a PSA to collect plant-specific data are in general suitable for an in-depth analysis of time-dependent trends, as long as all of the relevant raw information obtainable from operational and maintenance personnel is collected and stored, preferably in electronic format. In contrast with traditional data collection procedures for parameters without time dependence, it is important to obtain more accurate information on the

life history of the investigated components (date of commission, dates of replacement, dates of complete overhaul work) and on the time of occurrence of failure events. This issue is discussed in Chapter 2 of the Guideline.

The second step consists of qualitative analysis, including **visual inspection** and simple **nonparametric tests** to check for the presence of time trends. The results of this step have to be documented. If nonparametric tests do not show any notable time trend the investigation can be stopped. The investigation may have to be repeated after more years of operation, for example in conjunction with periodic safety reviews. Several techniques to handle these kind of analysis are presented in Chapter 3.1 of the Guideline.

If a significant time trend is suggested by qualitative analysis or nonparametric test, it more **detailed analysis** is needed to allow a more accurate quantitative description of the time trend. This detailed analysis serves to confirm the presence of a systematic trend and develops a model from which future failure events can be predicted. This is especially important with respect to ageing effects, as the capability to predict future failure events can be of some importance for other applications besides PSA, for example, for planning long-term maintenance work and investments. It may also be of some benefit with respect to financial risk management, because unpredicted failure events may cause production losses. The quantitative estimation and prediction model should be selected on the basis of statistical tests measuring the performance of different model alternatives. The questions related to model construction and validation are discussed in paragraph 3.1.4 (frequentist approach) and Chapter 5 (Bayesian analysis).

The first step with respect to an implementation of the developed models into PSA should consist in a **sensitivity analysis** aimed to study the effect of time trends of component failure rates on the overall PSA results. If the impact of time dependency on CDF or LERF is small the investigation can be stopped by documenting the results in a **final report**. Such a graded approach allows to avoid unnecessary effort in cases then the effects of time trends are small and well bounded by the subjective uncertainty distributions used in the plant specific PSA to model component failure rates. It might be the case that the results of the PSA are dominated by some specific initiating events (e.g. external events like earthquakes) or by the contribution of common cause failures and the effects of time dependency on the overall results of the PSA are small.

A good documentation of the study results should be mandatory, because the insights gained may be of benefit for other applications, especially for maintenance planning. If the impact of time dependency is not negligible (a typical criterion is a change of the mean core damage frequency by more than 20%) then it is recommended to **implement** the results of the investigation directly in the **plant specific PSA**. The implementation can be performed in different degrees. For example, the results of the sensitivity study may indicate, that time dependent effects are only relevant for a small group of components. In many cases it may be sufficient to adjust the existing data models by using a time dependent mean value for the failure rates (f. e. for living PSA applications) instead of implementing a time dependent model. Such a simplified approach may be requested due to limitations of the PSA software used. The more detailed predictive models are then used only to calculate the time dependent mean values for failure rates to be used in the PSA model. The issues related to specification and elaboration of reliability parameters for PSA model are discussed in paragraph 3.2.

2.1 Goals and objectives

This chapter specifies the data needed for time-dependent reliability analysis and describes the main steps of data collection and processing.

The objective of data collection and processing is to prepare an initial data set for one of the following tasks :

- analysis of ageing trend,
- elaboration of time-dependent reliability models.

As it was previously mentioned the first step of the analysis consists of systematic **data collection and processing**.

2.2 Data specification

A non-repairable population is one for which individual components that fail are removed permanently from the population. While the system may be repaired by replacing failed components from either a similar or a different population, the members of the original population dwindle over time until all have eventually failed.

An important parameter defined for non repairable populations is the failure rate. This is the (instantaneous) rate of failure of components during the next instant of time that survived to time t. It is a rate per unit of time. During the next instant of time the failure rate may change and the units that have already failed play no further role since only the survivors count.

A repairable component is one which can be restored to a satisfactory operational state by any action, including parts replacements or changes to adjustable settings. When discussing the rate at which components fail during system operation time (and are then repaired) we will define a Rate Of Occurrence Of Failure (ROCOF) or failure intensity.

A different approach is used for modelling the ROCOF. Time is measured by system power-on-hours from initial turn-on at time zero, to the end of system life. Failures occur at given system ages and the system is repaired to a state that may be the same as new, or better, or worse. The frequency of repairs may be

increasing, decreasing, or staying at a roughly constant rate.

Let $N(t)$ be a counting function that keeps track of the cumulative number of failures a given system has had from time zero to time t . $N(t)$ is a step function that increases by one every time a failure occurs and stays at the new level until the next failure.

Every system will have its own observed $N(t)$ function over time. If we observed the $N(t)$ curves for a large number of similar systems and "averaged" these curves, we would have an estimate of $M(t)$ = the expected number (average number) of cumulative failures by time t for these systems.

The derivative of $M(t)$, denoted $m(t)$, is defined to be the Rate Of Occurrence Of Failures at time t or ROCOF.

Censoring (when the observation period ends, not all units have failed - some are survivors)

The initial reliability data can be presented in two ways:

- individual times to failure data: times to failure and censoring times - $\{t_i, t_{fi}, t_{ri}\}, i=1, \dots, n$,
- binned data: failure intensities (rates) or failure probabilities on demand and their confidence intervals per bin - $\{\lambda_k^{low}, \lambda_k, \lambda_k^{up}\}, k=1, \dots, M$.

Here

t_i, t_{fi}, t_{ri} – are the time to left censoring, time to failure and time to right censoring correspondently,

n – is a number of observations (components) in the sample,

$\lambda_k, \lambda_k^L, \lambda_k^U$ – are the failure rate and its lower and upper confidence limits estimated at time interval k ,

M – is a number of intervals (bins).

In case of times to failure data “times” could represent any suitable measure of “age” of the component, i.e. hours in operation, calendar time, number of demands/start-ups, number of cycles, etc. The same interpretation of “age” is applicable for binned data. Here, failure intensity could be expressed as failure per hour, per demand, per number of cycles, etc.

When processing both kinds of data an important assumption has to be made concerning the nature of the repair process after the failure : same as new, same as old or partial renewal.

Binned data could be easily processed from individual times to failure data, but not vice versa. Certain information from initial data set is lost in this case. From these perspectives, the binned data are less informative then individual times to failure

In practical applications, it is important to correlate (synchronize) the “component age” with “age of the unit”.

- Example 1 Calculation of individual times to failure.

If the age of the component is measured in years of operation (calendar time) it is sufficient to have as initial raw data the dates of certain events. Knowing the dates of failure, the beginning and end of the observation period, and the date which corresponds to the beginning of operation (“age” = 0), the times to failure and censoring times could be calculated as follows:

$$t_i = D_S - D_{COM} - \text{left censoring (start of observation period),}$$

$$t_{ri} = D_E - D_{COM} - \text{right censoring (end of observation period),}$$

$$t_{fi} = D_F - D_{COM} - \text{time to critical failure.}$$

Here

D_S is the date at which observation begins,

D_E is the date at which observation ends,

D_F is a date of critical failure,

D_{COM} is the date of commissioning / beginning of operation.

• Example 2 Calculation of failure rates per bin

Assume that all components are commissioned at the same time and the repair of failed components is same-as-old. Data are collected as a number of failures per year for each component group. Dividing the observation period by M equal intervals (bins) of width Δt_k , the failure rates could be estimated as follows:

$$\lambda_k = x_k / (c \Delta t_k) \quad (1)$$

Here

x_k – total number of failures in interval k ,

c – number of components in the group.

Approximate confidence limits will be :

$$\lambda_k^L = \frac{\chi_{\alpha/2}^2(2x)}{2(c \Delta t_k)}$$

$$\lambda_k^U = \frac{\chi_{1-\alpha/2}^2[2(x+1)]}{2(c \Delta t_k)} \quad (2)$$

Where, χ^2 is the percentile of the chi-square distribution with corresponding degrees of freedom.

Separating out Failure Modes.

Note that when a data set consists of failure times that can be sorted into several different failure modes, it is possible (and often necessary) to analyze and model each mode separately. Consider all failures due to modes other than the one being analyzed as censoring times, with the censored run-time equal to the time it failed due to the different (independent) failure mode.

2.3 Collection and processing

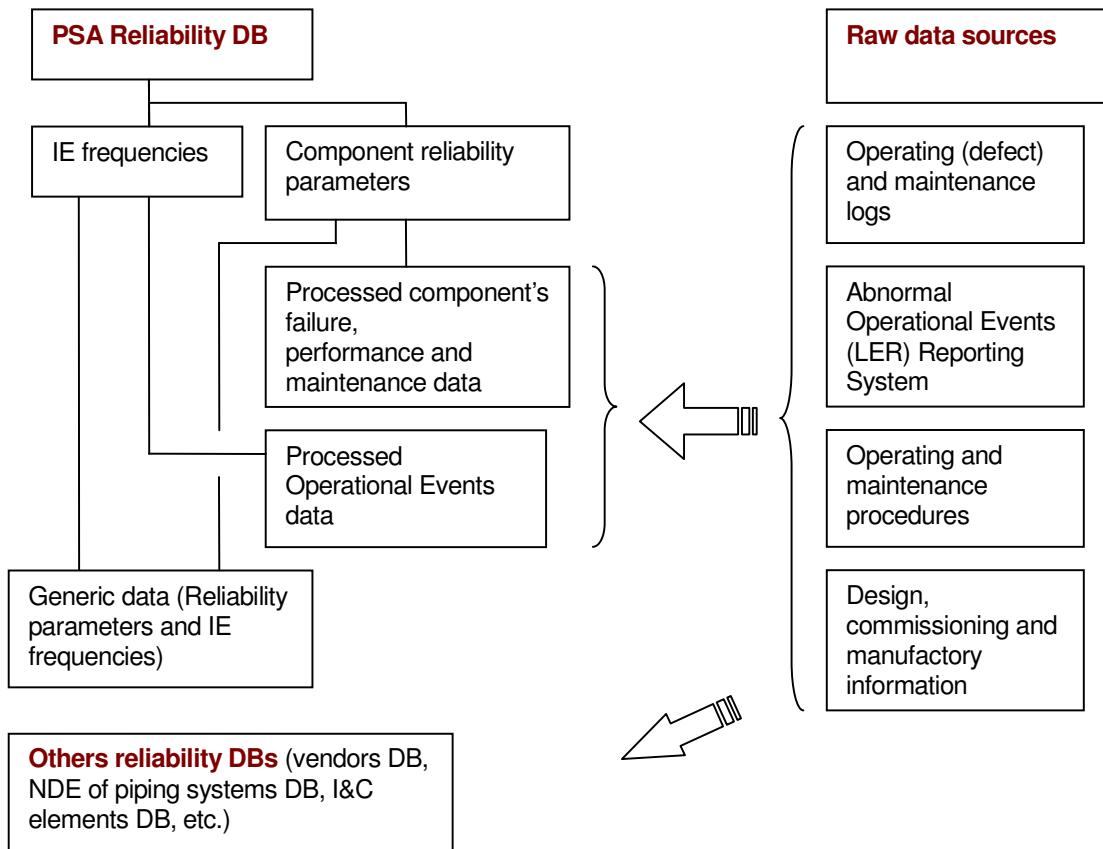
2.3.1 Checking available data sources

The first step of data collection consists of identification of available data sources.

Figure 2 is a greatly simplified representation of relations between three types of data considered potentially available for age-dependent reliability analysis:

- PSA reliability data,
- Other specific reliability data,

- Raw operating and maintenance data. Information about design, manufacturing, operating and maintenance.



• Figure 2 Potential data sources for age-dependent reliability analysis

PSA reliability data includes estimates of initiating event frequencies and component reliability parameters. These data are directly used in PSA Event Tree and Fault Tree models.

Most important information collected and analyzed during PSA reliability data collection and processing is usually documented in the relevant PSA task reports and/or databases. These data are well structured and of high quality.

An example of PSA Component Reliability Data nomenclature is presented in Figure 3.

It has to be underlined that only “processed data” about failures and component performance could be used for age-dependent reliability analysis, but even those data are not sufficient for this purpose. Consequently, additional data must be extracted and processed from raw data sources.

Another potential data source presented in Figure 2 is non-PSA reliability data. These data cover component types not included in the PSA model. Such data could be of interest if such a component has to be added to the PSA or if it will be necessary to enlarge the database for a similar PSA component group.

The third type of data that could be used for age-dependent reliability analysis, are the raw data collected from the plant, as well as operational, maintenance and design documentation.

Processed data	
Reliability parameters 1. Component group 2. Failure modes 3. Parameters : <ul style="list-style-type: none"> • Failure rate, • Failure prob. per demand, • Mean time to repair, • Unavailability due to the maintenance. 4. Uncertainty (confidence intervals or Error Factors)	Data for parameters estimation 1. Component group 2. Component description / limits 3. Operating mode 4. Component function 5. Failure modes 6. Failure criteria 7. Sample specification (# and list of components) 8. Observation period 9. Stressors (av. # of hours in operation or demands per year) 10. Cumulated stressors 11. # of failures per failure mode 12. Estimation method / assumptions
	Data for cumulated operation time 13. Component ID (sub-group) 14. # of compon. in sub-group 15. # of hrs or demands / year 16. Observation period 17. Total cumulated stressors Failure data 18. Unit / component ID 19. Failure date 20. Reactor state 21. Failure mode 22. Criticality factor 23. Repair time 24. Failure cause / description

• Figure 3 Nomenclature of PSA Component Reliability Data

The following procedure could be applied to check available data sources:

Step 1 : checking existing reliability data sources

1. Start from PSA Reliability data and then consider non-PSA Reliability data.
2. Check if “**processed data**” (see Figure 3) are available.
3. Check if all **failure types and modes** considered for ageing analysis are covered by collection process. Verify **compliance of failure criteria**.
4. Check if data **collection period is large enough** to perform an ageing analysis (observation period more than 8-10 years).
5. Check for which component groups **sufficient failure counts** are available (more than 10 failures are registered).
6. Verify if **reporting and failure criteria** remain the same during the observation period.
7. Verify if any important **modifications** or massive **replacements** of components have occurred during the observation period.

Step 2 : identify additional data categories needed for time-dependent ageing assessment :

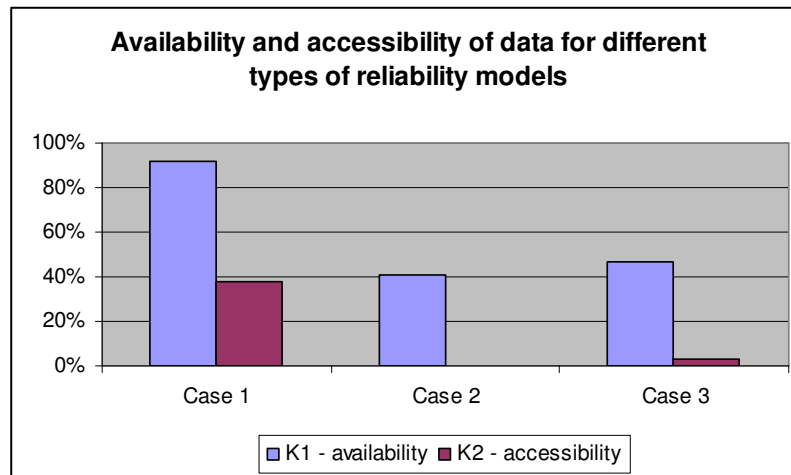
1. Define the type of **potential application** :
 - Type 1: simple age-dependent reliability models or trend assessment;
 - Type 2: age-dependent reliability models including test and maintenance effect evaluations;
 - Type 3: comprehensive age-dependent reliability models, for use in advanced PSA, maintenance optimisation and lifetime analysis.
2. Specify **categories of additional data** needed for chosen application

3. Identify potential **sources of data** and check their **availability and accessibility**. See example of potential data sources in Table 1.

• Table 1 Additional data sources

Application type	Data category	Possible data sources
Type 1: simple age-dependent reliability models or trend assessment	Component commissioning date (age 0)	In-plant components DB, maintenance DB; Unit commissioning date from IAEA PRIS
	Failure/censoring dates	LERs, operating and maintenance logs Component Reliability DB
	Component replacement – date and cause	LERs, operating and maintenance logs Component Reliability DB
Type 2: age-dependent reliability models including test and maintenance effect evaluations (categories of additional data needed are the same listed for Type 1, plus the followings)	Characteristics of applied tests and maintenance strategy – type and periodicity	Test and maintenance procedures
	Degree of component renewal during the maintenance	Maintenance procedures, Maintenance logs
Type 3: comprehensive age-dependent reliability models, for use in advanced PSA, maintenance optimisation and lifetime analysis (categories of additional data needed are the same listed for Type 1 and Type 2, plus the followings).	Component lifetime	Design specifications, Vendor documentation
	real cumulated number of hours in operation, number of demands	Operating logs, centralized signal registration system; System operating procedures and test programs
	information about average and extreme levels of operating stressors	Design and commissioning docs, operating procedures, safety report; Operating logs, centralized signal registration system; LERs, IRS reports about transients
	information about of average and extreme levels of environmental stressors	Design and commissioning docs, operating procedures, safety report; Operating logs, centralized signal registration system; LERs, IRS reports about transients

Unfortunately, raw data are not always available or easily extractable. An expert evaluation of availability and accessibility of data for different types of reliability models is shown in Figure 4 taken from (Rodionov A., 2007).



• Figure 4 Availability and accessibility of data for different types of reliability models

The diagram reflects the expert judgements about availability and accessibility of raw data. The experts belong to 8 different organisations that responded to the questionnaire. This diagram demonstrates that even for simple age-dependent reliability assessments for which most of the data are available, the cost of additional data processing could be quite high. If one has to apply reliability models to maintenance analysis and optimisation, or to lifetime evaluation and prediction, additional data collection and processing efforts would have to be made.

An improvement of reliability data collection could greatly help with age-dependent reliability analysis applications.

The difficulties with data availability and accessibility in some cases could be resolved by postulation of model assumptions. For example, assumption that the age of the components is equal to the unit age, permits the use of unit commissioning date (age 0), as individual component commissioning date.

2.3.2 Data processing : output formats

To assure the quality and transparency of data processing the following output formats are recommended :

Sample specification, see Table 2,

Data tables :

Table 3 for processing individual times to failure data,

Table 4 for binned data.

• Table 2 Output format for sample specification

Component group : (title/ID)	
Unit(s)	(list/IDs)

GUIDELINES FOR ANALYSIS OF DATA RELATED TO AGING OF NUCLEAR POWER PLANT COMPONENTS AND SYSTEMS

Failure Mode(s)	(list/coding)
Operating modes / failure mode	(list)
Sample size	
Sample	(list of components/IDs)
Sub-groups	(titles/IDs)
Statistic	
Data source(s)	(reference list)
Observation period	(dates)
Covered ages window	(years)
Number of registered failures	(critical/non-critical/CCF)
Number of recurrent failures	
Censoring rate	
Uncertain records	
Assumption on the type of repair after failure	(same as new/same as old/partial)
Assumption on the component state after preventive maintenance	(same as new/same as old/partial)
Reliability parameters (calculated under assumption of constant failure rate/probability on demand)	(for comparison)

• Table 3 Output format for individual times to failure data

Unit ID	Component ID	Date of commissioning / replacement	Date of the beginning of observation	Date of failure	Date of the end of observation	Failure mode	Criticality factor	t_{li}	t_{fi}	t_{ri}

GUIDELINES FOR ANALYSIS OF DATA RELATED TO AGING OF NUCLEAR POWER PLANT COMPONENTS AND SYSTEMS

• Table 4 Output format for binned data

Bins	1	2	3	M
Years in operation (age)						
Number of failures						
Cumulative operating time / number of demands						
λ_k						
λ_k^L						
λ_k^U						

Chapter
3

Statistical Techniques for Analysis of Reliability Data

3.1 Trend assessment

3.1.1 General provisions

The simplest task of statistical analysis is to investigate whether the SSC failure rate is approximately constant.

Various statistical tests could be used to validate or to refute the assumption of constant failure rate. Some of them are discussed in NUREG/CR-6823 (Atwood C.L. and all, 2003). According to the statistical technique used they can be divided into three groups:

- graphs (visual evaluation),
- nonparametric hypothesis tests,
- parametric hypothesis tests.

The purpose of this chapter is to present available statistical techniques that could be used for ageing analysis of reliability data.

3.1.2 Visual examination

These simple plots are a powerful diagnostic tool since the human eye can often detect patterns or anomalies in the data by studying graphs. That kind of invaluable information would be lost if the analyst only used quantitative statistical tests to check model fit. Every analysis should include as many visual tests as are applicable.

Advantages of Visual Tests

- easy to understand and explain,
- can occasionally reveal patterns or anomalies in the data,
- when a model "passes" a visual test, it is somewhat unlikely any quantitative statistical test will "reject" it (the human eye is less forgiving and more likely to detect spurious trends).

Disadvantages of Visual Tests

- visual tests are subjective,
- they do not quantify how well or how poorly a model fits the data,
- they are of little help in choosing between two or more competing models that both appear to fit the data.

3.1.2.1 Non-repairable components or repair same-as-new (renewal process)

Construction of **Cumulative Hazard plot** for times to failure data (use times between failures for repairable components)

1. Order the failure times and running times for each of the n units on test in ascending order from 1 to n. The order is called the rank of the unit. Calculate the reverse rank for each unit (reverse rank = n-rank +1).
2. Calculate a Hazard "value" for every failed unit (do this only for the failed units). The Hazard value for the failed unit with reverse rank k is just 1/k.
3. Calculate the cumulative hazard values for each failed unit. The cumulative hazard value corresponding to a particular failed unit is the sum of all the hazard values for failed units with ranks up to and including that failed unit.
4. Plot the cumulative hazard value vs the time of fail. If the exponential model is appropriate the plot will be approximately a straight line going through zero with slope λ .

- Example 3 Times to failure data for non-repairable components

Ten components were observed for up to 250 hours. Six failures occurred at 37, 73, 132, 195, 222 and 248 hours. Four components were taken off test without failing at the following run times: 50, 100, 200 and 250 hours.

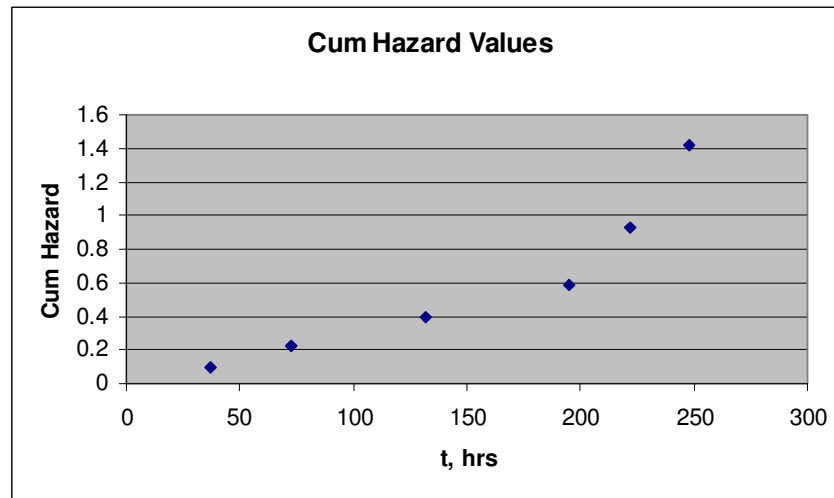
Solution : Cumulated hazard plot

The data from Example 3 were processed in accordance with the procedure presented earlier. The results are presented in the Table 5.

- Table 5 Estimation of cumulated hazard values.

(1) Time of Event	(2) 1= failure 0=runtime	(3) Rank	(4) Reverse Rank	(5) Haz Val (2) x 1/(4)	(6) Cum Hazard Value
37	1	1	10	1/10	0.10
50	0	2	9		
73	1	3	8	1/8	0.225
100	0	4	7		
132	1	5	6	1/6	0.391
195	1	6	5	1/5	0.591
200	0	7	4		
222	1	8	3	1/3	0.924
248	1	9	2	1/2	1.424
250	0	10	1		

The graph of cumulated hazard versus the time (column (1) vs column (6)) is shown on the Figure 5.



• Figure 5 Cumulated hazard plot for data from Example 3 Times to failure data for non-repairable components

The slope is approximately constant from 0 up to 200 hours and then the slope increases from 200 to 250 hours in operation. It could be an indication of failure rate increasing in time.

Construction of **failure rates side-by-side confidence intervals** for binned data

For each bin, calculate an estimate of failure rate λ , treating λ as if it were constant within the bin.

The simplest such estimate is the maximum likelihood estimate (MLE), which is the number of failures divided by total exposure time for the bin.

• Binned data for non-repairable components

The times to failure data of Example 3 were transferred to binned data using as a bin 50-hours times intervals, see Table 6.

Solution : side-by-side confidence intervals

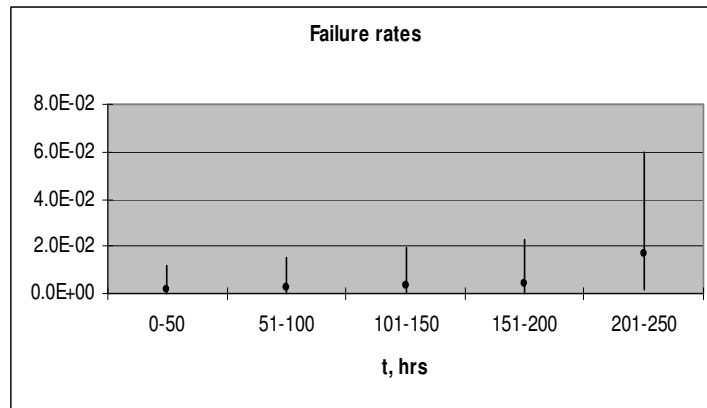
For each bin the point estimate (MLE), and lower and upper bounds of a 95% confidence interval were calculated.

• Table 6 Point estimates and side-by-side confidence intervals for binned data

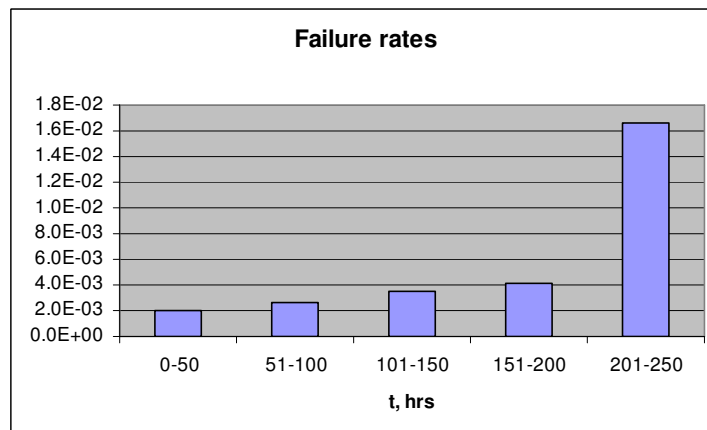
Bins	1	2	3	4	5
Operating hours	0-50	51-100	101-150	151-200	201-250
Number of failures	1	1	1	1	2
Components in operation	10	8	6	5	3
Cumulative operating time / number of demands	$9 \times 50 + 37 = 487$	$7 \times 50 + (73 - 50) = 373$	$5 \times 50 + (132 - 100) = 282$	$4 \times 50 + (195 - 150) = 245$	$50 + (222 - 200) + (248 - 200) = 120$

λ_k	0.0021	0.0027	0.0035	0.0041	0.017
λ_k^L	5.20E-05	6.79E-05	8.98E-05	1.03E-04	2.02E-03
λ_k^U	1.14E-02	1.49E-02	1.98E-02	2.27E-02	6.02E-02

Figure 6 presents the graphical interpretation of the results presented in the table.



a



b

• Figure 6 Side-by-side confidence intervals (a) and point estimates (a, b) of failure rates per bin

The graphs show a clear increase in failure rate for the last bin.

3.1.2.2 Repairable components (repair same-as-old)

Cumulative failures plot for times to failure data

The plot of cumulative number of failures versus system age is a step function that increases by one every time there is a new failure. The slope of this plot is an estimate of the rate of occurrence of failure (ROCOF). If

this plot looks linear, there is no obvious improvement (or degradation) trend. A bending downward indicates improvement; bending upward indicates degradation.¹

- Example 4 Times to failure data for repairable components

Ten components were observed up to 250 hours. Six failures occurred at 37, 73, 132, 195, 222 and 248 hours. All components had been repaired after the failures and continued to operate. Three components were taken off operation without failing at the following run times: 50, 100, 200 and the rest were operated until 250 hours.

Solution : Cumulative failures plot

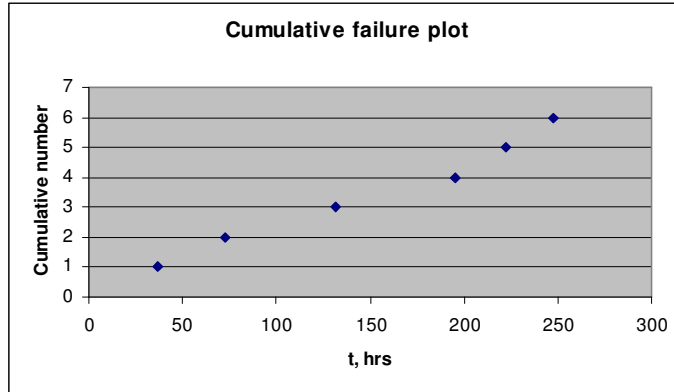
Cumulative number of failures was computed in the following table:

- Table 7 Calculation of cumulative number of failures

(1) Time of Event	(2) 1= failure 0=runtime	(6) Cumulative number of failures
37	1	1
50	0	
73	1	2
100	0	
132	1	3
195	1	4
200	0	
222	1	5
248	1	6
250	0	

Figure 7 presents the plot of cumulative failures versus cumulative age. The plot does not indicate an obvious trend.

¹ If repair is same-as-new (renewal process), the cumulative failure plot will be approximately a straight line regardless of how the failure rate is changing between failures.



• Figure 7 Cumulative failure plot for repairable components

Inter-arrival times between the failures.

This is another plot for visual examination of data for repairable components under the assumption that repair is same-as-old. The inter-arrival times between new failures (in other words, the waiting times between failures, with the time to the first failure used as the first "inter-arrival" time) are plotted against failure number. If the trend is increasing, this suggests improvement; if the trend is down, this suggests degradation.

• Example 4 Failure times of individual repairable component

The component was in operation for 250 hours. Six failures occurred at 37, 73, 132, 195, 222 and 248 hours. The component had been repaired after the failures and returned to operation. Three preventive maintenance activities were performed at the following times: 50, 100, 200 hours.

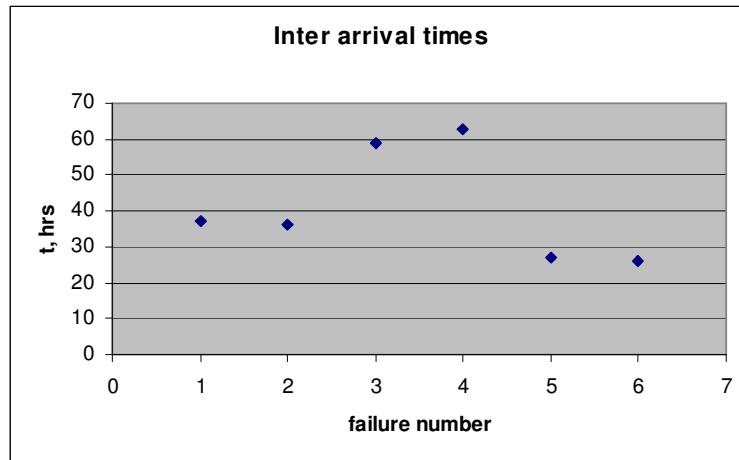
Solution : inter-arrival times between failures

Inter-arrival times are computed in the following table.

• Table 8 Calculation of inter-arrival times

Number of failure	Time of Event	1= failure 0=runtime	Inter-arrival time
1	37	1	37
	50	0	
2	73	1	36
	100	0	
3	132	1	59
4	195	1	63
	200	0	
5	222	1	27
6	248	1	26
	250	0	

The inter arrival times plot is presented on the Figure 8.



• Figure 8 Inter-arrival times plot for individual component.

The plot does not indicate a trend in this example.

3.1.3 Nonparametric tests

An alternative to a plot is a hypothesis test, which can give a quantitative answer to the question of whether ageing appears to be present, by measuring the strength of the evidence against the null hypothesis:

H_0 : no ageing occurs.

This contrasts with plots, which give a visual impression but nothing quantitative.

Three nonparametric tests are described below: Laplace test, inversion criteria test (a version of the Kendall τ -test), and a “two-cell” test (a simple variation of the chi-squared test).

3.1.3.1 Laplace test

The method considers the null hypothesis H_0 : λ is constant over time, and the alternative hypothesis H_1 : λ is either an increasing or a decreasing function of time.

This hypothesis says that the events tend to occur more at one end of the interval than at the other.

We can consider m components, with the i -th component observed from age t_{i1} to t_{i2} and having n_i failures occurring at successive random times during that time period.

Denote the random failure times by t_{ij} , for $i = 1$ to m and $j = 1$ to n_i . If no ageing is occurring, then each t_{ij} is uniformly distributed on (t_{i1}, t_{i2}) .

Let c_i denote the centre of the i -th interval, $(t_{i1} + t_{i2})/2$, and let w_i denote the width of the interval, $t_{i2} - t_{i1}$, then $(t_{ij} - c_i)/w_i$ is uniformly distributed on $(-1/2, +1/2)$, with mean 0 and variance $1/12$. The sum of all these random quantities,

$$\sum_{i=1}^m \sum_{j=1}^{n_j} \left(\frac{t_{ij} - c_i}{w_i} \right) \quad (3)$$

has mean 0 and variance $\Sigma n_i/12$. Therefore,

$$\sqrt{12/\Sigma n_i} \sum_{i=1}^m \sum_{j=1}^{n_j} \left(\frac{t_{ij} - c_i}{w_i} \right) \quad (4)$$

has mean 0 and variance 1, and by the Law of Large Numbers it is approximately normally distributed.

The approximation is very good when the t_j 's are uniform and there are 6 or more of them.

As ± 1.645 are the 95th and 5th percentiles, respectively, of the standard normal distribution, there is only about 5% chance to fail with the conclusion that null hypothesis is true if expression (4) is within the interval - 1.645 and +1.645. This assumes that no ageing occurs. In case when ageing is occurring, then more failures are expected late than early, and Expression (4) will tend to be larger than the case when no ageing was occurring. Thus, large values of Expression (4) give evidence of ageing.

3.1.3.2 Inversion test

The test could be applied for the statistical samples of repairable components installed at one or several units, when the components are identical in design and operation.

It is supposed that n observations of variable λ_i are distributed continuously and independently. The test considers the null hypothesis

H_0 : the λ_i are identically distributed

against the alternative that a trend is present. The method examines the sequence of age-ordered values of estimated failure rates $\{\lambda_i\}$, calculated for reasonably small, equal time intervals (Δt) belonging to the observation period.

For that, an "age window" is defined as a period of operation for which statistical information about failure events and component mission times is available. This "age window" is divided into M intervals ($t_i, t_i + \Delta t$). For the i th interval the estimated failure rate λ_i is

$$\lambda_i = n_i / \left(\sum_{k=1}^{c_i} \Delta t_k^i \right) \quad (5)$$

where:

n_i – total number of failures observed within the i th interval,

c_i – number of components been in operation during the i th interval,

Δt_k^i – time in operation of k th component at i th interval.

If all c_i components were in operation during whole interval duration the formula could be rewritten as

$$\lambda_i = n_i / (c_i \Delta t).$$

An “inversion” occurs if $i < j$ but $\lambda_i > \lambda_j$. There are $M(M-1)/2$ pairs (i, j) with $i < j$. If H_0 is true, about half of these pairs should correspond to inversions. The presence of many inversions indicates a decreasing trend, and the presence of few inversions indicates an increasing trend.

The inversion number A is

$$A = \sum_{i=1}^{M-1} \sum_{j=i+1}^M h_{ij}, \text{ where } h_{ij} = \begin{cases} 1 & \text{for } \lambda_i > \lambda_j \\ 0 & \text{for } \lambda_i \leq \lambda_j \end{cases} \tag{6}$$

The hypothesis of absence of trend, H_0 , is not rejected if

$$A_L < A \leq A_U.$$

The case of $A \leq A_L$ corresponds to the increasing trend of failure rate, which may be a result of ageing. If, instead, $A_U < A$ the trend is decreasing and it means that reliability is improved with time.

Here the lower (A_L) and the upper (A_U) inversion limits depend on the assumed confidence level α . The values of A_L and A_U as a function of M and α are given in Table 9.

• Table 9 Lower (A_L) and upper (A_U) inversion limits

M	A_L			A_U		
	$(1-\alpha/2) = 0.99$	$(1-\alpha/2) = 0.975$	$(1-\alpha/2) = 0.95$	$\alpha/2 = 0.05$	$\alpha/2 = 0.025$	$\alpha/2 = 0.01$
10	9	11	13	31	33	35
12	16	18	21	44	47	49
14	24	27	30	60	63	66
16	34	38	41	78	81	85
18	45	50	54	98	102	107
20	59	64	69	120	125	130
30	152	162	171	263	272	282
40	290	305	319	460	474	489
50	473	495	514	710	729	751
60	702	731	756	1013	1038	1067
70	977	1014	1045	1369	1400	1437
80	1299	1344	1382	1777	1815	1860
90	1668	1721	1766	2238	2283	2336
100	2083	2145	2198	2751	2804	2866

Two important remarks concerning the application of inversion test have to be noted:

- validity of the hypothesis about identical distribution of random variable on times intervals ;
- consideration of repeated, or identical, values of λ_i .

The first problem could be resolved by “equal-exposure” binning of the data, i.e. setting up the lengths of time intervals (Δt) not as equal calendar times, but as periods with approximately equal cumulative operating time of the components in operation in these periods.

The second remark concerns the fact that sometimes the statistic contains several intervals with repeated or zero failure counts, so that $\lambda_i = \lambda_j$, a “tie”. If the number of ties is more than 10% of M , it is recommended to use

Sean's correction, which essentially counts each tie as half an inversion.

- Example 5 Binned data for large sample of repairable NPP components

The data on cumulative number of failures and operating time within one-year intervals were collected for identical sensors installed in different technological systems of NPPs, see (Atwood L.C. and all, 2007). The data were arranged by year in operation and then failure intensities were estimated for each operating year, assuming a constant value within the year. Table 10 presents the distribution of the cumulative number of failures, operating time, and estimated failure intensities in one-year bins.

The data are “virtual”, but quite close to actual operating experience data collected on French, German, and US electrical generation plants. In particular, the data include large samples representative of components from the same technological group.

The failure rates were estimated for equal one-year intervals, the sequence of which represents the time in operation or age of the component. This data has two peculiarities :

- there are some intervals without failures, consequently, failure rates for these intervals are estimated as zero,
- the accumulated operating time is different from one interval to another, leading to differences in failure rate confidence intervals.

These peculiarities were taken into account during the data analysis.

- Table 10 Binned data for large sample of repairable NPP components

Interval number, i (age, years)	Number of failures, n	Cumulative operating time, Σt_i (comp.x years)	Failure intensity, λ_i (1/year)	Upper bound of λ_i	Lower bound of λ_i
1	-	-	-	-	-
2	0	9.59	0.00E+00	3.85E-01	0.00E+00
3	1	45.28	2.21E-02	1.23E-01	5.59E-04
4	0	78.00	0.00E+00	4.73E-02	0.00E+00
5	1	120.00	8.33E-03	4.64E-02	2.11E-04
6	2	295.43	6.77E-03	2.45E-02	8.19E-04
7	8	500.10	1.60E-02	3.15E-02	6.91E-03
8	7	585.00	1.20E-02	2.47E-02	4.81E-03
9	12	759.68	1.58E-02	2.76E-02	8.16E-03
10	15	1062.44	1.41E-02	2.33E-02	7.90E-03
11	15	1082.37	1.39E-02	2.28E-02	7.76E-03
12	10	1046.68	9.55E-03	1.76E-02	4.58E-03
13	21	1014.00	2.07E-02	3.16E-02	1.28E-02
14	25	971.88	2.57E-02	3.79E-02	1.67E-02
15	16	796.42	2.01E-02	3.26E-02	1.15E-02
16	17	591.71	2.87E-02	4.60E-02	1.68E-02
17	16	507.00	3.16E-02	5.12E-02	1.81E-02
18	5	332.01	1.51E-02	3.51E-02	4.89E-03
19	0	29.37	0.00E+00	1.26E-01	0.00E+00

Solution : inversion test

Consider the failure intensities for the bins from 2 up to 19. The matrix in Figure 9 permits us to calculate the number of inversions as defined by equation (4). In the matrix for each row from $i = 1$ to $M+1$, the inversion is defined for each cell $j = i+1, M$. For example, cell H9 presents the inversion value for $i = 4$ and $j = 6$, with the value defined by the formula $H9 = IF(\$F\$4>H\$4,1,0)$. The last row in the matrix provides the sums of the inversions for each column j and cell C24 gives the total number of inversions by summing across the columns.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	bins	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
2	# of failures	0	1	0	1	2	8	7	12	15	15	10	21	25	16	17	16	5	0	0
3	accum.operatin	9.594	45.279	78	120.003	295.425	500.097	585	759.681	1062.438	1082.367	1046.882	1014	971.88	796.419	591.708	507	332.007	29.367	0
4	$\lambda_i(t)$	0.00E+00	2.21E-02	0.00E+00	8.33E-03	6.77E-03	1.60E-02	1.20E-02	1.58E-02	1.41E-02	1.39E-02	9.55E-03	2.07E-02	2.57E-02	2.01E-02	2.87E-02	3.16E-02	1.51E-02	0.00E+00	0
5		$j=i, M+1$	$j=i, M$	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
6	1			0																
7	2				1		1		1		1		1		1		0	1		0
8	3					0	0		0		0		0		0		0	0		0
9	4						1		0		0		0		0		0	0		0
10	5							0			0		0		0		0	0		0
11	6								1		1		1		1		0	0		1
12	7									0			1		0		0	0		0
13	8										1		0		0		0	1		1
14	9											1		1		0	0	0		0
15	10												1		0	0	0	0		0
16	11													0		0	0	0		0
17	12														0	1	0	0		1
18	13															1	0	0		1
19	14																0	0		1
20	15																	0		1
21	16																		1	1
22	17																			1
23																				
24	Number of Inversions, A		49	0	1	1	2	1	2	2	3	3	5	1	0	3	1	1	8	15

• Figure 9 Calculation of inversions number

The result of the inversion test for the given data set shows an increasing trend in the failure intensity. For $M = 18$ the total number of inversions is $A = 49$ which is less than $A_L = 50$ at a 95% confidence limit.

3.1.3.3 Two-cells test

This very simple test is similar to a particular case of the chi-squared test. It begins by partitioning the data into two segments, "early" and "late", with approximately half of the exposure time in each segment.

Define

n = total number of failures

s_E = early exposure time, (that is, component-years in the early period)

s_L = later exposure time,

$s_T = s_E + s_L$ total exposure time

X_L = number of failures in the late period,

$p = s_L/s_T$.

The standard chi-squared test would construct a two-sided test of the hypothesis that the failure rate is constant. We prefer to have the possibility of a one-sided test, and so consider X_L itself, rather than the chi-squared statistic. If the failure rate is constant, then X_L has a binomial(n, p) distribution. The probability of observing a value as large as X_L (or larger) can be found with the help of many computer packages. Alternatively, an approximation can be used, because

$$Z = (X_L - np) / \sqrt{np(1-p)} \tag{7}$$

has approximately a standard normal distribution. If $Z > 1.645$, H_0 could be rejected at approximately the 5% significance level, in favor of the hypothesis that λ is increasing. A value greater than 1.28 shows evidence of ageing that is significant at approximately the 10% level.

3.1.4 Parametric tests

3.1.4.1 Task specification

For binned data on failure intensities we consider the following parametric models :

1. Constant failure intensity : $\lambda(t) = \varphi(\bar{\theta}; t) = \text{Const}$;
2. Linear failure intensity : $\lambda(t) = \varphi(\bar{\theta}; t) = \theta_1 + \theta_2 t$;
3. Log-linear or exponential failure intensity : $\lambda(t) = \varphi(\bar{\theta}; t) = \theta_1 \exp(\theta_2 t)$ or $\ln \varphi(\bar{\theta}; t) = \theta_1 + \theta_2 t$;
4. Power-law (Weibull) failure intensity model : $\lambda(t) = \varphi(\bar{\theta}; t) = \theta_1 t^{\theta_2}$

For models 2-4, $\theta_2 > 0$ implies an increasing trend in failure intensity with time, i.e., component failure intensity increases with age of the component.

The test consists of trying to find evidence to reject the null hypothesis

$$H_0: \lambda(t) = \text{const}$$

in favour of the alternative hypothesis

$$H_1: \lambda(t) \text{ is increasing in } t.$$

The test could be performed by checking if the “ageing” parameter is 0. In other words H_0 could be rejected if the confidence interval for the “ageing” parameter is strictly greater than 0.

Two steps are involved, and both must be performed. They involve questions about two different hypotheses.

1. Does the model fit the data adequately? Answering this question involves a goodness-of-fit test, or model validation.
2. If so, is the ageing parameter significantly greater than zero?

3.1.4.2 Goodness-of-fit test and parameter estimation

We will use Fisher’s chi-square test to check goodness-of-fit, the statistic of which is:

$$\chi^2(\bar{\theta}) = \sum_{i=1}^s \frac{[v(\Delta_i) - \varphi(\bar{\theta}; t_i) T_i]^2}{\varphi(\bar{\theta}; t_i) T_i} \tag{8}$$

where $\varphi(\bar{\theta}; t)$ is one of the four functions proposed to model the failure rate $\lambda(t)$.

Here

$\Delta_1, \Delta_2, \dots, \Delta_s$ are the durations of each interval (bin),

$\nu(\Delta_i)$ is the number of failures in interval Δ_i ,

T_i is the cumulative operating time of all components that have been in operation within the interval Δ_i .

The hypothesis to be tested is presented as follows :

$$H_0 : \exists \bar{\theta} : \lambda_i = \varphi(\bar{\theta}; t_i) , \quad (9)$$

where λ_i is the averaged failure intensity per interval Δ_i .

To calculate the Fisher chi-square statistic, an unknown value of $\bar{\theta}$ is replaced with an estimate $\hat{\bar{\theta}}$, obtained using the method of minimum χ^2 :

$$\hat{\bar{\theta}} = \arg \min_{\bar{\theta}} \chi^2(\bar{\theta}) . \quad (10)$$

The criterion for testing the null hypothesis of conformity with the proposed model is a simple comparison of p -value and a chosen confidence level value α . A p -value is calculated from:

$$p = \int_z^{\infty} f_{\chi^2_{s-r}}(t) dt \quad (11)$$

where

$$z = \chi^2(\hat{\bar{\theta}}) ;$$

$f_{\chi^2_{s-r}}(t)$ is the density of a chi-square distribution with $s-r$ degrees of freedom,

s is the number of group intervals Δ_i (where $\nu(\Delta_i)$ should differ from 0),

r is the number of parameters estimated. For the constant failure intensity model (model 1) $r = 1$, for the other proposed models (2 – 4) $r = 2$.

The *null* hypothesis is not rejected if $p > \alpha$, otherwise it is rejected. If several models cannot be rejected, preference is given to the model with the largest p -value.

The detailed procedure for parameter estimation and model verification is presented in the example below using MS Excel tools.

To illustrate the method application let's use the data from Example 5.

Solution : goodness-of-fit test

Step 1 : merge bins in the initial data set in order to avoid having bins for which the estimated failure intensity is zero. For this example the following bins were merged : bins 2 and 3, 4 and 5, and bins 18 and 19, see Table 10 and Figure 10 Processed data for goodness-of-fit test Figure 10.

Step 2 : in the Excel worksheet create a table for each of the four potential models, ordering data as shown in Figure 10.

	A	B	C	D	E	F	G
1	Intervals	t	v(Δ)	Op. time, T	λ(t)	Ψ(t)	Const
2	2-3	2	1	54.87	1.86E-02	0.0005	θ1
3	4-5	4	1	198.00	1.86E-02	1.9622	0.01864
4	6	5.5	2	295.43	1.86E-02	2.2337	
5	7	6.5	8	500.10	1.86E-02	0.1877	
6	8	7.5	7	585.00	1.86E-02	1.3988	opt ΣΨ(t)
7	9	8.5	12	759.68	1.86E-02	0.3301	24.3964
8	10	9.5	15	1062.44	1.86E-02	1.1664	p-value
9	11	10.5	15	1082.37	1.86E-02	1.3287	0.04102
10	12	11.5	10	1046.68	1.86E-02	4.6376	
11	13	12.5	21	1014.00	1.86E-02	0.2325	
12	14	13.5	25	971.88	1.86E-02	2.6139	
13	15	14.5	16	796.42	1.86E-02	0.0895	
14	16	15.5	17	591.71	1.86E-02	3.2301	
15	17	16.5	16	507.00	1.86E-02	4.5368	
16	18-19	18	5	361.37	1.86E-02	0.4478	

• Figure 10 Processed data for goodness-of-fit test

Here, column A (“Intervals”) shows how the initial one-year bins were merged.

Column B (“t”) is the middle of the bin, the point in time where the failure intensity function $\varphi(\bar{\theta}; t)$ is estimated. For example, the first bin (cell A2) covers the period between 1 and 3 years of operation, so the middle point (cell B2) is 2.0.

Columns C and D (“v(Δ)”) and (“Op.time, T”) provide the cumulative number of failures and operating time within the bin.

Column E (“λ(t)”) shows the estimated intensity function $\varphi(\bar{\theta}; t)$ using the parameters θ_1, θ_2 obtained by minimizing equation (10). Figure 10 illustrates this for the constant failure intensity model, so here $\lambda(t) = \theta_1 = \$G\$3$.

Column F (“Ψ(t)”) represents the summed elements of equation (8) : for example, for cell F2, “Ψ(t)” =((C2-D2*E2)^2)/(D2*E2).

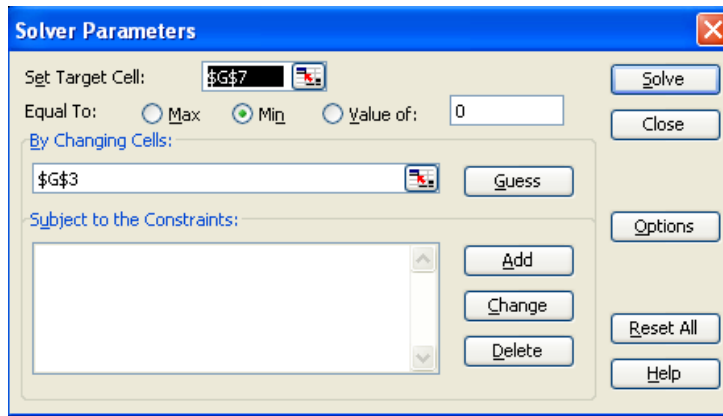
Column G (“const”) contains the estimated parameters θ_1, θ_2 , “opt ΣΨ(t)”, and the p-value. In the given example, the estimated value of θ_1 is stored in cell G3, the sum of equation (8) “opt ΣΨ(t)” =SUM(F2:F16) is placed in cell G7, and the “p-value” =CHIDIST(G7,14) is in cell G9.

For the p-value calculation the degrees of freedom ($s-r$) of the chi-square distribution is entered manually. In the example, the degrees of freedom value is $(s-r) = (15-1) = 14$.

Step 3 : resolve the equation (10) with the help of the Excel add-in SOLVER. In order to do that go to the “Tools” and open the “Solver” dialog window, see Figure 11. Specify “opt $\Sigma\Psi(t)$ ” as the target cell, i.e., cell G7 in the present example. Set this cell as equal to “min” with the “changing cells” as the parameters to be estimated, i.e., cell G3 in this example for a constant failure intensity model. Then activate the calculation by pressing the “Solve” button.

Note : before launching SOLVER initial estimates of the model parameters (θ, θ_2) have to be specified. One could use a least-squares estimate (see Excel Graph Wizard options) or a simple constant failure intensity approximation.

For the initial estimates, do not use more than two significant figures.



• Figure 11 Excel SOLVER dialog window

Step 4 : compare the results for different models and choose the best fitting model (the one with the largest p-value). Figure 12 presents the results of the goodness-of-fit test for models 1-4 applied to the data from Example 5.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Intervals	t	v(Δ)	Op. time, T	$\lambda(t)$	$\Psi(t)$	Const	Intervals	t	v(Δ)	Op. time, T	$\lambda(t)$	$\Psi(t)$	Linear
2	2-3	2	1	54.87	1.86E-02	0.0005	01	2-3	2	1	54.87	6.70E-03	1.0865	01
3	4-5	4	1	198.00	1.86E-02	1.9622	0.01864	4-5	4	1	198.00	9.14E-03	0.3618	0.00427
4	6	5.5	2	295.43	1.86E-02	2.2337		6	5.5	2	295.43	1.10E-02	0.4735	02
5	7	6.5	8	500.10	1.86E-02	0.1877		7	6.5	8	500.10	1.22E-02	0.5988	0.00122
6	8	7.5	7	585.00	1.86E-02	1.3988	opt $\Sigma\Psi(t)$	8	7.5	7	585.00	1.34E-02	0.0892	opt $\Sigma\Psi(t)$
7	9	8.5	12	759.68	1.86E-02	0.3301	24.3964	9	8.5	12	759.68	1.46E-02	0.0730	13.2438
8	10	9.5	15	1062.44	1.86E-02	1.1664	p-value	10	9.5	15	1062.44	1.58E-02	0.1962	p-value
9	11	10.5	15	1082.37	1.86E-02	1.3287	0.04102	11	10.5	15	1082.37	1.70E-02	0.6447	0.42916
10	12	11.5	10	1046.68	1.86E-02	4.6376		12	11.5	10	1046.68	1.83E-02	4.3458	
11	13	12.5	21	1014.00	1.86E-02	0.2325		13	12.5	21	1014.00	1.95E-02	0.0790	
12	14	13.5	25	971.88	1.86E-02	2.6139		14	13.5	25	971.88	2.07E-02	1.1874	
13	15	14.5	16	796.42	1.86E-02	0.0895		15	14.5	16	796.42	2.19E-02	0.1206	
14	16	15.5	17	591.71	1.86E-02	3.2301		16	15.5	17	591.71	2.31E-02	0.8029	
15	17	16.5	16	507.00	1.86E-02	4.5368		17	16.5	16	507.00	2.43E-02	1.0836	
16	18-19	18	5	361.37	1.86E-02	0.4478		18-19	18	5	361.37	2.62E-02	2.1007	
17	Intervals	t	v(Δ)	Op. time, T	$\lambda(t)$	$\Psi(t)$	Exponent	Intervals	t	v(Δ)	Op. time, T	$\lambda(t)$	$\Psi(t)$	Weibull
18	2-3	2	1	54.87	8.31E-03	0.6484	01*	2-3	2	1	54.87	5.19E-03	1.7952	01
19	4-5	4	1	198.00	9.74E-03	0.4469	0.0071	4-5	4	1	198.00	8.58E-03	0.2881	0.00314
20	6	5.5	2	295.43	1.10E-02	0.4744	02	6	5.5	2	295.43	1.08E-02	0.4471	02
21	7	6.5	8	500.10	1.19E-02	0.7176	0.07916	7	6.5	8	500.10	1.22E-02	0.5875	0.72557
22	8	7.5	7	585.00	1.28E-02	0.0354	opt $\Sigma\Psi(t)$	8	7.5	7	585.00	1.35E-02	0.1078	opt $\Sigma\Psi(t)$
23	9	8.5	12	759.68	1.39E-02	0.1952	12.4410	9	8.5	12	759.68	1.48E-02	0.0475	14.1202
24	10	9.5	15	1062.44	1.51E-02	0.0614	p-value	10	9.5	15	1062.44	1.61E-02	0.2542	p-value
25	11	10.5	15	1082.37	1.63E-02	0.3932	0.49185	11	10.5	15	1082.37	1.73E-02	0.7375	0.36544
26	12	11.5	10	1046.68	1.76E-02	3.8747		12	11.5	10	1046.68	1.85E-02	4.5056	
27	13	12.5	21	1014.00	1.91E-02	0.1401		13	12.5	21	1014.00	1.96E-02	0.0611	
28	14	13.5	25	971.88	2.07E-02	1.2070		14	13.5	25	971.88	2.07E-02	1.1586	
29	15	14.5	16	796.42	2.24E-02	0.1835		15	14.5	16	796.42	2.19E-02	0.1134	
30	16	15.5	17	591.71	2.42E-02	0.5014		16	15.5	17	591.71	2.29E-02	0.8657	
31	17	16.5	16	507.00	2.62E-02	0.5567		17	16.5	16	507.00	2.40E-02	1.2061	
32	18-19	18	5	361.37	2.95E-02	3.0050		18-19	18	5	361.37	2.56E-02	1.9449	

- Figure 12 Presentation of goodness-of-fit test results. Base case

The p -values are as follows :

- Constant failure intensity model (1), $p = 0,04$,
- Linear model (2), $p = 0,43$,
- Exponential or log-linear model (3), $p = 0,49$,
- Weibull or power low model (4), $p = 0,37$.

These results show that the constant model does not fit the data, but the other three models give a reasonable fit (p -values for all three models are quite close to each other). The best fitting model is the exponential one.

Cases where the data cannot be fit to any model (insufficient, inadequate or sparse data) have to be also clearly identified and documented. The cut-off criteria have to be established. For example, if all models fit the data with p -value $< 0,1$ this component group could be eliminated from subsequent analysis.

Consideration of burn-in failures could improve the result of the goodness of fit test and could change the conclusion about existence or absence of an ageing trend. If visual examination shows some indication of burn-in failures, a sensitivity analysis for the impact of burn-in failures is recommended. Figure 13 presents the results of such an analysis for the data of Example 5. Here, the first bin (2-3) was excluded and calculations were performed for bins 2-15 (marked in yellow)

The results of the sensitivity analysis for this example are as follows :

- Constant failure intensity model (1), $p = 0,03$,
- Linear model (2), $p = 0,468$,
- Exponential or log-linear model (3), $p = 0,467$,
- Weibull or power low model (4), $p = 0,466$.

Again, all three time-dependent models fit the data at about the same level of significance, but now, the best fitting model is the linear.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Intervals	t	v(Δ)	Op. time T	λ(t)	Ψ(T)	Const	Intervals	t	v(Δ)	Op. time T	λ(t)	Ψ(t)	Linear
2	2-3	2	1	54.87	1.86E-02	0.0005	θ1	2-3	2	1	54.87	4.33E-03	2.4502	θ1
3	4-5	4	1	198.00	1.86E-02	1.9626	0.01864	4-5	4	1	198.00	7.23E-03	0.1306	0.00142
4	6	5.5	2	295.43	1.86E-02	2.2343		6	5.5	2	295.43	9.42E-03	0.2197	θ2
5	7	6.5	8	500.10	1.86E-02	0.1881		7	6.5	8	500.10	1.09E-02	1.2088	0.00145
6	8	7.5	7	585.00	1.86E-02	1.3996	opt ΣΨ(t)	8	7.5	7	585.00	1.23E-02	0.0061	opt ΣΨ(t)
7	9	8.5	12	759.68	1.86E-02	0.3306	24.3959	9	8.5	12	759.68	1.38E-02	0.2241	11.7239
8	10	9.5	15	1062.44	1.86E-02	1.1674	p-value	10	9.5	15	1062.44	1.52E-02	0.0868	p-value
9	11	10.5	15	1082.37	1.86E-02	1.3299	0.02768	11	10.5	15	1082.37	1.67E-02	0.5195	0.4681
10	12	11.5	10	1046.68	1.86E-02	4.6394		12	11.5	10	1046.68	1.81E-02	4.2561	
11	13	12.5	21	1014.00	1.86E-02	0.2320		13	12.5	21	1014.00	1.96E-02	0.0640	
12	14	13.5	25	971.88	1.86E-02	2.6119		14	13.5	25	971.88	2.11E-02	1.0073	
13	15	14.5	16	796.42	1.86E-02	0.0892		15	14.5	16	796.42	2.25E-02	0.2067	
14	16	15.5	17	591.71	1.86E-02	3.2282		16	15.5	17	591.71	2.40E-02	0.5617	
15	17	16.5	16	507.00	1.86E-02	4.5346		17	16.5	16	507.00	2.54E-02	0.7527	
16	18-19	18	5	361.37	1.86E-02	0.4482		18-19	18	5	361.37	2.76E-02	2.4797	
17	Intervals	t	v(Δ)	Op. time T	λ(t)	Ψ(t)	Exponenti	Intervals	t	v(Δ)	Op. time T	λ(t)	Ψ(t)	Weibull
18	2-3	2	1	54.87	7.85E-03	0.7515	θ1*	2-3	2	1	54.87	3.63E-03	3.2133	θ1
19	4-5	4	1	198.00	9.29E-03	0.3836	0.00664	4-5	4	1	198.00	6.88E-03	0.0965	0.00192
20	6	5.5	2	295.43	1.05E-02	0.3993	θ2	6	5.5	2	295.43	9.23E-03	0.1933	θ2
21	7	6.5	8	500.10	1.15E-02	0.8929	0.0842	7	6.5	8	500.10	1.08E-02	1.2734	0.92097
22	8	7.5	7	585.00	1.25E-02	0.0123	opt ΣΨ(t)	8	7.5	7	585.00	1.23E-02	0.0046	opt ΣΨ(t)
23	9	8.5	12	759.68	1.36E-02	0.2760	11.7433	9	8.5	12	759.68	1.38E-02	0.2245	11.7514
24	10	9.5	15	1062.44	1.48E-02	0.0303	p-value	10	9.5	15	1062.44	1.53E-02	0.0914	p-value
25	11	10.5	15	1082.37	1.61E-02	0.3281	0.46651	11	10.5	15	1082.37	1.67E-02	0.5362	0.46584
26	12	11.5	10	1046.68	1.75E-02	3.7591		12	11.5	10	1046.68	1.82E-02	4.2998	
27	13	12.5	21	1014.00	1.90E-02	0.1538		13	12.5	21	1014.00	1.97E-02	0.0576	
28	14	13.5	25	971.88	2.07E-02	1.1943		14	13.5	25	971.88	2.11E-02	0.9858	
29	15	14.5	16	796.42	2.25E-02	0.2054		15	14.5	16	796.42	2.25E-02	0.2109	
30	16	15.5	17	591.71	2.45E-02	0.4377		16	15.5	17	591.71	2.40E-02	0.5620	
31	17	16.5	16	507.00	2.66E-02	0.4633		17	16.5	16	507.00	2.54E-02	0.7625	
32	18-19	18	5	361.37	3.02E-02	3.2071		18-19	18	5	361.37	2.75E-02	2.4527	

• Figure 13 Presentation of goodness-of-fit test results. Early failures excluded

The next stage of the analysis consists of formal verification of the trend hypothesis for the best fitting model.

3.1.4.3 Trend verification

As previously mentioned, for models 2-4, the fact that parameter $\theta_2 > 0$ implies an increasing trend in failure intensity with time, i.e., component failure rate increases with age of the component. A verification of such a trend can be made using the following hypothesis test.

The null hypothesis (absence of trend) is specified as : $H_0 : \theta_2 = 0$

The alternative hypothesis (increasing trend) is $H_1 : \theta_2 > 0$

The test is intimately related to the confidence interval for θ_2 .

The null hypothesis (H_0) is rejected in favour of hypothesis H_1 at significance level ϵ if and only if the $100(1-\epsilon)\%$ confidence interval for θ_2 is to the right of zero.

The confidence interval is obtained numerically as follows.

The chi-square statistic with fixed parameter θ_1 can be used:

$$\chi^2(\hat{\theta}_1, \theta_2) = \sum_{i=1}^s \frac{[v(\Delta_i) - \varphi(\hat{\theta}_1, \theta_2; t_i) \cdot T_i]^2}{\varphi(\hat{\theta}_1, \theta_2; t_i) \cdot T_i} \tag{12}$$

Here $\hat{\theta}_1$ is the point estimate of θ_1 obtained by solving equation (10).

The same procedure described above for the goodness-of-fit test is then applied.

Solving the equation

$$\varepsilon = \int_{\mu_\varepsilon}^{\infty} f_{\chi^2_{s-1}}(t) dt$$

with a given ε value, determines the parameter μ_ε .

It should be noted that the degrees of freedom has increased by 1, due to $r = 1$.

Then the following transcendental inequality is solved numerically:

$$\chi^2(\hat{\theta}_1, \theta_2) = \sum_{i=1}^s T_i \frac{\left[\frac{\nu(\Delta_i)}{T_i} - \varphi(\hat{\theta}_1, \theta_2; t_i) \right]^2}{\varphi(\hat{\theta}_1, \theta_2; t_i)} \leq \mu_\varepsilon \quad (13)$$

The solution of (13) is the confidence interval for parameter : $\theta_2 : (\underline{\theta}_2; \bar{\theta}_2)$.

It should be stressed that calculation of lower and upper bounds of confidence interval $(\underline{\theta}_2; \bar{\theta}_2)$ is only possible, if the p -value of the point estimate $\hat{\theta}$ meets the condition: $p > \varepsilon$.

For the purpose of detecting an increasing trend, it is sufficient to estimate only the lower bound of θ_2 .

Solution : estimation of lower bound for θ_2

To find the lower bound of “ageing” parameter the inequality (13) was transformed to an equality;

$$\sum_{i=1}^s T_i \frac{\left[\frac{\nu(\Delta_i)}{T_i} - \varphi(\hat{\theta}_1, \hat{\theta}_2 - y; t_i) \right]^2}{\varphi(\hat{\theta}_1, \theta_2; t_i)} = \mu_\varepsilon \quad (14)$$

here,

$\underline{\theta}_2 = \hat{\theta}_2 - y$ is a lower bound,

$\hat{\theta}_2$ is the point estimate obtained from (10),

$y = \hat{\theta}_2 - \underline{\theta}_2$ is the deviation of the parameter.

To perform the calculation with Excel Solver, the following additional parameters have to be added to the Excel tables used in 3.1.4.2 (see Figure 14) :

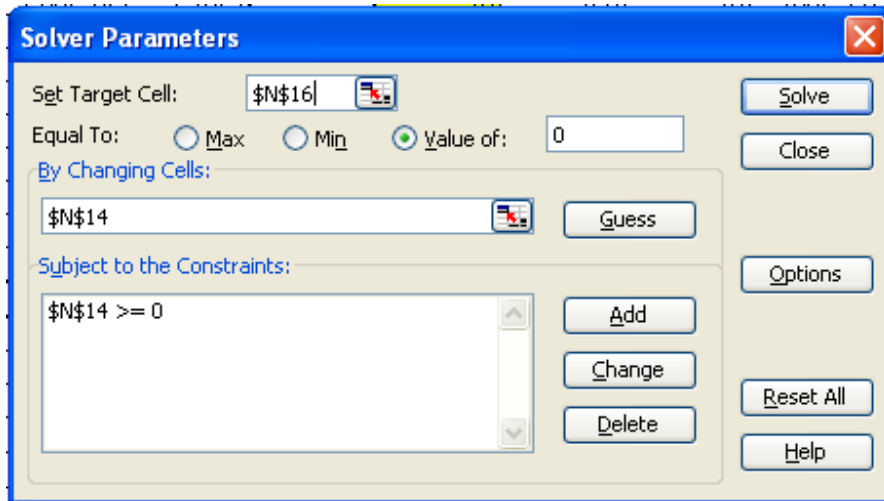
- for the linear, exponential and Weibull models the significance level ε is specified in cells N12, G28, and N28 respectively,

- ☑ the initial estimate for y is specified in cells N14, G30, and N30,
- ☑ the calculations for $\lambda(t)$ are modified in accordance with (14) for the linear, exponential and Weibull models. For the linear model, for example, the values in cells L2-L16 are calculated now as $\lambda(t) = \varphi(\hat{\theta}_1, \hat{\theta}_2 - y, t_i) = =\$N\$3+(\$N\$5-N\$14)*I2$
- ☑ equation (14) is entered in cell N16 for the linear model : N16 =N7-CHINV(N12,13), here $\mu_e = \text{CHINV}(N12,13)$ is on the left side of equation (14). The same operations are specified for the exponential and Weibull models in cells G32 and N32, respectively,
- ☑ finally, the value of $\theta_2 = \hat{\theta}_2 - y$ is entered in cell O5 (O21 and G34 for the exponential and Weibull models).

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Intervals	t	$v(\Delta)$	Op. time T	$\lambda(t)$	$\Psi(t)$	Const	Intervals	t	$v(\Delta)$	Op. time T	$\lambda(t)$	$\Psi(t)$	Linear	
2	2-3	2	1	54.87	1.86E-02	0.0005	#1	2-3	2	1	54.87	3.72E-03	3.1060	#1	
3	4-5	4	1	198.00	1.86E-02	1.9626	0.018645	4-5	4	1	198.00	6.02E-03	0.0308	0.001417	
4	6	5.5	2	295.43	1.86E-02	2.2343		6	5.5	2	295.43	7.74E-03	0.0362	#2	$\theta_2 - y$
5	7	6.5	8	500.10	1.86E-02	0.1881		7	6.5	8	500.10	8.89E-03	2.8364	0.001454	0.00115
6	8	7.5	7	585.00	1.86E-02	1.3996	opt $\Sigma\Psi(t)$	8	7.5	7	585.00	1.00E-02	0.2149	opt $\Sigma\Psi(t)$	
7	9	8.5	12	759.68	1.86E-02	0.3306	24.39585	9	8.5	12	759.68	1.12E-02	1.4365	19.81193	
8	10	9.5	15	1062.44	1.86E-02	1.1674		10	9.5	15	1062.44	1.23E-02	0.2705		
9	11	10.5	15	1082.37	1.86E-02	1.3299	0.027678	11	10.5	15	1082.37	1.35E-02	0.0105	0.070728	
10	12	11.5	10	1046.68	1.86E-02	4.6394		12	11.5	10	1046.68	1.46E-02	1.8532		
11	13	12.5	21	1014.00	1.86E-02	0.2320		13	12.5	21	1014.00	1.58E-02	1.5496	ϵ	
12	14	13.5	25	971.88	1.86E-02	2.6119		14	13.5	25	971.88	1.69E-02	4.4170	0.1	
13	15	14.5	16	796.42	1.86E-02	0.0892		15	14.5	16	796.42	1.81E-02	0.1747	y	
14	16	15.5	17	591.71	1.86E-02	3.2282		16	15.5	17	591.71	1.92E-02	2.7642	0.000304	
15	17	16.5	16	507.00	1.86E-02	4.5346		17	16.5	16	507.00	2.04E-02	3.0955	opt2	
16	18-19	18	5	361.37	1.86E-02	0.4482		18-19	18	5	361.37	2.21E-02	1.1220	0.0000	
17	Intervals	t	$v(\Delta)$	Op. time T	$\lambda(t)$	$\Psi(t)$	Exponenti	Intervals	t	$v(\Delta)$	Op. time T	$\lambda(t)$	$\Psi(t)$	Weibull	
18	2-3	2	1	54.87	7.49E-03	0.8438	#1*	2-3	2	1	54.87	3.34E-03	3.6342	#1	
19	4-5	4	1	198.00	8.46E-03	0.2716	0.006636	4-5	4	1	198.00	5.82E-03	0.0203	0.00192	
20	6	5.5	2	295.43	9.26E-03	0.1979	#2	6	5.5	2	295.43	7.51E-03	0.0217	#2	$\theta_2 - y$
21	7	6.5	8	500.10	9.84E-03	1.9274	0.084202	7	6.5	8	500.10	8.59E-03	3.1958	0.920966	0.800442
22	8	7.5	7	585.00	1.05E-02	0.1280	opt $\Sigma\Psi(t)$	8	7.5	7	585.00	9.63E-03	0.3312	opt $\Sigma\Psi(t)$	
23	9	8.5	12	759.68	1.11E-02	1.5043	27.68825	9	8.5	12	759.68	1.06E-02	1.8930	27.68825	
24	10	9.5	15	1062.44	1.18E-02	0.4839		10	9.5	15	1062.44	1.16E-02	0.5622		
25	11	10.5	15	1082.37	1.25E-02	0.1507	0.006143	11	10.5	15	1082.37	1.26E-02	0.1343	0.006143	
26	12	11.5	10	1046.68	1.33E-02	1.1146		12	11.5	10	1046.68	1.36E-02	1.2387		
27	13	12.5	21	1014.00	1.42E-02	3.0814	ϵ	13	12.5	21	1014.00	1.45E-02	2.7013	ϵ	
28	14	13.5	25	971.88	1.50E-02	7.3826	0.01	14	13.5	25	971.88	1.54E-02	6.6964	0.01	
29	15	14.5	16	796.42	1.60E-02	0.8440	y	15	14.5	16	796.42	1.63E-02	0.6918	y	
30	16	15.5	17	591.71	1.70E-02	4.8191	0.023616	16	15.5	17	591.71	1.72E-02	4.5531	0.120523	
31	17	16.5	16	507.00	1.80E-02	5.1429	opt2	17	16.5	16	507.00	1.81E-02	5.0702	opt2	
32	18-19	18	5	361.37	1.97E-02	0.6397	0.0000	18-19	18	5	361.37	1.94E-02	0.5782	0.0000	
33							$\theta_2 - y$								
34							0.060586								

• Figure 14 Calculation of “ageing” parameter lower confidence limit

To calculate the lower bound of the “ageing” parameter for the linear model, open SOLVER and set cell N16 as the target cell, which has to be equal to 0, see Figure 15. Next, specify the cells to change (“y”) as N14 and add the constrains for the calculation process, as N14 ≥ 0 . This will assure a positive value of y , so the second solution of equation (14), the upper bound of θ_2 , will not be considered.



• Figure 15 Excel SOLVER dialog window for finding the lower bound of the ageing parameter

The analysis for the linear model (the best fitting model for the example data with burn-in failures excluded) gives the following results:

lower 90% bound of “ageing” parameter $\frac{\theta_2}{2} = 1,15 \cdot 10^{-3}$.

This suggests the presence of a monotone increasing trend of failure intensity in time (at a 0,1 significance level).

3.2 Reliability parameters specification and elaboration

3.2.1 Generic process for PSA time-dependent parameter estimation

The ageing could affect the reliability of one or more SSCs. In case of sufficient reliability data, age-dependent reliability models could be constructed and introduced into the PSA.

For active components the age-dependent reliability parameters could be considered at the level of the fault trees by assigning time-dependent unavailabilities for corresponding basic events in the fault trees.

The component time-dependent unavailability could be used in the calculation of the probability of functional events, as well as for estimation of initiating event frequencies. Both of the cases have to be considered in the input parameter specification.

Depending on the available PSA code and the expected applications the technique for introducing time-dependent unavailability at the level of basic events could be different.

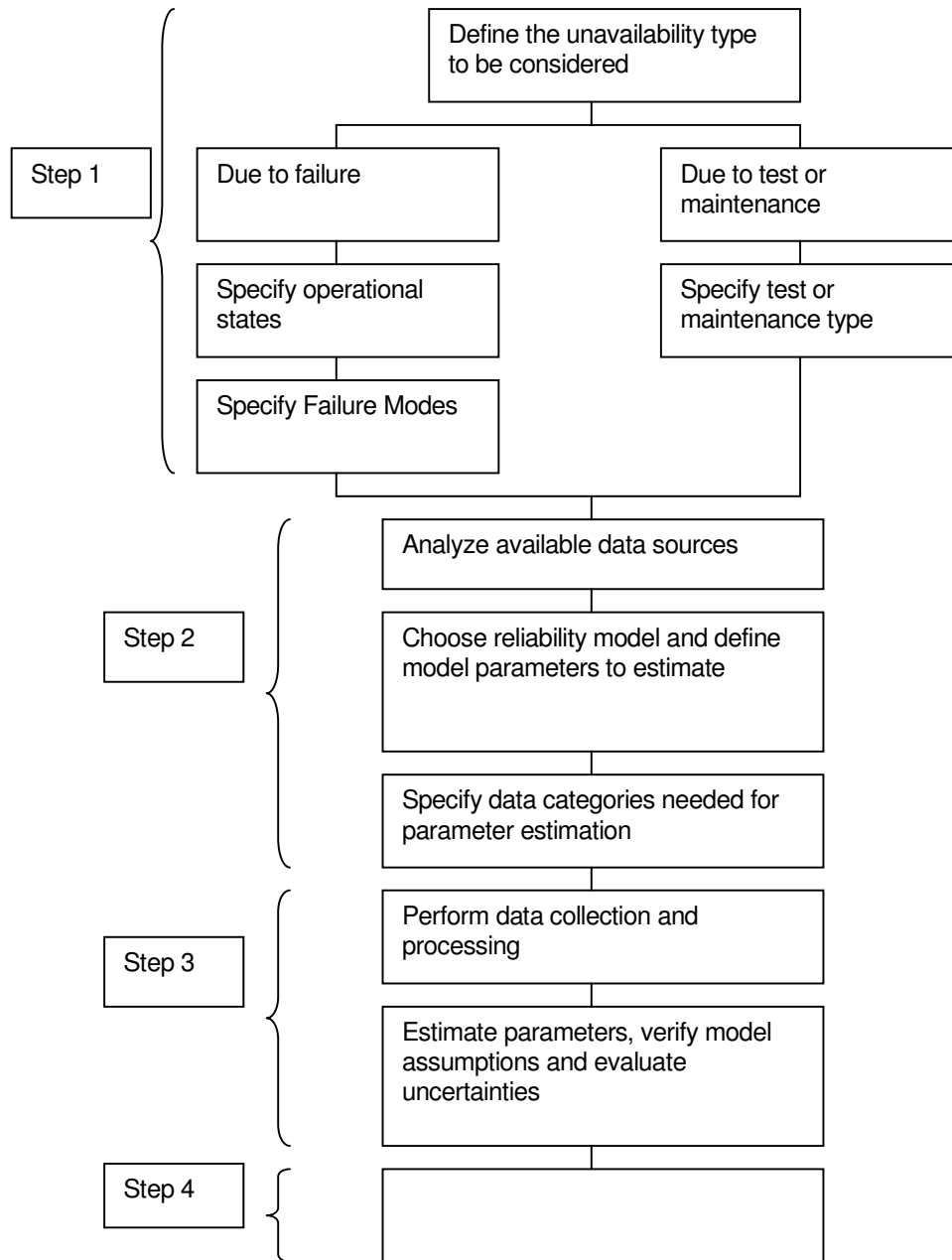
For example, in RiskSpectrum, which is used in many European countries, there is no possibility to specify basic event unavailability as a function of time, and then to perform a risk analysis of the entire model at different time points. One possible solution could be to use the Exchange Event option with pre-calculated unavailability values estimated at different time points assigned for the exchange events.

This chapter presents an approach to defining input data needed for particular Ageing PSA applications.

The general process of evaluating age-dependent input data consists of four steps:

1. Specify unavailability type and attributes associated with a basic event,
2. Choose the model to be applied and data categories needed for parameter estimation,
3. Collect and process data,
4. Estimate time-dependent unavailability factors to be considered in PSA.

Schematically, this process is illustrated in Figure 16. The following paragraphs describe the details of each step.



• Figure 16 Generic process for evaluating age-dependent input data.

3.2.2 Specification of unavailability type and attributes associated with basic events

3.2.2.1 Check-in procedure

As soon as one component is selected for ageing assessment in the PSA and the purpose of the PSA application is specified, the following issues have to be clarified :

- Is this component already modeled in the PSA?
- Where in the PSA model is it taken into account (FT to calculate initiating event frequency or FT for functional event probability)?
- Which type of component unavailability is to be considered :
 - Unavailability due to the failure,
 - Tests and maintenance unavailability,
- Are the relations between dominant ageing mechanisms and failure modes clearly identified?
- On which time-scale is the unavailability of the component to be modeled or extrapolated?
- Have tests and preventive maintenance characteristics been considered?

The answers to these questions will help to understand the scope of needed modifications of the FT models and the PSA as a whole.

Three particular aspects could be mentioned as potentially important at this stage :

- New failure modes and operational states,
- Time-scale (or time points) to be considered,
- Unavailability due to the maintenance.

These aspects are discussed below.

3.2.2.2 Failure modes and operational states

The reasons for selecting a component as susceptible for ageing could be, for example, the existence of an ageing trend in reliability, the appearance of unexpected or previously neglected ageing mechanisms, or the lack of a surveillance program for the component. From the PSA modeling point of view, those reasons could lead to the need to consider ageing for the component:

- Either by existing/modeled failure modes and operating states, in which case the FT models need not be changed (the Ageing PSA task will concern only the input parameters at the level of basic events),
- Or with a new failure mode, new operating states, or new unavailability due to maintenance, In this case new basic events have to be introduced into the FT models.

- Example 6 External leakage of HPSI valves.

The external leakage of HPSI valves is identified as a potentially important failure mode but it was neglected in the PSA reference model. The task consists of screening all HPSI FTs, creation and introduction of new basic events, specification of the attributes and parameters for these basic events.

If the reference PSA model considers the external leakage of the valve, but only in accidental actuation of HPSI, and analysis of the valve surveillance program shows potential latent (undetectable) failure in standby mode, this could be modeled by introducing a new basic event.

Another case of the same failure mode could be if such a leakage is considered in a FT used for calculation of Small LOCA initiating event frequency in a shutdown state. For such a case, the parameters of the basic event have to be modified.

A third possible case for this failure mode could be if external leakage is not a critical failure, but due to the increasing frequency of such events the unplanned unavailability of a HPSI train is increased. This would require introduction of new basic events in the PSA model.

In all these examples, the input data sets are different.

In some cases, due to insufficient data, several failure modes with the same failure mechanisms could be grouped together in order to develop time-dependent models for particular failure modes. For example, failure to close or open of the HPSI valve could be represented as failure to change position.

Table 11 provides an example of operating states / failure modes specification.

• Table 11 Operating states / failure modes specification

Modeling level	Component state	Unavailability type	Failure modes	Time-dependent parameters
FT for IE frequency estimation	Normal operation	Unavailability due to sudden failure (on demand)	Fail to remain in position (Spurious actuation), Fail to start, Fail to open, Fail to close, ...	$Q(t) = f(p(w(t)))$ $p(w, t)$ – probability to fail on demand as a function of number of demands, $w(t)$ – demands intensity
		Unavailability due to degradation failure	Fail to function : short circuit, open circuit, clogging, etc., Fail to run, Internal leak, External leak, ...	$Q(t) = f(\lambda(t))$ $\lambda(t)$ – failure rate/intensity as a function of age
		Unplanned unavailability due to non-critical failures		$Q(t) = f(\phi(t), Tr(t))$ $\phi(t)$ – failure intensity for non-critical failures, $Tr(t)$ – time to repair as a function of age
FT for System / Safety Function unavailability estimation	In standby	Unavailability due to sudden failure (on demand)	Fail to start, Fail to open, Fail to close, ...	$Q(t) = f(p(w(t)))$ $p(w, t)$ – probability to fail on demand as a function of number of demands, $w(t)$ – demands intensity
		Unavailability due to degradation failure	Fail to function : short circuit, open circuit, clogging, etc., Internal leak, External leak, ...	$Q(t) = f(\lambda(t), T_{PT})$ $\lambda(t)$ – failure rate/intensity as a function of age T_{PT} – test period
		Unplanned unavailability due to non-critical failures		$Q(t) = f(\phi(t), Tr(t))$ $\phi(t)$ – failure intensity for non-critical failures, $Tr(t)$ – time to repair as a function of age
	Accidental operation	Unavailability due to degradation failure	Fail to function, Internal leak, Rupture, ...	$Q(t) = f(\lambda(t), T_m)$ $\lambda(t)$ – failure rate/intensity as a function of age T_m mission time
		Unavailability due to sudden failure (on demand)	Fail to start (restart) Fail to open/close (repeated actions),	$Q(t) = f(p(w(t)))$ $p(w, t)$ – probability to fail on demand as a function of number of demands, $w(t)$ – demands intensity

3.2.2.3 Time scale to be considered

Depending on the PSA application, the time scale or time points for risk estimation could be defined differently.

For demonstrating the impact of ageing on the current safety level of the plant, the time-dependent reliability models could be quantified with point estimates of unavailability, corresponding to the current age of the unit. A practical approach could be to compare these point estimates with the reference values, at the component / failure mode level, and the resulting CDF level.

Another possible case could be the predictive risk evaluation at the end of expected lifetime (40 years, for example). As in the previous case, for input data set the time-dependent unavailabilities could be calculated at the time point of 40 years and then the results could be compared with those of the reference PSA.

To see the change in risk profile with the age of the plant over the specified period of time (remaining life, for example), this period could be divided into one-year intervals and corresponding unavailability values could be estimated for the input data set. Then the results could be compared with those of the reference PSA.

3.2.2.4 Unavailability due to the maintenance

Three types of unavailability due to the maintenance usually are considered in PSA :

- unavailability due to corrective maintenance after critical failures (this could be explicitly modeled in the reliability equation by introducing a repair rate parameter, for example),
- unavailability due to corrective maintenance after non-critical failure,
- unavailability due to planned maintenance or tests.

In some cases, the increase in component failure rate may not be observed because of:

- increased number of corrective maintenance activities due to non-critical failures,
- or, increase number of preventive / conditional maintenance interventions.

Both of these factors may imply increased downtime for the component. In this case some modifications to the FT model are needed.

3.2.3 Reliability models and parameters definition

Depending on the type of component and the data available, three types of reliability modeling could be applied :

- Statistical models for failure intensity (rate), these models are based on specific reliability data collected from operating experience. The frequentist approach could be used to estimate the model parameters.
- Statistical models for failure intensity (rate), which are based on combination specific and generic reliability data. The Bayesian approach could be used to estimate the model parameters.

- Physical reliability models for failure probability or failure rate. These models apply the knowledge of degradation mechanisms and operational and environmental stressors to calculate a failure probability.

Physical reliability models are not covered in the Guideline. The Bayesian approach to parameter estimation is presented in Chapter 5. The following analysis deals only with the frequentist approach.

The last column of Table 11 presents in very general terms the types of reliability models and parameters to consider :

- failure rate/intensity as a function of age, $\lambda(t)$,
- probability to fail on demand as a function of number of demands, $p(w, t)$,
- demand intensity, $w(t)$,
- failure intensity for non-critical failures, $\phi(t)$,
- time to repair as a function of age, $Tr(t) = \mu(t)$.

In practice, any of these functions could be represented or approximated by one of the cases of Generalized Linear Model discussed in Chapter 3.1.4. The notion of “time” or “age” could be expressed in different units. In particular, for failure on demand, it could be the cumulative number of demands before failure. This approximation is acceptable in the case of frequently demanded components, if the “age” of the component is much larger than the times between demands.

Taking into account the two possible formats of initial data presentation discussed in chapter 2, the following techniques for parameter estimation are proposed :

- Maximum Likelihood Estimation (MLE) in case of times to failure/censoring data,
- minimization of chi-square statistic for binned data.

The chi-square minimization approach is presented in detail in Chapter 3.1.4.

The MLE approach is presented below for the case of a two-parameter Weibull model with right-censored data.

Assign for the parameters θ_1 and θ_2 of the power-law model (4) presented in paragraph 3.1.4.1 $\theta_1 = \beta/(\eta)^\beta$, $\theta_2 = \beta - 1$. In this case the failure intensity (rate) function can be written as :

$$\lambda(t) = \frac{\beta}{\eta} \left(\frac{t}{\eta} \right)^{\beta-1} \tag{15}$$

For right-censored data $\{t_1, t_2, \dots, t_m, t_{m+1}, \dots, t_n\}$, $i=1, \dots, n$, where m is a number of observed failures and $(n-m)$ is a number of censoring times, the parameters β and η could be defined by solving the system of equations :

$$\frac{1}{\beta} + \frac{1}{m} \sum_{i=1}^m \ln(t_i) - \frac{\sum_{i=1}^n t_i^\beta \ln(t_i)}{\sum_{i=1}^n t_i^\beta} = 0 \quad (16)$$

$$\eta = \left[\frac{\sum_{i=1}^n t_i^\beta}{m} \right]^{1/\beta} \quad (17)$$

Unfortunately, due to the data problems, in many practical cases this solution is not applicable or provides unreliable estimations. As a consequence, the Bayesian procedure for parameter estimation is recommended.

3.2.4 Unavailability as a function of unit age

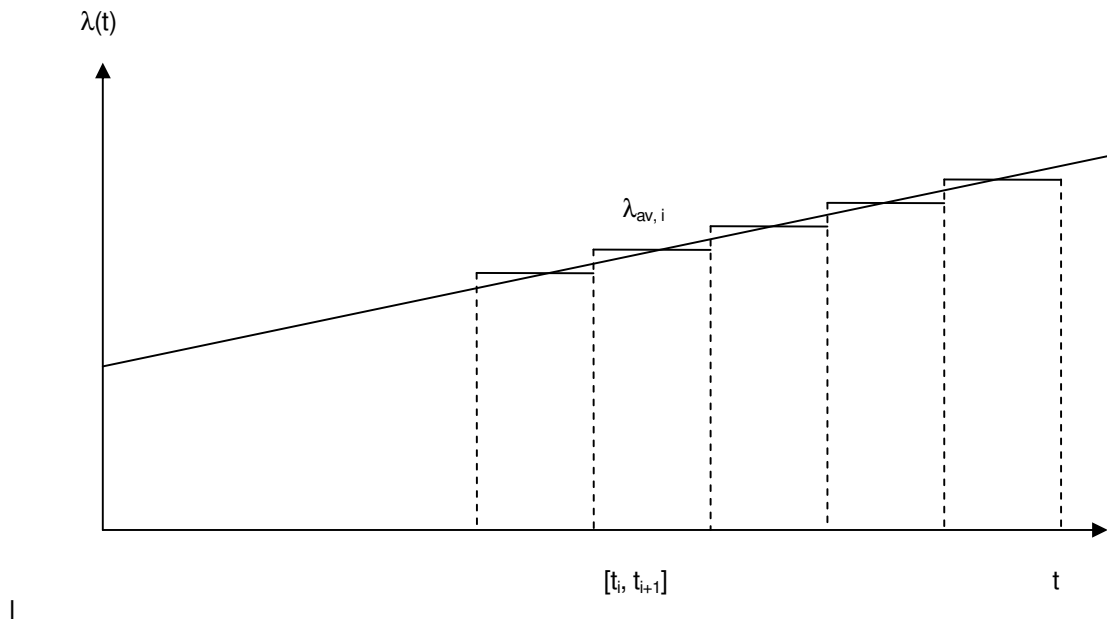
The results of the previous steps permit us to specify the following for each basic event :

- unavailability type,
- component state / operational mode,
- failure mode,
- time-dependent reliability model and parameters.

The next step is to estimate unavailability values to introduce into the PSA model.

Note : most existing PSA codes do not permit time-dependent failure intensity (rate) models and do not perform calculations considering unit age. One possible solution to this problem could be first to prepare a set of pre-calculated unavailability values corresponding to selected age points, and then to introduce these data into the PSA model.

One could use a step-wise approximation for unavailability as a function of time, as shown in Figure 17. The time or age scale is divided into small time intervals and for each interval $[t_i, t_{i+1}]$ the average value of failure intensity (rate), failure probability per demand, or unavailability due to corrective maintenance is assigned to that interval.



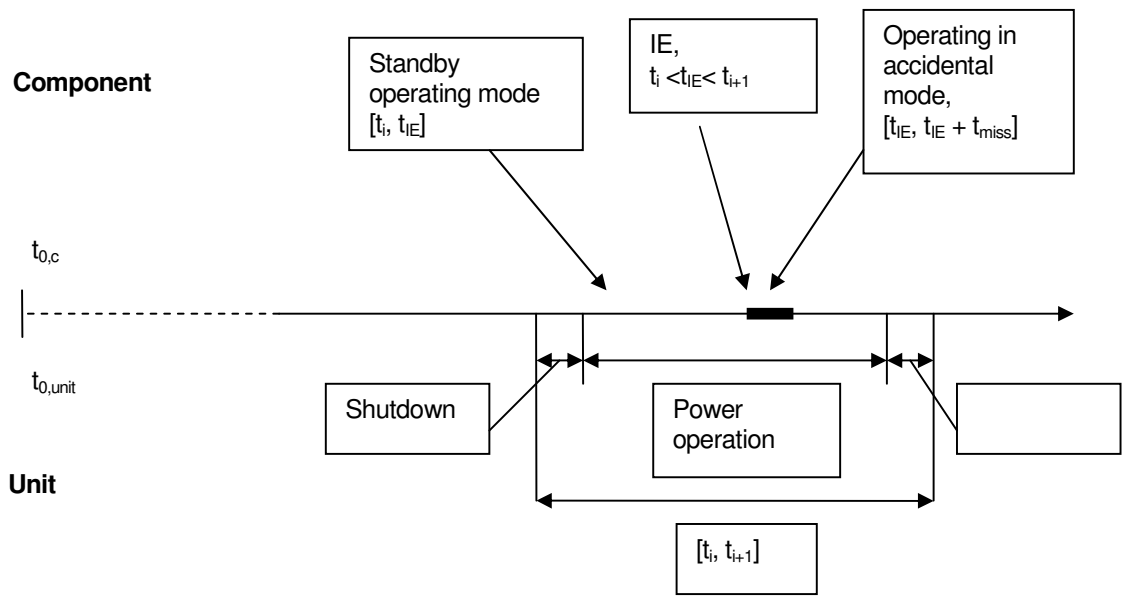
• Figure 17 Step-wise approximation of time-dependent failure intensity function.

The average value is calculated as :

$$\lambda_{av} = \frac{1}{(t_{i+1} - t_i)} \int_{t_i}^{t_{i+1}} \lambda(t) dt \quad (18)$$

This approach permits one to use for basic events the same options and types of reliability models as in the reference PSA. In addition, in standard PSA the Core Damage Frequency is calculated for a one-year interval; for comparison, one should use the same interval in the PSA considering ageing effects.

In calculating the averaged failure intensity values for the interval $[t_i, t_{i+1}]$, it is necessary to consider a correlation of the age of component with the unit age. For example, if a unit commissioning date is assumed as a starting point (age zero) for both the unit and the component, and there were no replacements of components during unit operation, then it could be supposed that if an initiating event occurs in the interval $[t_i, t_{i+1}]$, the component age and unit age will be the same, see Figure 18. In case of component replacement at time t_R , the age of the component at the time when an initiating event occurs, t_{IE} , will be $(t_{IE} - t_R)$.



• Figure 18 Unit and component age correlation.

4.1 Introduction

The use of Bayesian inference methods, as noted by risk researchers, uses “all of the (available) information” and “leads to better parameter estimates and to better decisions.” These aspects are important since organizations are often asked to make inference using sparse data. Consequently, waiting to obtain the “long run” objective frequency before making a decision is in many cases simply not possible. Furthermore, additional information (beyond data) usable in the Bayesian inference framework should not be ignored.

In the context of PSA, where we use probability distributions to represent our state of knowledge regarding parameter values in the models, Bayes’ Theorem gives the **posterior** (or updated) distribution for the parameter (or multiple parameters) of interest, in terms of the prior distribution, failure model, and the observed data, which in the general continuous form is written as:

$$\pi_1(\theta | x) = \frac{f(x | \theta)\pi(\theta)}{\int f(x | \theta)\pi(\theta)d\theta}. \quad (19)$$

In this equation, $\pi_1(\theta|x)$ is the posterior distribution for the parameter of interest, denoted as θ (note that θ can be vector-valued). The posterior distribution is the basis for all inferential statements about θ , and will also form the basis for model validation approaches to be discussed later. The observed data enters via the likelihood function, $f(x|\theta)$, and $\pi(\theta)$ is the prior distribution of θ .

The denominator of Bayes’ Theorem is sometimes denoted $f(x)$, and is called the marginal or unconditional distribution of x . Note that it is a weighted average distribution, with the prior distribution for θ acting as the weighting function. In cases where X is a discrete random variable (e.g., number of events in some period of time), $f(x)$ is the probability of seeing x events, unconditional upon a value of θ . In this context, which will become useful for model validation, $f(x)$ will be referred to as the **prior predictive** distribution for X .

The likelihood function, or just likelihood, is also known by another name in PSA applications – it is the aleatory model describing an observed physical process. For example, a battery failure may be modeled in an electric power system fault tree as a Poisson process. Consequently, there is a fundamental modeling tie from the PSA to the data collection and evaluation process.

The **likelihood function**, $f(x|\theta)$, is most often binomial, Poisson, or exponential in traditional PSA applications.² This function represents the observed failure generation mechanism (the aleatory failure model) and is also a function of the data collection process. Note that the symbol “|” represents a conditionality, which in the case of the likelihood function is described as “the probability we see the observed data **given** the parameter θ takes on a certain value.”³

Priors can be classified broadly as either informative or noninformative. Informative priors, as the name suggests, contain substantive information about the possible values of the unknown parameter, θ . Noninformative priors, on the other hand, are intended to be in some sense objective, letting the data speak

² When dealing with repair or recovery times, the likelihood function may be lognormal, Weibull, or gamma. We will describe these, and other, examples in Chapter 5.

³ When the observed random variable is discrete, the likelihood function is a probability. For continuous random variables, the likelihood is a density function, which is proportional to a probability.

for themselves; thus, they contain little substantive information about the parameter of interest. Other terms for noninformative priors are diffuse priors, vague priors, flat priors, formal priors, and reference priors. They are often mathematically derived, and there are numerous types in use.

Bayesian Parameter Estimation and Model Validation

5.1 Preface to Chapter 5

This chapter is not written in the typical style of other data analysis guides and textbooks, in the sense that it contains few equations and references to the technical literature. Instead WinBUGS scripts are used in place of equations in the hope that they will better convey essential concepts to a more practically oriented audience. Furthermore, these scripts can be adapted to problems at hand by analysts at the plants. For an overview of WinBUGS and associated scripts, see Appendix B. Details of the underlying mathematics can be found in Appendix C.

In this document, WinBUGS scripts are indicated by way of formatted text inside a shaded box:

```

model {
x ~ dbin(p, n)           #Binomial dist for number of failures in n demands
...
}

```

Within these scripts, the following notation is used:

- “~” indicates that the variable to the left of “~” follows the *distribution* to the right of “~.” Examples of distributions include binomial (dbin), beta (dbeta), gamma (dgamma), Poisson (dpois), normal (dnorm), and lognormal (dlnorm). A list of distributions available in WinBUGS can be found by selecting Help/Distributions within WinBUGS.
- “#” indicates that the text to the right of “#” is a comment.
- “<-” indicates that the variable to the left of “<-” is assigned the expression to the right of “<-.”

A directed acyclic graph (DAG) is a common way of displaying a Bayesian inference problem and is the underlying model used by WinBUGS (in script form). In a DAG, stochastic variables are displayed as ovals. Variables that influence x , referred to as the parents of x , are shown at a higher level in the DAG, and are connected by arrows to the variables they influence. Constant parameters (such as n in the binomial distribution) are shown in the DAG as diamonds. We will display the DAGs associated with many of the WinBUGS scripts in this document; however, it is not necessary to develop the DAG, as WinBUGS uses the script representation for its analysis.

For relatively simple problems, a DAG can be an aid in understanding the problem, particularly for an analyst who is new to WinBUGS. However, as the complexity of the problem increases, most analysts will find that the script representation of the problem is clearer. We will use the following conventions for DAGs in this document. Note that all variables, which are referred to as *nodes* by WinBUGS, can be scalars, vectors, matrices, or arrays.

- Ovals represent stochastic variables whose uncertainty (either aleatory or epistemic) is represented by a probability distribution.

- Diamonds represent constant parameters (no uncertainty).
- Rectangles represent calculated parameters. As such, their probability distribution is not specified by the analyst but is calculated by WinBUGS from an equation within the script.⁴
- Dashed lines are sometimes used for clarification when certain parameters are entered or calculated in the script as part of other nodes.
 - In cases where the node is used as inference, the arrow will be connected to the dashed symbol.
 - In cases where the parameter within the node is used as inference, the arrow will be connected to the symbol within the dashed node.

5.2 Inference for a Single Parameter

We begin with a short review of the most commonly encountered situations in PSA, which meet the first or second and third of the following three conditions:

- The probabilistic model for failure (corresponding to the likelihood function in Bayes' Theorem) is a Bernoulli process.
- The probabilistic model for failure is a (homogeneous) Poisson process. In this second case, the analysis depends on the form of the data:
 - a. Number of failures and exposure time, leading to Poisson distribution
 - b. Individual failure times, leading to exponential distribution.
- The observed data are complete, homogeneous, and are known with certainty.

More complicated inference, where one or more of these conditions are not satisfied, is described in later sections of this chapter. Some of the terms in these conditions are defined below.

Homogeneous	A set of data made up of similar constituents. A homogeneous population is one in which each item is of the same type.
Complete	All failures are observed, no data are missing
Censoring	Usually applied to time-to-failure data. Can be of many different kinds, but generally the term means that more failures could have occurred than were actually observed.

Type I: the observation period ended at a pre-specified time, and some components under observation had not yet failed. The number of components that fail during the test is therefore random.

⁴ Note that within WinBUGS, rectangles always represent constant (nonstochastic) nodes. We have deviated from this convention to more clearly indicate which variables, stochastic or constant, are being calculated inside a WinBUGS script.

Type II: the observation period ended after a specified number of components had failed, so that some components were still operating at the end of the test. The length of the test is therefore random.

Left censoring: failures occurring before the censoring time are missing (see below).

Right censoring: failures occurring after the censoring time are missing (see below).

Interval censoring: failure times are only recorded inside the censoring interval.

Missing	Data were not recorded. For example, a failure was known to occur in a given time interval, but the actual time of failure was not recorded.
Uncertain data	The failure counts, number of demands, or exposure time are not known with complete accuracy.

5.2.1 Binomial Distribution for Failures on Demand

This model is often used when a component must change state in response to a demand. For example, a relief valve may need to open to relieve pressure upon receipt of a signal from a controller that an over-pressure condition exists. The following assumptions underlie the binomial distribution:

- There are two possible outcomes of each demand, typically denoted success and failure.
- There is a constant probability of failure (this is typically success in reliability engineering) on each demand, denoted herein as p .
- The outcomes of earlier demands do not influence the outcomes of later demands.

The unknown parameter in this model is p , and the observed data are the number of failures, denoted x , in a specified number of demands, denoted n .

5.2.1.1 Estimating p

Frequentist estimation of p typically proceeds by maximizing the likelihood function, leading to the following estimate:

$$\hat{p} = \frac{x}{n} \quad (20)$$

Confidence intervals, which are conservative because of the discrete nature of x, can be obtained using the BETAINV() function in Excel or other spreadsheet packages. For the lower percentile, the parameters are x and n - x + 1. For the upper percentile, they are x + 1 and n - x.

Bayesian inference requires a prior distribution. The simplest type of prior distribution from the standpoint of the mathematics of Bayesian inference is the so-called **conjugate** prior, in which the prior and posterior distribution are of the same functional type (e.g., beta, gamma), and the integration in Bayes' Theorem is circumvented. For the binomial distribution, the conjugate prior is a beta distribution. Two parameters are needed to describe the beta prior distribution completely, and these are denoted α_{prior} and β_{prior} . Conceptually, α_{prior} can be thought of as the number of failures contained in the prior distribution, and the sum of α_{prior} and β_{prior} is like the number of demands over which these failures occurred. Thus, small values of α_{prior} correspond to less information, and this translates into a broader, more diffuse prior distribution.

With the observed data consisting of x failures in n demands, the conjugate nature of the prior distribution and likelihood function allows the posterior distribution to be written down immediately using simple arithmetic: the posterior distribution is also a beta distribution, with new (adjusted) parameters given by

$$\begin{aligned} \alpha_{\text{post}} &= \alpha_{\text{prior}} + x \\ \beta_{\text{post}} &= \beta_{\text{prior}} + n - x \end{aligned} \quad (21)$$

From the properties of the beta distribution, the prior and posterior mean of p are given by $\alpha_{\text{prior}}/(\alpha_{\text{prior}} + \beta_{\text{prior}})$ and $\alpha_{\text{post}}/(\alpha_{\text{post}} + \beta_{\text{post}})$, respectively. Credible intervals (the Bayesian analog of confidence intervals) for either distribution can be found using the BETAINV() function, using either the prior or posterior parameters as inputs, as appropriate.

It is also possible to use a prior distribution that has little influence on the posterior. Such a prior is referred to variously as a *noninformative prior*, *diffuse prior*, *reference prior*, *formal prior*, *vague prior*, and *objective prior*. For this document, we will stick with the term noninformative prior.

As the name suggests, a noninformative prior distribution contains little information about the parameter of interest, which in this case is p. Such priors originated in a (continuing) quest to find a mathematical representation of complete uncertainty. There are two situations in which a noninformative prior may be useful. The first is where the observed data are abundant enough to dominate the information contained in any reasonable prior, so it does not make sense to expend resources developing an informative prior distribution. The second is where the analyst wishes to use a prior that has little influence on the posterior, perhaps as a point of reference.

With lots of observed data, the prior will have little influence on the posterior. So why not just use the data alone? Remember that a probability distribution is required to propagate uncertainty, so Bayes' Theorem is still used to obtain a posterior distribution.

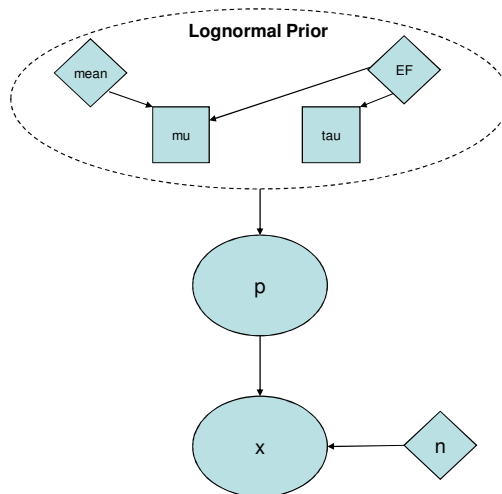
Unfortunately, there are many routes that lead to slightly different noninformative priors, and the intuitive choice of a uniform prior is not what is usually used in PSA. The most common noninformative prior for single-parameter inference in PSA is the Jeffreys prior. Its functional form is dependent upon the likelihood function, so there is not a single "Jeffreys prior;" rather, there is a different Jeffreys prior for each likelihood function. For the case here, where the likelihood function is the binomial distribution, the Jeffreys prior is a beta distribution with both parameters equal to 0.5. Thus, inference with the Jeffreys prior is a special case of inference with a beta conjugate prior. All of the results described above apply in the case of the Jeffreys prior, but with α_{prior} and β_{prior} both equal to 0.5. This leads to a posterior mean of $(x + 0.5)/(n + 1)$.

A nonconjugate prior is one in which the prior and posterior distribution are not of the same functional form. In such cases, numerical integration is required for the denominator of Bayes' Theorem. In the past, this has been a limitation of Bayesian inference, and is one reason for the popularity of conjugate priors. However, cases often arise in which a nonconjugate prior is desirable, despite the increased mathematical difficulty. As an example, generic databases often express epistemic uncertainty in terms of a lognormal distribution, which is not conjugate with the binomial likelihood function.

Although spreadsheets can be used to carry out the required numerical integration for the case of a single unknown parameter, the easiest way to deal with nonconjugate priors today is with WinBUGS. We illustrate the case of a lognormal prior with the following example.

- Example 7 Relief valve fails to open (binomial model) and lognormal prior.

Assume that we are interested in the probability of a relief valve failing to open on demand, and we have a generic database that provides a lognormal prior. Assume the database lists the mean failure probability as 10^{-3} with an error factor of 10. Assume that our observed data are 2 failures in 285 demands. Figure 19 shows the DAG for this problem. The WinBUGS script shown below is used to analyze this example.



- Figure 19 DAG for binomial inference with lognormal prior

```

model {
x ~ dbin(p, n)           #Binomial distribution for number of failures
p ~ dlnorm(mu, tau)     #Lognormal prior distribution for p
tau <- 1/pow(log(prior.EF)/1.645, 2)  #Calculate tau from lognormal error factor
#Calculate mu from lognormal prior mean and error factor
mu <- log(prior.mean) - pow(log(prior.EF)/1.645, 2)/2
}
data
list(x=2,n=285, prior.mean=0.001, prior.EF=10)

```

- Script 1. WinBUGS script for Bayesian inference with binomial likelihood function and lognormal prior.

Solution

Running this script for 100,000 iterations, discarding the first 1,000 iterations to allow for convergence to the posterior distribution, gives a posterior mean for p of 3.54×10^{-3} and a 90% credible interval of $(5.57 \times 10^{-4}, 9.23 \times 10^{-3})$.

When working with a lognormal prior distribution from a generic database, be sure to know whether the distribution is expressed in terms of the prior mean or median value.

Generic databases may not always describe the lognormal distribution in terms of a mean value and an error factor; quite often the median (50th percentile) replaces the mean value. This may also be the case when eliciting information from experts; the expert may be more comfortable providing a median value. In this case, the analysis changes only slightly. In Script 1, the line that calculates the mean from the lognormal prior mean and error factor is replaced by the following line:

```
mu <- log(prior.median)
```

and prior.median is loaded in the data statement instead of prior.mean.

Beware of lognormal priors for p when p is expected to be near one.

Cases may arise where the value of p could be approaching unity. In such cases, using a lognormal prior is problematic because it allows for values of p greater than unity, which is not meaningful since p is a probability (either failure probability or reliability, depending on the context). In such cases, a logistic-normal prior is a “lognormal-like” distribution, but one that constrains the values of p to lie between zero and one. The WinBUGS script below uses the lognormal mean and error factor (e.g., from a generic database), but “constrains” the distribution to lie between zero and one by replacing the lognormal distribution with a logistic-normal distribution.


```

model {
x ~ dbin(p, n)                #Binomial distribution for number of failures
p <- exp(p.constr)/(1 + exp(p.constr)) #Logistic-normal prior distribution for p
p.constr ~ dnorm(mu, tau)
tau <- 1/pow(log(prior.EF)/1.645, 2) #Calculate tau from lognormal error factor
#Calculate mu from lognormal prior mean and error factor
mu <- log(prior.mean) - pow(log(prior.EF)/1.645, 2)/2
}
data
list(x=2,n=285, prior.mean=0.001, prior.EF=10)

```

- Script 2 WinBUGS script for Bayesian inference with binomial likelihood function and logistic-normal prior

5.2.2 Poisson Distribution for Initiating Events or Failures in Time

This model is often used for failures of normally operating components, failures of standby components that occur at some point in time prior to a demand for the component to change state, and for initiating events. As an example, we might use a Poisson distribution as our stochastic model for failure of an operating cooling pump. The following assumptions underlie the Poisson distribution:

- The probability of an event (e.g., a failure) in a small time interval is approximately proportional to the length of the interval. The constant of proportionality is denoted by λ .
- The probability of simultaneous events in a short interval of time is approximately zero.
- The occurrence of an event in one time interval does not affect the probability of occurrence in another, non-overlapping time interval.

The unknown parameter in this model is λ , and the observed data are the number of events, denoted x , in a specified time period, denoted t . Both x and t are assumed to be known with certainty in this section.

Note that λ is a rate and has units of inverse time. Also note that lambda is not a function of time, so the simple Poisson distribution cannot be used for reliability growth or aging.

5.2.2.1 Estimating λ

Frequentist estimation, as for p above, usually proceeds by maximizing the likelihood function leading to the following point estimate of λ :

$$\hat{\lambda} = \frac{x}{t} \quad (22)$$

Confidence intervals, which are again conservative because the observed events are discrete counts, can be obtained from percentages of the chi-square distribution, using the following:

$$\lambda_L = \frac{\chi_{\frac{x}{2}}^2(2x)}{2t} \tag{23}$$

$$\lambda_U = \frac{\chi_{1-\frac{x}{2}}^2[2(x+1)]}{2t}$$

As was the case with the binomial distribution, a conjugate prior is sometimes chosen for purposes of mathematical convenience with respect to Bayesian inference. For the Poisson distribution, the conjugate prior is a gamma distribution. Two parameters are needed to describe the gamma prior distribution completely, and these are denoted α_{prior} and β_{prior} . Conceptually, α_{prior} can be thought of as the number of events contained in the prior distribution, and β_{prior} is like the period of time over which these events occurred. Thus, small values of α_{prior} correspond to less information, and this translates into a broader, more diffuse prior distribution.

With the observed data consisting of x failures in time t , the conjugate nature of the prior distribution and likelihood function allows the posterior distribution to be written down immediately using simple arithmetic: the posterior distribution is also a gamma distribution, with new (adjusted) parameters given by

Note that β_{prior} and β_{post} have units of time, and the units have to be the same as for t .

$$\alpha_{post} = \alpha_{prior} + x$$

$$\beta_{post} = \beta_{prior} + t \tag{24}$$

Caution: Be sure to know how your spreadsheet software parameterizes the gamma distribution. Most packages use the reciprocal of β as the parameter.

From the properties of the gamma distribution the prior and posterior mean of lambda are given by $\alpha_{prior}/\beta_{prior}$ and $\alpha_{post}/\beta_{post}$, respectively. Credible intervals for either distribution can be found using the GAMMAINV() function built into modern spreadsheets.

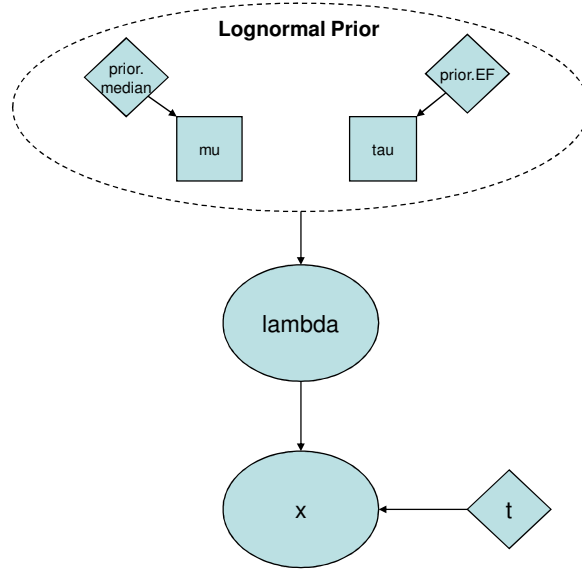
WinBUGS can only accept proper distributions, so enter the Jeffreys prior for the Poisson likelihood as `dgamma(0.5, 0.0001)`

As was the case for the binomial distribution, there are many routes to a noninformative prior for λ , with the most commonly used one in PSA being the Jeffreys prior. In the case of the Poisson likelihood the Jeffreys noninformative prior is like a gamma distribution with $\alpha_{prior} = 0.5$ and $\beta_{prior} = 0$. This is not a proper distribution, as the integral over all possible values of λ is not finite. However, it always yields a proper posterior distribution, with parameters $\alpha_{post} = x + 0.5$ and $\beta_{post} = t$. Thus, the posterior mean of λ is given by $(x + 0.5)/t$. Inference with the Jeffreys prior can be thought of as a special case of inference with a gamma conjugate prior, with the results above applying.

As was the case for the parameter p in the binomial distribution, a lognormal distribution is a commonly encountered nonconjugate prior for λ in the Poisson distribution. The analysis can be carried out with WinBUGS, exactly as was done for p in the binomial distribution. Here, however, there is no concern about values of λ greater than one, because λ is a rate instead of a probability, and can take on any positive value, in principle.

- Example 8 Cooling pump fails to operate (Poisson model) and lognormal prior.

Assume that the prior distribution for the failure rate of a cooling pump is lognormal with a median of 5×10^{-7} /hour and an error factor of 14. Assume the observed data are no failures in 200 days. Figure 20 shows the DAG underlying this problem. The WinBUGS script below can be used to find the posterior mean and 90% interval for lambda.



• Figure 20 DAG for Poisson inference with lognormal prior

```

model {
x ~ dpois(mean.poisson)      #Poisson distribution for number of events
mean.poisson <- lambda*time.hr #Poisson parameter
time.hr <- time*24          #Convert days to hours
lambda ~ dlnorm(mu, tau)    #Lognormal prior distribution for lambda
#Calculate tau from lognormal error factor
tau <- 1/pow( log(prior.EF)/1.645, 2)
#Calculate mu from lognormal median and error factor
mu <- log(prior.median)
}
data
list(x=0, time=200, prior.median=5.E-7, prior.EF=14)
  
```

• Script 3 WinBUGS script for Bayesian inference with Poisson likelihood function and lognormal prior

Running this script for 100,000 iterations, discarding the first 1,000 iterations to allow for convergence to the posterior distribution, gives a posterior mean for p of 1.62×10^{-6} /hour and a 90% credible interval of $(3.50 \times 10^{-8}$ /hour, 6.45×10^{-6} /hour).

5.2.3 Exponential Distribution for Random Durations

The analysis in this section is employed when the times at which random events occur are observed, instead of the number of such events in a specified period of time. If the assumptions for the Poisson distribution as described in Sec. 5.2.2 are met, then the times between events are exponentially distributed with unknown parameter λ ; this is the same λ that appears as the unknown parameter in the Poisson distribution. Because the observed data consist of n failure times (when the data are complete), with n specified, the form of the likelihood function changes from a Poisson distribution to a product of n exponential distributions. However, much of the analysis is very similar to the analysis done for the Poisson distribution in Sec. 5.2.2. In this section we treat only the case in which all failure times are observed and known

If the times at which Poisson-distributed events occur are observed, then the likelihood function is now based on the exponential distribution.

with certainty. In later sections we will cover cases in which not all components fail, and in which observed failure times are uncertain.

5.2.3.1 Estimating λ

The frequentist maximum likelihood estimate of λ is given by

$$\hat{\lambda} = \frac{n}{\sum_{i=1}^n t_i} \tag{25}$$

Confidence intervals, which are now exact because time is a continuous random variable, are obtained from percentages of the chi-square distribution using the following:

$$\lambda_L = \frac{\chi_{\alpha/2}^2(2n)}{2\sum t_i}$$

$$\lambda_U = \frac{\chi_{1-\alpha/2}^2(2n)}{2\sum t_i} \tag{26}$$

As was the case for the Poisson distribution in Sec. 5.2.2, the conjugate prior for Bayesian inference with the exponential likelihood is again a gamma distribution, with parameters denoted α_{prior} and β_{prior} . Once again, β_{prior} has units of time, and these units must match the units of the observed times that constitute the data. The posterior distribution will again be a gamma distribution with parameters $\alpha_{\text{post}} = \alpha_{\text{prior}} + n$ (the number of observed times), and $\beta_{\text{post}} = \beta_{\text{prior}} + t_{\text{total}}$, where t_{total} is the sum of the observed times. From the properties of the gamma distribution, the prior and posterior mean of λ are given by $\alpha_{\text{prior}} / \beta_{\text{prior}}$ and $\alpha_{\text{post}} / \beta_{\text{post}}$, respectively. Credible intervals for either distribution can be found using the GAMMAINV() function built into modern spreadsheets.

The Jeffreys noninformative prior for the exponential likelihood is like a gamma distribution with both parameters equal to zero. This might seem odd, given the relationship between the exponential and Poisson distributions mentioned above. In fact, it is odd that the Jeffreys prior changes, depending on whether one counts failures or observes actual failure times. However, it is not important to delve into the reasons for this difference and its philosophical implications here. Again, the Jeffreys prior is an improper distribution, but it always results in a proper posterior distribution. The parameters of the posterior distribution will be n and t_{tot} , resulting in a posterior mean of n/x_{tot} . This is

WinBUGS can only accept proper distributions, so enter the Jeffreys prior for exponential data as dgamma(0.0001, 0.0001). An initial value for lambda will have to be provided in this case.

numerically equal to the frequentist MLE, and credible intervals will be numerically equal to confidence intervals from a frequentist analysis of the data.

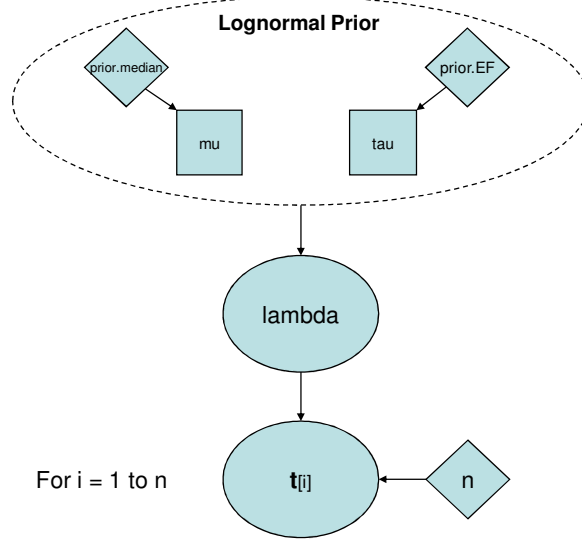
Again, the lognormal distribution is a commonly encountered nonconjugate prior for a failure rate. The only thing that changes from the earlier discussion in Sec. 5.2.2 is the likelihood function, which is now a product of exponential distributions. We again use WinBUGS to carry out the analysis.

- Example 9 Circulating pump fails to operate (exponential model) and lognormal prior.

The following failure times (in minutes) were recorded for a component: 55707, 255092, 56776, 111646, 11358772, 875209, 68978. Using the lognormal prior from Example 8, find the posterior mean and 90% credible interval for the failure rate, λ .

Solution

Figure 21 shows the underlying DAG model. The WinBUGS script for this example is shown below.



• Figure 21 DAG for exponential inference with lognormal prior

```

model {
for(i in 1:n) {
    time[i] ~ dexp(lambda) #Exponential likelihood function for n failure times
  }
  lambda ~ dlnorm(mu, tau) #Lognormal prior for lambda
  #Calculate tau from lognormal error factor
  tau <- 1/pow( log(prior.EF)/1.645, 2)
  #Calculate mu from lognormal median and error factor
  mu <- log(prior.median)
}
data
list(time=c(55707, 255092, 56776, 111646, 11358772, 875209, 68978), n=7, prior.median=5.E-7,
prior.EF=14)
  
```

• Script 4 WinBUGS script for Bayesian inference with exponential likelihood function and lognormal prior

Using 100,000 iterations, with 1,000 burn-in iterations discarded to allow for convergence to the posterior distribution, the posterior mean is found to be 5.46×10^{-7} /hour, with a 90% credible interval of $(2.64 \times 10^{-7}$ /hour, 9.16×10^{-7} /hour).

5.3 Checking Model Assumptions

A common assumption underlying the simple stochastic models commonly used in PSA is that the model parameter (p or λ) does not change with time. Note that this assumption is never strictly satisfied, but in many cases the deviations are not significant. However, if aging is present, then the violation of this assumption may be of practical significance from a PSA perspective.

5.3.1 Binomial distribution

Recall that the binomial distribution, used as a stochastic model for failures on demand, is derived from three underlying assumptions. We restate these assumptions here for convenience, but our focus in this section will be on the second assumption regarding the constant nature of p .

- There are two possible outcomes of each demand typically denoted success and failure.
- There is a constant probability of failure (typically in PSA; typically success in reliability engineering) on each demand, denoted herein as p .
- The outcomes of earlier demands do not influence the outcomes of later demands.

- Example 10 Check for time trend in relief valve leakage probability in cooling water system

Consider the case where we are concerned about leakage from the relief valve on the surge tank in a closed-cooling system. Assume that a check is performed once a week on this valve and a notation is made of whether or not the valve is leaking. The analyst decides to use a binomial distribution for the number of times the valve is leaking in a year, so that n is 52. The unknown parameter of interest is p , the probability that the valve is leaking in any given week. Assume that over time, the data below has been collected for this valve.

<i>Year</i>	<i>Failures</i>	<i>Demands</i>
1	4	52
2	2	52
3	3	52
4	1	52
5	4	52
6	3	52
7	4	52
8	9	52
9	6	52

The analyst decides to use a Jeffreys prior with the data pooled across the nine years of leak-testing (36

failures in 486 demands). This assumption of constant p across time results in a posterior mean for p of 0.08 and a 90% credible interval of (0.06, 0.099). Investigate the validity of the analyst's approach.

Solution

One of the assumptions underlying the binomial distribution is that p does not change from one demand to the next. In our example, where data has been collected over a period of time, this also implies that p should not change over time. In pooling the data, the analyst is tacitly accepting the validity of this assumption. A qualitative way to check this assumption is to calculate a confidence interval or credible interval for p in each of the nine years, plot them side by side, and look for overlap. If the intervals all overlap with one another, this is evidence that p is not varying significantly over time, so pooling the data may be reasonable. Note that p could exhibit a time trend, perhaps increasing over time due to wearout of the valve, or it could vary significantly from one year to the next. Either situation could make the simple binomial likelihood a poor model.

The WinBUGS script used to construct a side-by-side plot of the 95% credible intervals for each year (using a Jeffreys noninformative prior) is shown below.

```

model {
  for (i in 1:N) {
    x[i] ~ dbin(p[i], n[i])    #Binomial distribution for failures in each year
    p[i] ~ dbeta(0.5, 0.5)    #Jeffreys prior for p, used in developing interval plot of p
  }
}

Data
x[]    n[]
4      52
2      52
3      52
1      52
4      52
3      52
4      52
9      52
6      52

END
list(N=9)

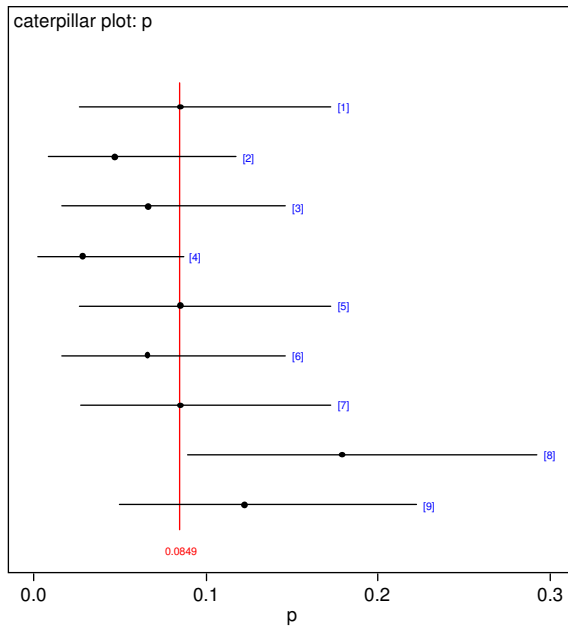
```

- Script 5 WinBUGS script to generate caterpillar plots for p with multiple data sources (block data entry)

Note the use of the block format for entering data. Also, there are now two steps to loading data.

This script was run for 100,000 iterations, discarding the first 1,000 iterations to allow for convergence. The figure below shows the 95% credible intervals for p in each year. The plot suggests that p might be increasing with time, but there is a lot of uncertainty, so it is

difficult to judge whether a trend may actually be present. A more quantitative measure is needed.



• Figure 22 Side-by-side plot of 95% credible intervals for p in each of the 9 years. Dots indicate posterior mean, and red line is the average of the posterior means.

The WinBUGS script below is used to generate a quantitative measure of how well a model that pools the data can replicate the observed data in each year. This test is similar in concept to the frequentist chi-square test, but does not require further binning of the data and does not rely upon asymptotic approximations. A Jeffreys prior is used for the single parameter (p .constant) that is estimated. Running this script in the usual way gives a mean for the p .value node of 0.18. If the model were good at replicating the observed data, we would expect the mean of this node to be near 0.5. The p -value from the frequentist chi-square test is 0.08. Monitoring the x .rep node shows that the constant- p model tends to predict too many failures in early years and too few in later years, suggesting that a trend model might be better at replicating the observed data. We will examine this model in Sec. 5.4.1 below.

```

model {
  for (i in 1:N) {
    x[i] ~ dbin(p[i], n[i])    #Binomial distribution for failures in each year
    x.rep[i] ~ dbin(p[i], n[i]) #Replicate value from posterior predictive distribution
    diff.rep[i] <- pow(x.rep[i] - n[i]*p[i], 2)/(n[i]*p[i]*(1-p[i]))
    diff.obs[i] <- pow(x[i] - n[i]*p[i], 2)/(n[i]*p[i]*(1-p[i]))
    p[i] <- p.constant    #Constant parameter in each year, resulting in pooling of data
  }

  chisq.rep <- sum(diff.rep[])
  chisq.obs <- sum(diff.obs[])
  p.value <- step(chisq.rep - chisq.obs)
  p.constant ~ dbeta(0.5, 0.5) #Jeffreys prior for p.constant
}

Data
x[]      n[]
4        52
2        52
3        52
1        52
4        52
3        52
4        52
9        52
6        52

END
list(N=9)

```

• Script 6. WinBUGS script for testing assumption that p in binomial distribution does not vary over time.

5.3.2 Poisson Distribution

Recall the three underlying assumptions for the Poisson distribution:

- The probability of an event (e.g., a failure) in a small time interval is approximately proportional to the length of the interval. The constant of proportionality is denoted by λ .
- The probability of simultaneous events in a short interval of time is approximately zero.
- The occurrence of an event in one time interval does not affect the probability of occurrence in another, non-overlapping time interval.

Again, we will focus on the constant nature of λ over time.

- Example 11 Check for constant component failure rate

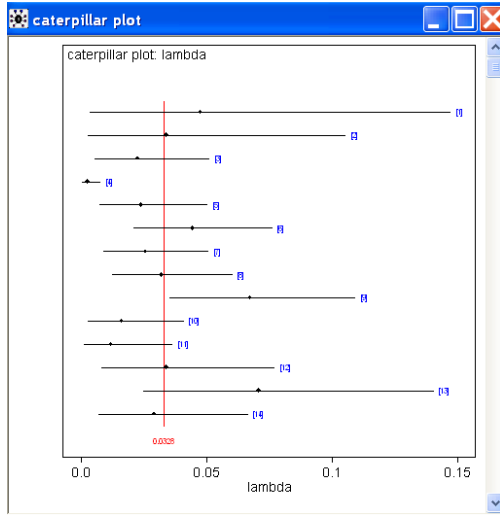
Consider the following virtual data for a particular component in a French reactor. Investigate whether λ for this component appears to be constant over time.

T	X	Exp. Time
0.5	1	31.64
1.5	1	44.135
3	3	156.407
6	1	608.38
8.5	4	189.545
9.5	9	214.63
10.5	5	216.355
11.5	6	204.091
12.5	12	185.926
13.5	2	157.536
14.5	1	127.608
15.5	3	104.105
16.5	5	77.86
20	3	121.167

Solution

One of the assumptions underlying the Poisson distribution is that λ does not change from one interval of time to the next. In this example, where data has been collected over a period of time, this also implies that λ should not change over time. A qualitative way to check this assumption is to calculate a 95% confidence or credible interval for each of the 14 years, plot them side by side, and look for overlap. If the intervals all overlap with one another, this is evidence that λ is not varying significantly over time. Note that λ could exhibit a time trend, perhaps increasing over time due to aging, or it could vary significantly from one year to the next. Either situation could make the simple Poisson likelihood a poor model.

Script 7 below can be used to construct a side-by-side plot of the credible intervals for each year (using a Jeffreys noninformative prior). The side-by-side interval plot is shown below.



• Figure 23 Side-by-side interval plot illustrating varying λ over time

The plot does not suggest a trend over time, but does suggest significant bin-to-bin variability.

Because WinBUGS can have difficulty generating initial values from the Jeffreys prior for the exponential distribution, we give it a starting value.

Script 7 also provides a quantitative measure of how well a model that pools the data can replicate the observed data in each year. A Jeffreys prior is used for the single parameter (λ .constant) that is estimated. Running this script in the usual way gives a mean for the p.value node of 0.004, indicating a poorly fitting model, in the sense that a model with constant λ is highly unlikely to be able to reproduce the variability in the observed data. Note that the p-value (from Fisher's asymptotic chi-square test) was 0.002. The MLE for the pooled data is 0.0197, with an (at least) 90% confidence interval of (0.015, 0.025).

```

model {
for(i in 1:N) {
  x[i] ~ dpois(mu[i])           #Poisson distribution for number of failures in each source
  x.rep[i] ~ dpois(mu[i])      #Replicate value from posterior predictive distribution
  mu[i] <- lambda[i]*t[i]      #Parameter of Poisson distribution
#lambda[i] ~ dgamma(0.5, 0.0001) #Jeffreys prior for caterpillar plot of intervals in each year
  lambda[i] <- lambda.constant  # Use this line to test for poolability
  diff.obs[i] <- pow(x[i] - mu[i], 2)/mu[i]
  diff.rep[i] <- pow(x.rep[i] - mu[i], 2)/mu[i]
}
chisq.obs <- sum(diff.obs[])
chisq.rep <- sum(diff.rep[])
p.value <- step(chisq.rep - chisq.obs)  #Mean of this node should be near 0.5
lambda.constant ~ dgamma(0.5, 0.0001) #Jeffreys prior for lambda
}
Data
Load data from file
inits
list(lambda.constant=0.001)

```

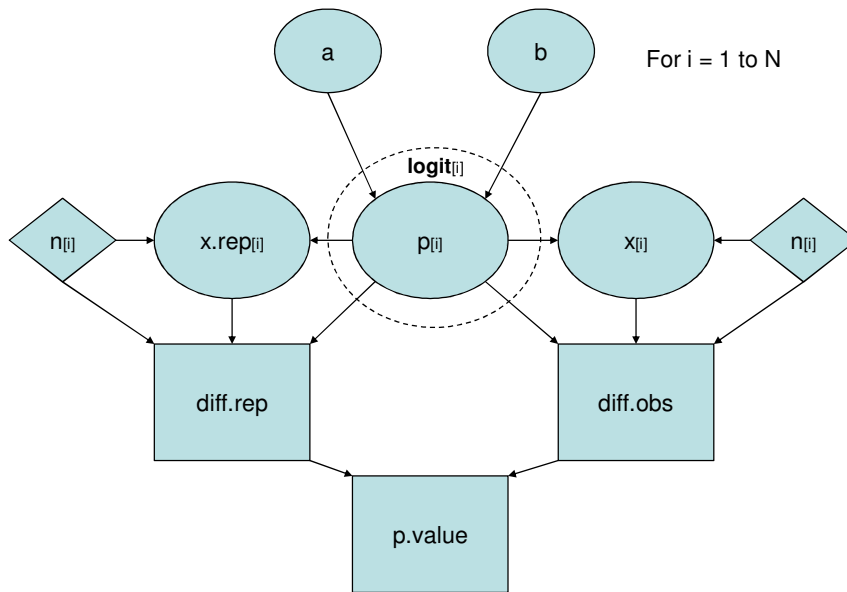
• Script 7 WinBUGS script to test for poolability of Poisson data

5.4 Time-Trend Models for p and λ

Example 10 analyzed valve leakage data over a period of nine years and concluded that p might be increasing with time over that period. In this section, we will see how to develop models in which p and λ are explicit functions of time, relaxing the assumptions of constant p and constant λ in the binomial and Poisson distribution, respectively. This introduces new unknown parameters of interest and makes the statistical inference significantly more complicated mathematically. However, modern tools such as WinBUGS make this analysis only slightly more complicated than the single-parameter cases analyzed earlier.

5.4.1 Time-Dependent Model for p in the Binomial Distribution

The standard way to allow for p to vary in time (monotonically) is by letting $f(p) = a + bt$, where $f(p)$ is a user-defined function of p , a and b are constants, and t is time. If $b = 0$, then this reduces to the constant- p model discussed earlier. If $b > 0$, then p is increasing with time. If $b < 0$, p is decreasing with time. The new parameters to be estimated are a and b . For Bayesian inference it is common to use noninformative prior distributions for these parameters. Of course, if prior information is available about the parameters a and b , an informative prior can be used, but this is not common in practice. The WinBUGS script below carries out inference with the most commonly used function of p . The underlying DAG model is shown in Figure 24. Other functions that are sometimes used are shown in the script, but are commented out. In many cases, changing the link function will not usually make any practical difference in the inference.



• Figure 24 DAG model for inference with time trend in p

```

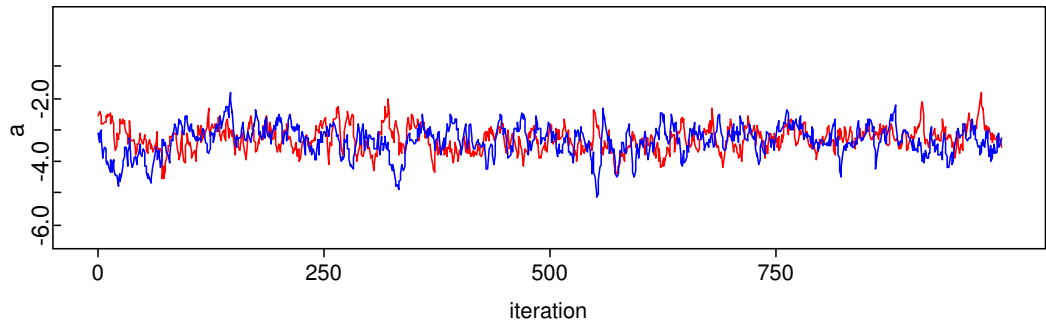
model {
for (i in 1:N) {
    x[i] ~ dbin(p[i], n[i])    #Binomial distribution for failures in each year
    logit(p[i]) <- a + b*i    #Use of logit() link function for p[i]
    #probit(p[i]) <- a + b*i    #Use of probit() link function for p[i]
    #cloglog(p[i]) <- a + b*i    #Use of complementary loglog function for p[i]
    #Model validation
    x.rep[i] ~ dbin(p[i], n[i])
    diff.obs[i] <- pow(x[i] - n[i]*p[i], 2)/(n[i]*p[i])
    diff.rep[i] <- pow(x.rep[i] - n[i]*p[i], 2)/(n[i]*p[i])
  }
  logit(p[10]) <- a + b*10    #Used to predict p in 10th year
  a~dflat()                  #Noninformative prior for a
  b~dflat()                  #Noninformative prior for b
  chisq.obs <- sum(diff.obs[])
  chisq.rep <- sum(diff.rep[])
  p.value <- step(chisq.rep - chisq.obs)
}
Data
x[]      n[]
4        52
2        52
3        52
1        52
4        52
3        52
4        52
9        52
6        52
END
list(N=9)
Inits
list(a=1, b=0)
list(a=-1, b=0.1)

```

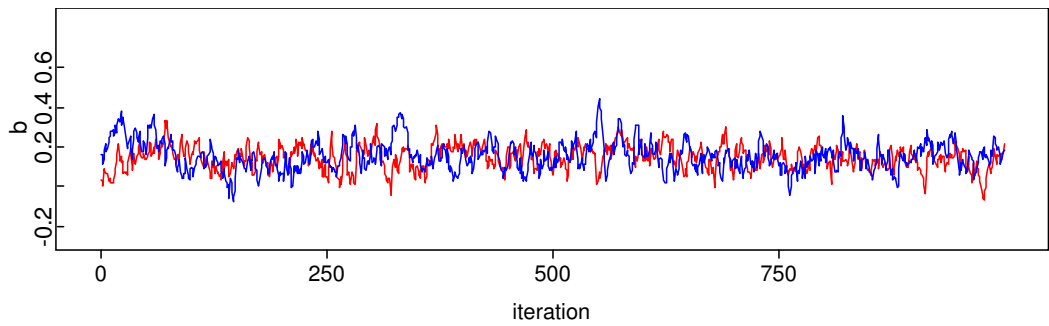
- Script 8. WinBUGS script for modeling a time trend in p.

Because this is a more complicated model, we will run two chains, starting at different points, as an aid in deciding when convergence to the posterior distribution has been achieved. Also, we must give starting values for each chain. These are listed in the Inits portion of the script. Running 1,000 iterations gives the

following history plots for the parameters a and b. It appears from these plots that the two chains are well mixed, indicating that we have achieved convergence within the first 1,000 iterations.

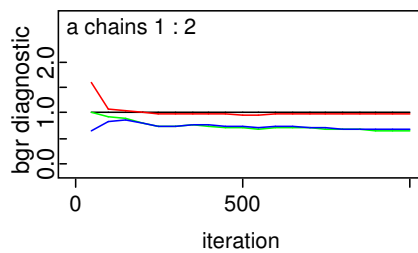


• Figure 25. Plot of values of first 1,000 values for “a” parameter, illustrating convergence.



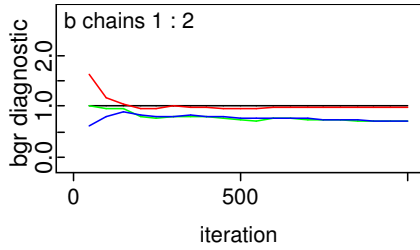
• Figure 26. Plot of values of first 1,000 values for “b” parameter, illustrating convergence.

We can also plot the BGR diagnostic, since we have run more than one chain. The plots for both parameters are indicative of convergence.



• Figure 27. BGR diagnostic for “a” parameter, indicating convergence.

For convergence, the red line in the BGR plot should be close to 1.0, and the blue and green lines should be stable.



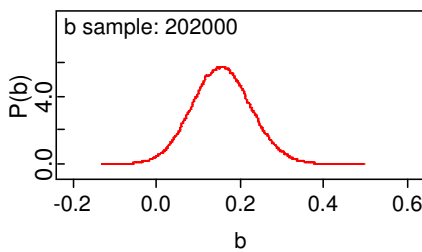
• Figure 28. BGR diagnostic for “b” parameter, indicating convergence.

Now that we are confident of convergence, we run another 100,000 iterations to estimate parameter values. The posterior mean and 90% credible interval for each of the parameters is listed below, along with the iteratively reweighted least-squares frequentist estimates.

• Table 12 Parameter estimates for logistic model for time trend in valve leakage

Parameter	Post. Mean	Frequentist Point Estimate	90% Credible Interval	Approx. 90% Confidence Interval
a	-3.36	-3.32	(-4.13, -2.65)	(-4.05, -2.59)
b	0.156	0.153	(0.04, 0.27)	(0.04, 0.27)

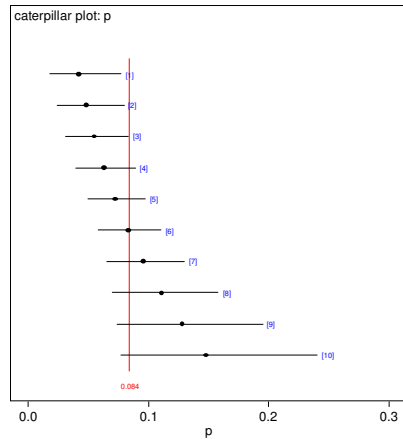
Examining the posterior density for the “b” parameter will help us judge the significance of any trend that might be present: if the posterior distribution is mostly to the right of zero, this indicates an increasing trend, and vice versa if the posterior distribution is mostly to the left of zero. By monitoring the b node, we obtain a posterior probability of at least 0.975 that $b > 0$, suggesting a significant increasing trend in p. The plot of the posterior distribution for b below shows this graphically.



• Figure 29. Posterior distribution for “b” parameter. Values below zero are very unlikely, suggesting an increasing trend in p over time.

We can also quantify the ability of this model to replicate the observed data by monitoring the mean of the p.value node. As before, a mean near 0.5 indicates good replicative ability. In this example, the mean is 0.47. Comparing this with the value of about 0.18 obtained for a constant-p model, we see that a model in which p increases with time is much better at replicating the observed data.

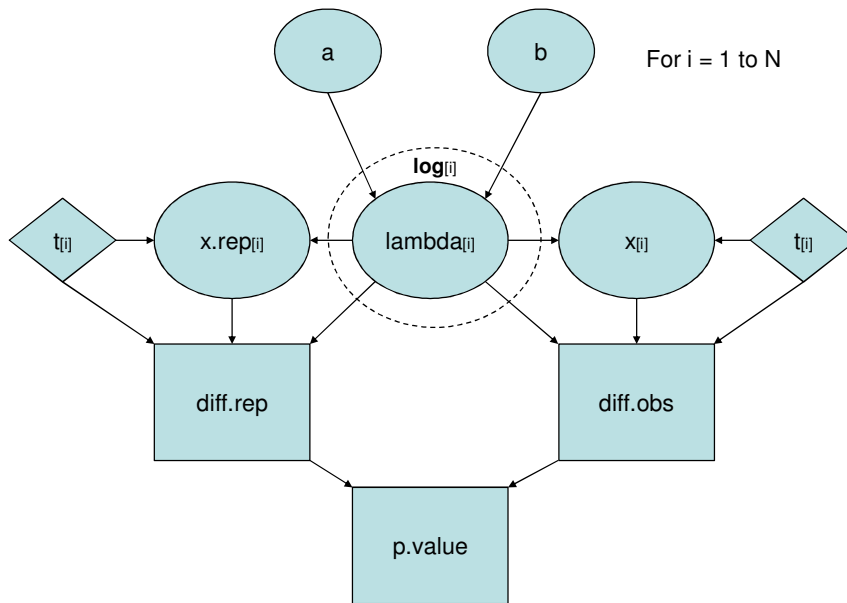
Finally, we can use this model to estimate p in the next year (year 10). This is given by node $p[10]$ in the script above, and would be what we would use in a PSA. Monitoring this node, we find a posterior mean of 0.15 and a 90% interval of (0.085, 0.22). Compare this with the estimates from the constant- p model of 0.08 for the mean and (0.06, 0.099) for the 90% interval. The figure below shows the 95% credible intervals for each year, based on the trend model for p .



• Figure 30 Plot of 95% interval for p in each year based on a trend model in which p is increasing with time.

5.4.2 Time-Dependent Model for λ in the Poisson Distribution

The standard way to allow for λ to vary in time (monotonically) is by letting $f(\lambda) = a + bt$, where $f(\lambda)$ is a user-defined function of λ , a and b are constants, and t is time. If $b = 0$, then this reduces to the constant- λ model discussed earlier. If $b > 0$, then λ is increasing with time. If $b < 0$, λ is decreasing with time. The new parameters to be estimated are a and b . It is common to use noninformative prior distributions for these parameters in Bayesian inference. Of course, if prior information is available, and informative prior can be used, but this is not common in practice. The WinBUGS script below carries out inference for the data of Example 11 with the most commonly used function of λ . The underlying DAG model is shown in Figure 31.



• Figure 31 DAG model for inference with time trend in λ .

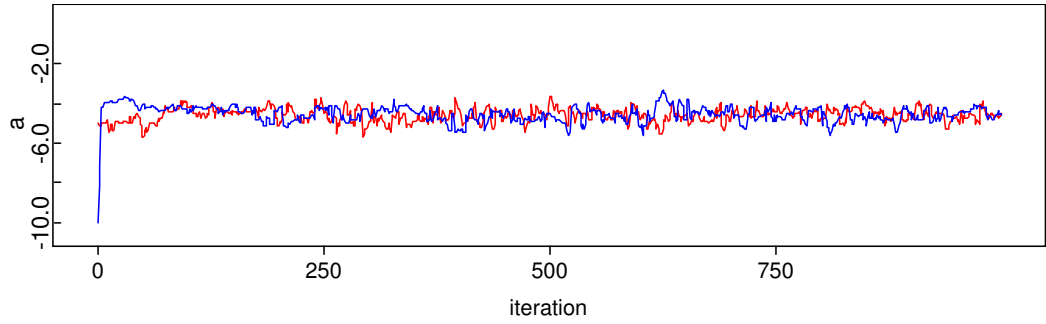
```

model {
for(i in 1:N) {
year[i] <- T[i]
x[i] ~ dpois(mu[i])      #Poisson distribution for number of failures in each source
x.rep[i] ~ dpois(mu[i]) #Replicate value from posterior predictive distribution
mu[i] <- lambda[i]*t[i]  #Parameter of Poisson distribution
log(lambda[i]) <- a + b*T[i]    #Loglinear model for lambda
diff.obs[i] <- pow(x[i] - mu[i], 2)/mu[i]
diff.rep[i] <- pow(x.rep[i] - mu[i], 2)/mu[i]
}
#Predict future values of lambda
log(lambda[15]) <- a + b*15
chisq.obs <- sum(diff.obs[])
chisq.rep <- sum(diff.rep[])
p.value <- step(chisq.rep - chisq.obs)    #Mean of this node should be near 0.5
a~dflat()      #Noninformative priors for a and b
b~dflat()
}
data
Load data from file
inits
list(a=-5, b=0.1)
list(a=-10, b=0.01)

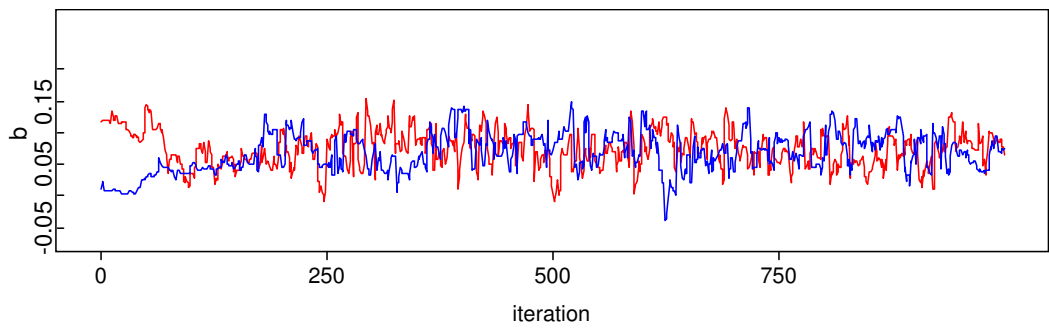
```

- Script 9. WinBUGS script for modeling time trend in lambda.

Because this is a more complicated model, we will run two chains, starting at different points, as an aid in deciding when convergence has taken place. Also, we must give starting values for each chain. These are listed in the Inits portion of the script. Running 1,000 iterations gives the following history plots for the parameters a and b. It appears from these plots that we have convergence within the first 1,000 iterations.



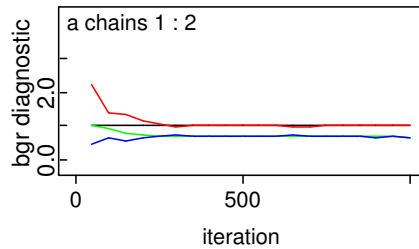
• Figure 32. Plot of values of first 1,000 values for “a” parameter, illustrating convergence.



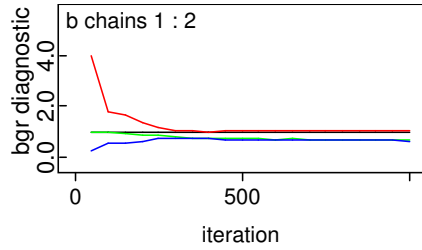
• Figure 33. Plot of values of first 1,000 values for “b” parameter, illustrating convergence.

We can also plot the BGR diagnostic, since we have run more than one chain. The plots for both parameters are indicative of convergence.

For convergence, the red line in the BGR plot should be close to 1.0, and the blue and green lines should be stable.

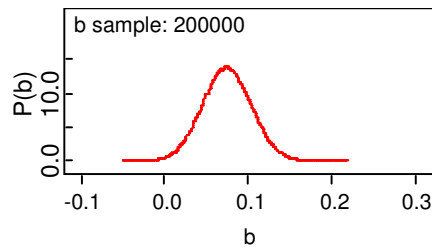


• Figure 34. BGR diagnostic for “a” parameter, indicating convergence.



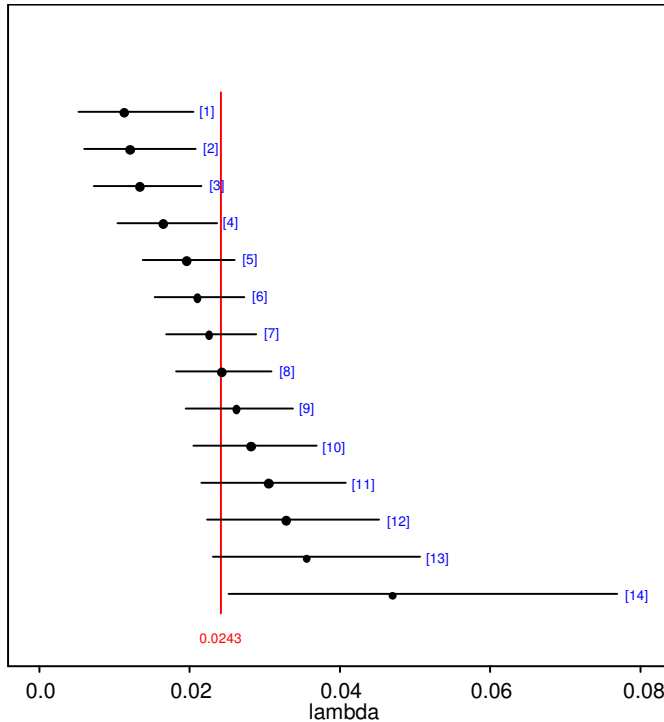
• Figure 35. BGR diagnostic for “b” parameter, indicating convergence.

Now that we are confident of convergence, we run another 100,000 iterations to estimate parameter values. Examining the posterior density for the “b” parameter will help us judge the significance of any trend that might be present: if the posterior distribution is mostly to the right of zero, this indicates an increasing trend, and vice versa if the posterior distribution is mostly to the left of zero. The posterior mean of “a” was -4.525, with a 90% interval of (-5.1, -3.98). The posterior mean of b was 0.07. The 90% credible interval for b was (0.026, 0.12). The marginal posterior density of b is shown below.



• Figure 36 Posterior distribution for “b” parameter. Values above zero are very likely, suggesting an increasing trend in λ over time.

We can also quantify the ability of this model to replicate the observed data by monitoring the mean of the p.value node. As before, a mean near 0.5 indicates good replicative ability. The Bayesian p-value for the loglinear model was 0.01, compared with the frequentist p-value of 0.0055. These results are again indicative of a poorly fitting model. Therefore, although we can use this model to predict λ in the next year, we would not want to use this predicted value in a PSA. The figure below shows the 95% credible intervals for each year, based on the loglinear trend model for λ .



• Figure 37 Plot of 95% interval for λ in each year based on a trend model in which λ is increasing with time

5.5 Population Variability Models

5.5.1 Population Variability Model for p in the Binomial Distribution

Often, we have data from multiple sources that we would like to use in estimating p for a component. The data are from similar components in similar systems, and we want to develop a prior distribution that captures both the central estimate from these sources and the variability in p indicated by the data. Such analysis is referred to as *hierarchical Bayesian* analysis, because the prior distribution is developed in stages or hierarchies. We first specify a prior to represent the variability in the data sources (the “population variability curve”, or PVC) and then specify a second prior describing the epistemic uncertainty in the parameters of the first-stage prior. The analysis is quite complicated mathematically. However, WinBUGS again makes the analysis straightforward.

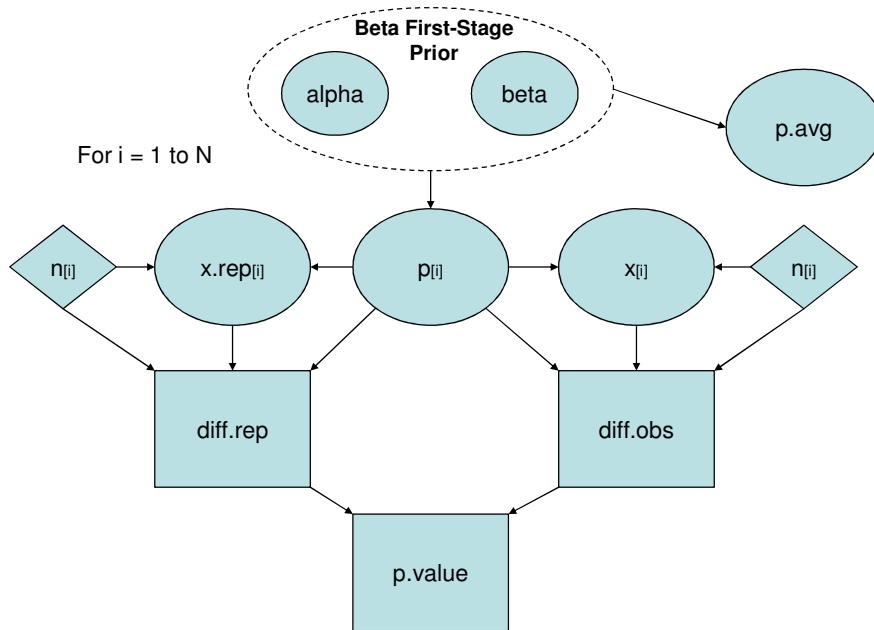
Sensitivity to the functional form of the PVC should always be checked.

We must begin by choosing a functional form for the PVC. Often a conjugate form is chosen (e.g., beta distribution if the parameter is p in a binomial distribution), but this is not a requirement. In fact, in cases where there is extremely high variability in p , choosing a conjugate form can lead to erroneous results. In such cases, a distribution with a heavier tail, such as the lognormal, may be a better choice. For our example, we will take the PVC to be a beta distribution with parameters α and β .

As the second stage in the hierarchy, we must specify a second-stage prior distribution for α and β (α and β are called *hyperparameters* and the second-stage prior is referred to as a *hyperprior*). This is usually done by choosing independent, diffuse distributions for both parameters. Choosing hyperpriors that are diffuse, yet that avoid numerical difficulties can be a bit of an art, and this is one of the reasons why analysis of population variability should always be reviewed by an expert analyst.

Because of potential convergence problems and other numerical difficulties, population variability analysis should always be checked by an expert.

The WinBUGS script used to perform the analysis is shown below. The underlying DAG model is shown in Figure 38. Note that the parameters of the beta first-stage prior are now stochastic nodes rather than constant nodes, as in earlier problems. Node p.avg will display the PVC representing variability in p across the sources. It will be the prior distribution for inference about the component failure probability, and will be updated with data collected from operation or testing of our system. Note that this script also calculates the p.value node, which is used to check the ability of the model to replicate the observed data.



• Figure 38 DAG model for hierarchical Bayesian inference in binomial distribution

```

model {
for(i in 1:N) {
    x[i] ~ dbin(p[i], n[i])           #Binomial model for number of events in each source
    p[i] ~ dbeta(alpha, beta)       #First-stage beta prior
    x.rep[i] ~ dbin(p[i], n[i])     #Replicate value from posterior predictive distribution
    #Generate inputs for Bayesian p-value calculation
    diff.obs[i] <- pow(x[i] - n[i]*p[i], 2)/(n[i]*p[i]*(1-p[i]))
    diff.rep[i] <- pow(x.rep[i] - n[i]*p[i], 2)/(n[i]*p[i]*(1-p[i]))
  }

  p.avg ~ dbeta(alpha, beta)        #Average beta population variability curve
#Calculate Bayesian p-value
  chisq.obs <- sum(diff.obs[])
  chisq.rep <- sum(diff.rep[])
  p.value <- step(chisq.rep - chisq.obs) #Mean of this node should be near 0.5
#Hyperpriors for beta first-stage prior
  alpha ~ dexp(1)
  beta ~ dgamma(0.0001, 0.0001)
}

data
x[]    n[]
0      164
1      322
0      13
2      186
2      6151
0      3264
1      4747
1      3211
4      457
4      456
3      277

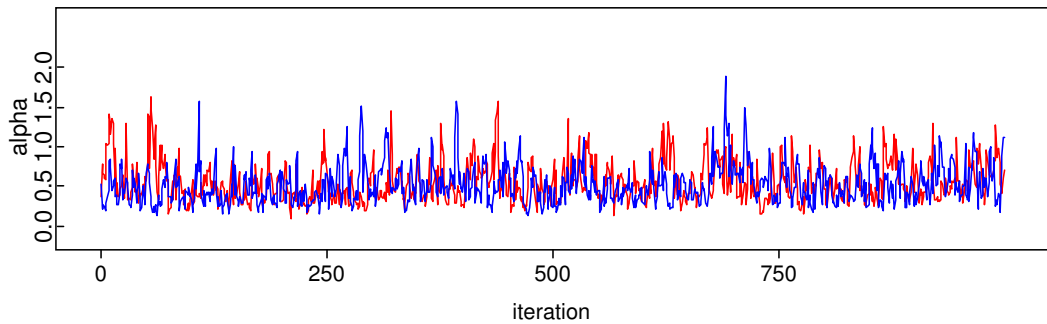
END
list(N=11)
inits
list(alpha=0.5, beta=200) #Chain 1
list(alpha=2, beta=100) #Chain 2

```

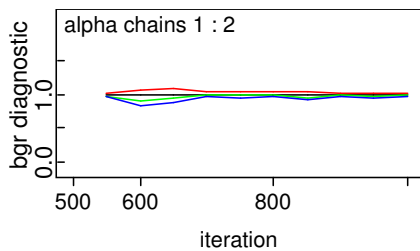
• Script 10. WinBUGS script for analyzing population variability in p.

The initial values for each chain are estimates distributed around what are felt to be likely values of α and β . Often, accurate estimates are not needed. However, in some cases, more care will have to be used in picking initial values, and more than two chains may be needed. Such problems will require expert assistance, but should not be too common in practice.

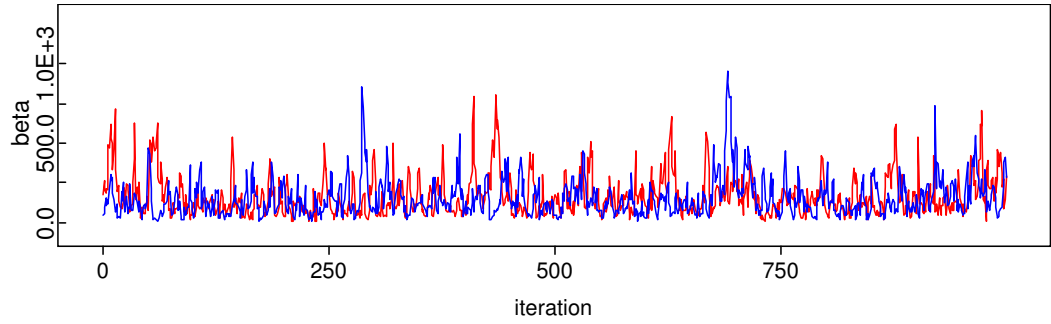
We run this script for 1,000 iterations, and check for convergence by examining history plots and BGR diagnostics for each parameter. These are shown below, and indicate convergence within the first 1,000 iterations, so we will discard these as burn-in samples.



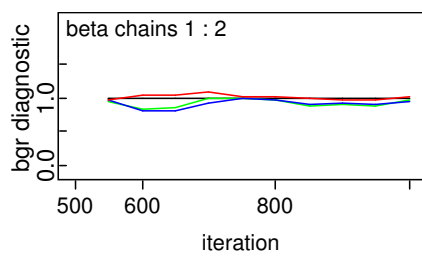
• Figure 39 Plot of first 1,000 values of α , indicating convergence.



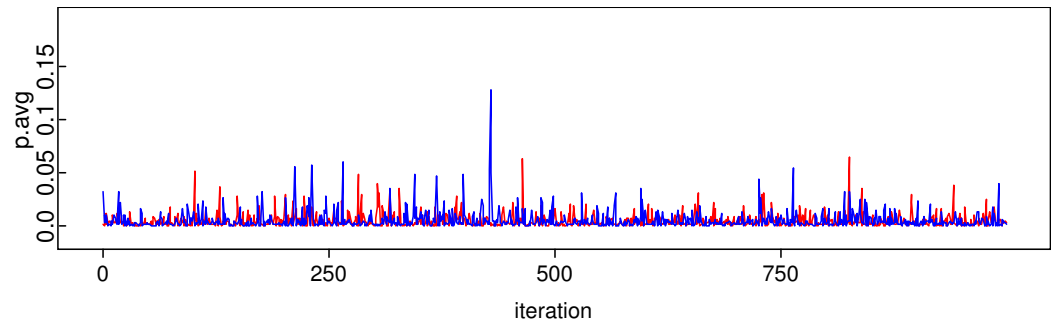
• Figure 40 BGR diagnostic for α , indicating convergence.



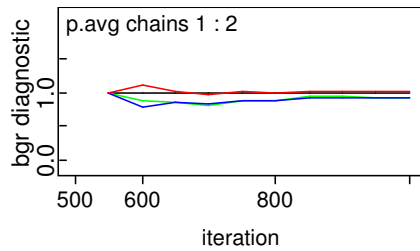
• Figure 41 Plot of first 1,000 values of β , indicating convergence.



• Figure 42 BGR diagnostic for β , indicating convergence.



• Figure 43 Plot of first 1,000 values of $p.avg$, indicating convergence.



• Figure 44 BGR diagnostic for p.avg, indicating convergence.

This script runs considerably more slowly than earlier scripts we have used, so we may want to make parameter estimates with less than the usual 100,000 iterations. Let us try 10,000 additional iterations and examine the Monte Carlo error for each parameter to judge if enough samples have been taken to accurately estimate each parameter. The results are listed below.

	mean	sd	MC_error	start	sample
alpha	0.5595	0.2737	0.004982	501	21000
beta	169.3	141.4	2.665	501	21000
p.avg	0.004746	0.01073	8.187E-5	501	21000

A heuristic to decide if more samples are needed is that the Monte Carlo error should be no more than about 5% of the standard deviation. This is satisfied for all three parameters, so we can stop after 10,000 iterations (11,000 including burn-in). The posterior mean of p.avg is 0.0047, with a 90% credible interval of (6.04×10^{-6} , 0.018). This is a very wide range of variability.

Note that the distribution of node p.avg is not a beta distribution, despite what one might be led to believe by looking at Script 10. It is a weighted average of beta distributions, with the weights provided by the posterior distribution of α and β . Thus, it cannot be written in closed form. However, if desired, it can be approximated by a distribution of known form to aid in further analysis, although this is not necessary in WinBUGS.

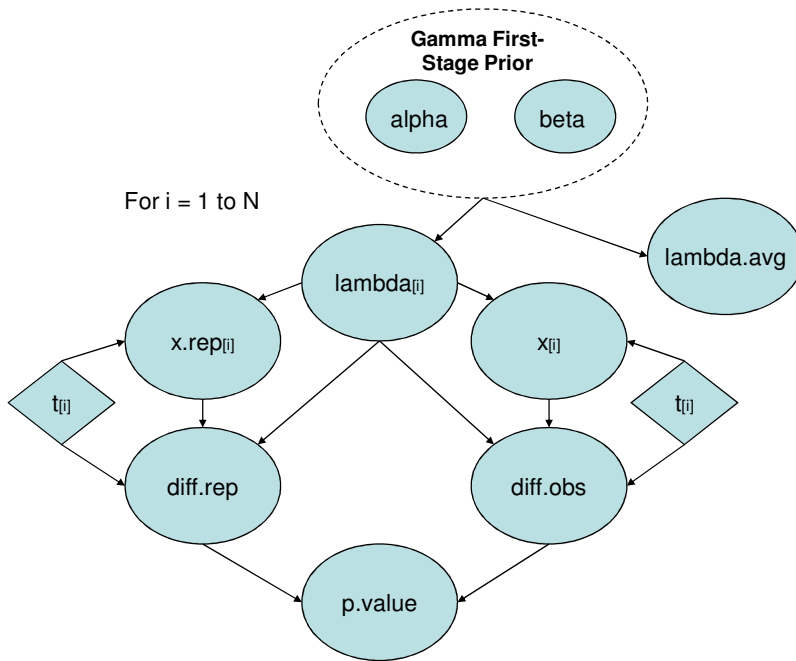
5.5.2 Population Variability Model for λ in the Poisson Distribution

In Example 11, we had data collected for a component over a period of time. We decided that we could not pool the data from these sources to generate a prior distribution for λ . Also, we decided that a model in which a trend in λ was present over time did not perform well in replicating the observed data. We need to develop a prior distribution that captures both the central estimate from these sources and the variability in lambda indicated by the data. The approach is the same as for p in the binomial distribution described in Sec. 5.5.1.

We begin by specifying a first-stage prior describing the source-to-source variability in lambda, and then a second-stage prior (hyperprior) for the parameters of the first-stage prior.

We will use a conjugate gamma distribution as the first-stage prior, although other functional forms can be used. We will use independent diffuse priors for the parameters of the first-stage gamma prior. The underlying DAG model is shown in Figure 45. Note that the parameters of the gamma first-stage prior are now stochastic nodes rather than constant nodes, as in earlier problems. The WinBUGS script used to carry out the analysis is shown below.

Because of potential convergence problems and other numerical difficulties, population variability analysis should always be checked by an expert.



• Figure 45 DAG model for hierarchical Bayesian inference in Poisson distribution

```

model {
for(i in 1:N) {
  x[i] ~ dpois(mu[i])           #Poisson distribution for number of failures in each source
  mu[i] <- lambda[i]*t[i]      #Parameter of Poisson distribution
  lambda[i] ~ dgamma(alpha, beta) #First-stage gamma prior
  x.rep[i] ~ dpois(mu[i])      #Replicate value from posterior predictive distribution
  diff.obs[i] <- pow(x[i] - mu[i], 2)/mu[i] #Inputs to calculate Bayesian p-value
  diff.rep[i] <- pow(x.rep[i] - mu[i], 2)/mu[i]
}
#Overall average population variability for lambda
lambda.avg ~ dgamma(alpha, beta)
chisq.obs <- sum(diff.obs[])    #Calculate Bayesian p-value
chisq.rep <- sum(diff.rep[])
p.value <- step(chisq.rep - chisq.obs) #Mean of node should be near 0.5
#Reparameterize in terms of mean and coefficient of variation
alpha <- pow(CV, -2)
beta <- alpha/mean
CV ~ dexp(1)                   #Maximum entropy hyperprior for coefficient of variation
mean ~ dgamma(0.0001, 0.0001) #Diffuse hyperprior on mean
}
data
x[]      t[]
0        87600
7        525600
1        394200
0        87600
8        4555200
0        306600
0        394200
0        569400
5        1664400
1        3766800
4        3241200
2        1051200
END
list(N=12)
inits
list(CV=1, mean=0.00001)
list(CV=2, mean=0.000001)

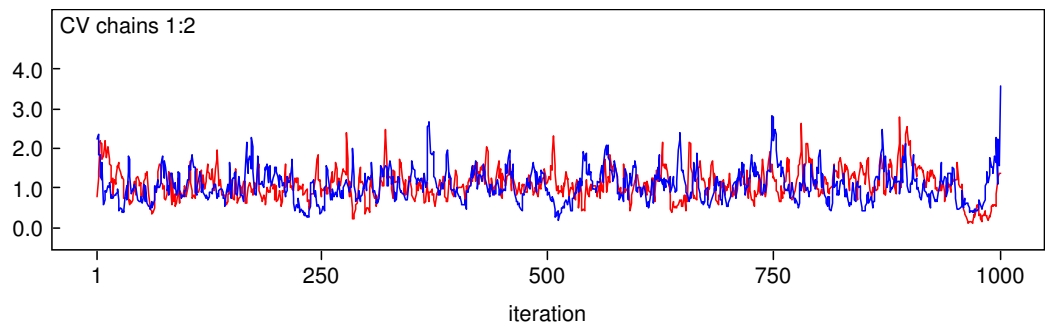
```

- Script 11 WinBUGS script for analyzing population variability in lambda with Poisson likelihood.

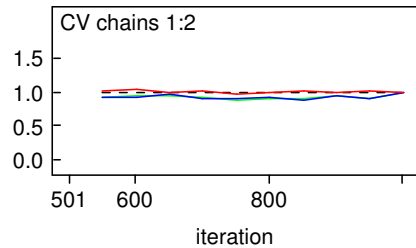
The initial values for each chain are estimates distributed around what are felt to be likely values of CV and the mean. Often, accurate estimates are not needed. However, in some cases, more care will have to be used in picking initial values, and more than two chains may be needed. Such problems will require expert assistance, but should not be too common in practice.

Note that Script 11 changes the parameters from alpha and beta to different parameters to improve numerical performance. This re-parameterization is an example of where expert consultation may be necessary in complex problems.

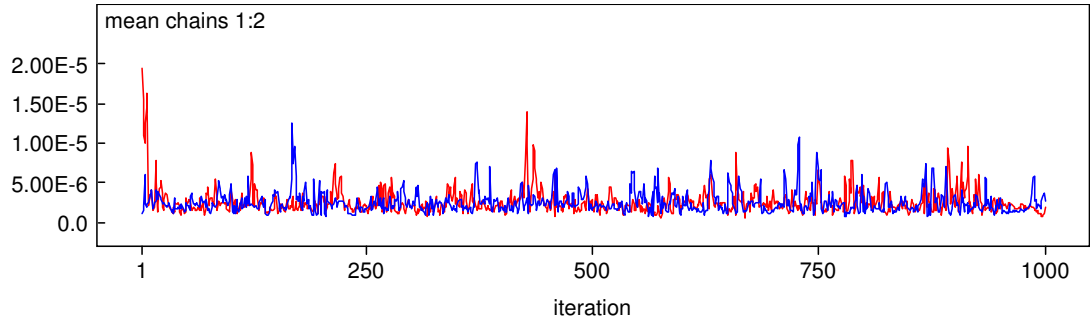
We run this script for 1,000 iterations, and check for convergence by examining history plots and BGR diagnostics for each parameter. These are shown below, and indicate convergence within the first 1,000 iterations, so we will discard these as burn-in samples.



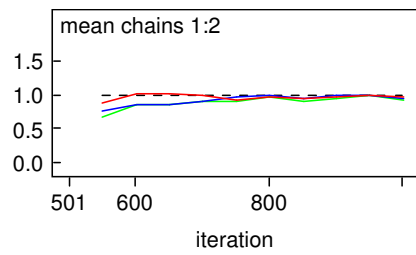
• Figure 46 Plot of first 1,000 values of CV, indicating convergence.



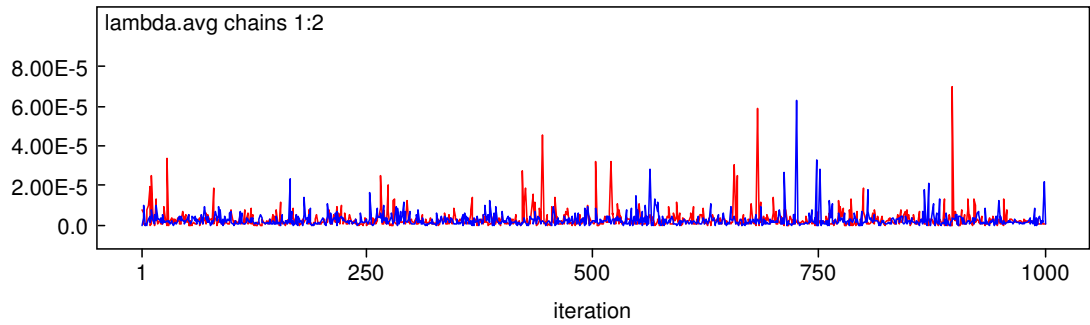
• Figure 47 BGR plot for CV, indicating convergence.



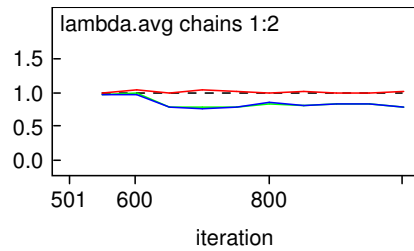
• Figure 48 Plot of first 1,000 values of mean parameter, indicating convergence.



• Figure 49 BGR plot for mean parameter, indicating convergence.



• Figure 50 Plot of first 1,000 values of lambda.avg, indicating convergence.



• Figure 51. BGR plot for lambda.avg, indicating convergence.

The sampling is much faster with a gamma first-stage prior distribution than for a beta distribution, so WinBUGS runs much faster than it did above when we were estimating a hierarchical model for p in the binomial distribution. Running 100,000 iterations gives a posterior mean for lambda.avg of 2.45×10^{-6} /hour, with a 90% credible interval of $(3.91 \times 10^{-8}, 7.77 \times 10^{-6})$.

Note that the distribution of node lambda.avg is not a gamma distribution, despite what one might be led to believe by looking at Script 11. It is a weighted average of gamma distributions, with the weights provided by the posterior distribution of the hyperparameters, CV and the mean. Thus, it cannot be written in closed form. However, if desired, it can be approximated by a distribution of known form to aid in further analysis, although this is not necessary in WinBUGS. To update the average PVC with additional data, add another line to the data block, change N to 13, and rerun the script. Then, monitoring lambda[13] will give the posterior distribution for the circuit board in our ATCS system.

The p.value for this model, which captures source-to-source variability in lambda, is 0.43, compared with 0.002 for the model that pools the data. Thus, the population variability model is much better able to replicate the data for the 12 sources under consideration.

5.6 More Complex Models for Time to Failure – Nonrepairable Failure

There are cases in which the simplest model for time to failure, the exponential distribution, is not adequate. In these cases, more complicated stochastic models may need to be used. In this section, we analyze cases in which the failed component is not repaired. We will treat failure with repair in a subsequent section, where we will also cover statistical tests that can be useful in deciding when a given model is appropriate.

5.6.1 Alternatives to Exponential Distribution for Random Durations

There are three commonly used distributions that are alternative aleatory models to the exponential distribution. These are the gamma distribution, the Weibull distribution, and the lognormal distribution. There are other less commonly used possibilities, which we will not treat here. We treat only the case of a complete sample in this section; examples of censoring are discussed in Sec. 3.8.

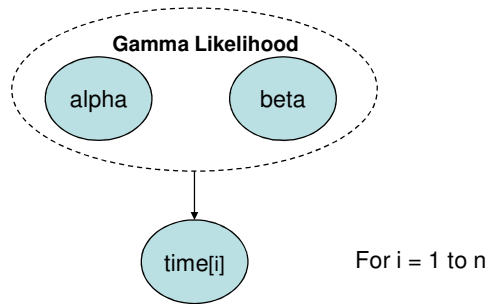
5.6.1.1 Gamma Distribution as Likelihood for Random Durations

As we saw earlier when we used the gamma distribution as a prior for λ , two parameters are required to specify the distribution, so now we must do Bayesian inference for both α and β . Again, WinBUGS makes this inference straightforward.

- Example 12 Times between failure for analysis with gamma distribution

The following times to failure (in hours) have been collected for a cooling pump: 1258, 1388, 1022, 1989, 2024, 1638, 390, 4362, 2240, 1215, 2146, 655. Estimate the parameters of a gamma distribution describing these times to failure. Also, compare the gamma distribution with the fit provided by an exponential distribution.

In the script below, we have specified a gamma distribution as the aleatory model (i.e., likelihood) for each failure time, and we have placed independent noninformative priors on α and β ; if information about α or β were available, an informative prior could be used. Figure 52 shows the underlying DAG for this problem. Note that the parameters of the gamma distribution are stochastic nodes.



• Figure 52 DAG model for gamma-distributed random durations

```

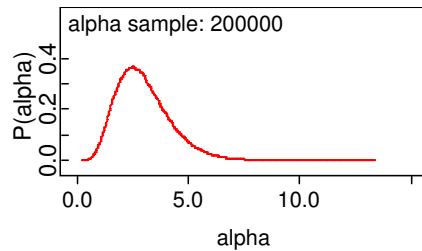
model {
for(i in 1:N) {
  t[i] ~ dgamma(alpha, beta)      #Gamma likelihood function for N failure times
}
#Model validation section
for(i in 1:N) {
  t.rep[i] ~ dgamma(beta, alpha)
  t.rep.ranked[i] <- ranked(t.rep[], i)
  t.ranked[i] <- ranked(t[], i)
#Test for gamma distribution
  #F.obs[i] <- (gammap(alpha, t.ranked[i])/rate)/exp(loggam(alpha))
  #F.rep[i] <- (gammap(alpha, t.rep.ranked[i])/rate)/exp(loggam(alpha))
  diff.obs[i] <- pow(F.obs[i] - (2*i-1)/(2*N), 2)
  diff.rep[i] <- pow(F.rep[i] - (2*i-1)/(2*N), 2)
}
CVM.obs <- sum(diff.obs[])
CVM.rep <- sum(diff.rep[])
p.value <- step(CVM.rep - CVM.obs)
alpha ~ dgamma(0.0001, 0.0001)    #Diffuse priors for alpha and beta
beta ~ dgamma(0.0001, 0.0001)
}
data
list(t=c(1258, 1388, 1022, 1989, 2024, 1638, 390, 4362, 2240, 1215, 2146, 655), N=12)
inits
list(alpha=2, beta=0.1)
list(alpha=0.5, beta=1)

```

• Script 12 WinBUGS script for Bayesian inference of random durations using gamma likelihood.

With relatively few observed times, it can be difficult to discriminate between the exponential distribution and alternative models.

Running 1,000 iterations and checking α and β for convergence using the history and BGR plots as we have done in earlier sections, we see that the model converges very rapidly. We discard the first 1,000 iterations and run another 100,000 iterations to estimate parameter values, obtaining the posterior distribution for α shown below. If $\alpha = 1$, then the gamma distribution reduces to the exponential distribution. In this example the posterior mean of α is 2.97, with a 90% credible interval of (1.34, 5.19). The posterior mean and 90% credible interval for β are 0.0018 and $(7.1 \times 10^{-4}, 0.0032)$. The marginal posterior density for α is mostly to the right of one, suggesting an increasing failure rate as a function of run time for this component.



• Figure 53 Posterior distribution for α indicates that values above 1 are likely

The MLE for α is 3.36, even further from 1. Note also that the MLE is known to be biased, but the bias is difficult to correct. Also, the MLE can be tricky to obtain, as numerically maximizing the log-likelihood is somewhat delicate.

To test whether an exponential distribution could describe these failure times adequately, we use a Bayesian analog of the Cramer-von Mises test, which is illustrated in the script below.

Bayesian analog of Cramer-von Mises test

```

model {
for(i in 1:N) {
  t[i] ~ dexp(lambda) #Exponential distribution for observed times
  t.ranked[i] <- ranked(time[], i) #Order observed times
  t.rep[i] ~ dexp(lambda) ##Replicate time from posterior predictive distribution
  F.obs[i] <- 1 - exp(-lambda*time.ranked[i]) #CDF for observed ranked times
  F.rep[i] <- 1 - exp(-lambda*time.rep.ranked[i]) #CDF for replicated ranked times
  diff.obs[i] <- pow(F.obs[i] - (2*i-1)/(2*N), 2)
  diff.rep[i] <- pow(F.rep[i] - (2*i-1)/(2*N), 2)
}
CVM.obs <- sum(diff.obs[]) #Observed CVM statistic
CVM.rep <- sum(diff.rep[]) #Replicated CVM statistic
p.value <- step(CVM.rep - CVM.obs) #Mean should be near 0.5

```

• Script 13 WinBUGS script for Bayesian analog of Cramer-von Mises test of exponential distribution for failure times

Running this script in the usual way gives a Bayesian p-value of 0.16, indicating that the exponential distribution may be adequate but not outstanding for modeling the observed failure data. In a similar way, we can check the fit of the gamma distribution using Script 12. Running this script, we find a Bayesian p-value of 0.xx.⁵

5.6.1.2 Weibull Distribution as Likelihood for Random Durations

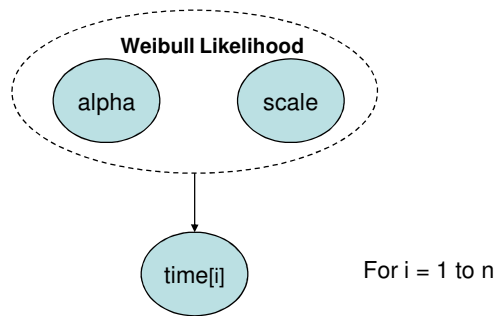
The Weibull is another two-parameter aleatory model for random durations. Like the gamma distribution, it has a shape parameter, which we will denote as α . If $\alpha = 1$, the Weibull distribution reduces to the exponential distribution, just as did the gamma distribution above. When α is less than one, the rate of occurrence is

⁵ The latest version of WinBUGS (OpenBUGS) has a problem with the gammap() logical function, so this test cannot be used. We anticipate this problem will be corrected in the next release.

decreasing with time, and when α is greater than one, it is increasing with time, analogously to the gamma distribution.

- Example 13 Inference for Weibull distribution as aleatory model for circulating pump failure times in ATCS system

We will carry out Bayesian inference for the times in Example 12, under the assumption that the aleatory model is a Weibull distribution with parameters called alpha and scale. The WinBUGS script shown below is used for this analysis with diffuse priors on the Weibull parameters. The underlying DAG is shown in Figure 54.



• Figure 54 DAG model for Weibull-distributed random durations

```

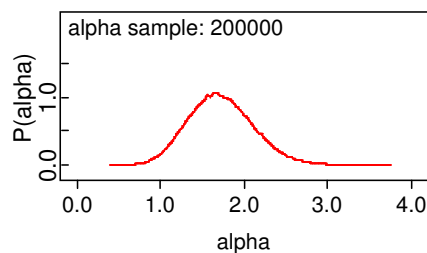
model {
for(i in 1:N) {
    t[i] ~ dweib(alpha, scale)      #Weibull likelihood function for n times
}
alpha ~ dgamma(0.0001, 0.0001)    #Diffuse priors for alpha and beta
scale ~ dgamma(0.0001, 0.0001)
}
data
list(t=c(1258, 1388, 1022, 1989, 2024, 1638, 390, 4362, 2240, 1215, 2146, 655), N=12)
inits
list(alpha=2, scale=1)
list(alpha=0.5, scale=10)

```

- Script 14 WinBUGS script for Bayesian inference of random durations using Weibull likelihood.

Running the script in the usual way produces the posterior distribution for α shown below. As for the gamma distribution above, the posterior distribution indicates that values of α greater than one are most likely, suggesting that the failure rate is increasing with operating time, and that the exponential model (with constant repair rate) may not be adequate. For reference, the MLE of α is 1.85, with an asymptotic standard error of 0.39.

We can also carry out the Bayesian analog of the Cramer-von Mises test to see how well a Weibull distribution can replicate the observed data. We find a Bayesian p-value of 0.49, significantly closer to 0.5 than was the case for the exponential distribution above.



- Figure 55 Posterior distribution of α for Example 13.

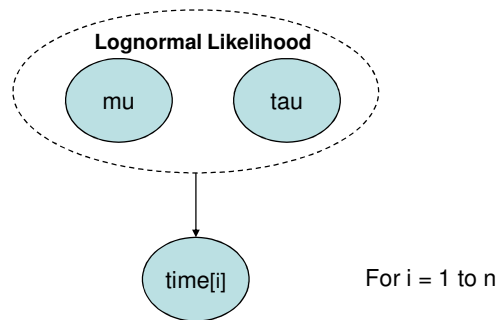
5.6.1.3 Lognormal Distribution as Likelihood for Random Durations

In addition to its use as a nonconjugate prior as described earlier, the lognormal distribution is also a popular aleatory model for random durations. It has the interesting property that the rate of the process initially

increases, and then decreases monotonically. The period over which the rate increases is often very short, so that it can be used as an alternative to the gamma or Weibull distribution with shape parameter less than one. It has a heavier tail than the gamma or Weibull distribution, and may therefore produce more conservative results. The parameters of the lognormal distribution (μ and τ) cannot be interpreted in the manner of the gamma and Weibull shape parameter, and the lognormal distribution does not reduce to the exponential distribution for certain parameter values, as do the gamma and Weibull distribution.

- Example 14 Inference for lognormal distribution as aleatory model for failure times

We will carry out Bayesian inference for the times in Example 12, under the assumption that the aleatory model is a lognormal distribution with parameters called μ and τ . The WinBUGS script shown below is used for this analysis with diffuse priors on the lognormal parameters. The underlying DAG model is shown in Figure 56.



• Figure 56 DAG model for lognormally-distributed random durations

```

model {
for(i in 1:N) {
    t[i] ~ dlnorm(mu, tau) #Lognormal likelihood function for N failure times
    t.rep[i] ~ dlnorm(mu, tau)
  }

  #Model validation section
for(i in 1:N) {
  
```

```

t.ranked[i] <- ranked(t[], i)
t.rep.ranked[i] <- ranked(t.rep[], i)

#Test for lognormal distribution
F.obs[i] <- phi((log(t.ranked[i]) - mu)/sigma)
F.rep[i] <- phi((log(t.rep.ranked[i]) - mu)/sigma)
diff.obs[i] <- pow(F.obs[i] - (2*i-1)/(2*N), 2)
diff.rep[i] <- pow(F.rep[i] - (2*i-1)/(2*N), 2)
}

CVM.obs <- sum(diff.obs[])
CVM.rep <- sum(diff.rep[])

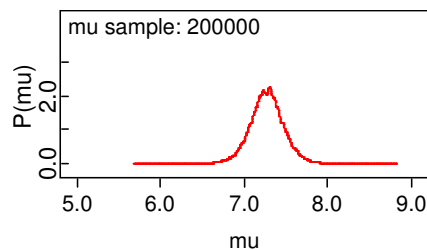
p.value <- step(CVM.rep - CVM.obs)
mu ~ dflat() #Diffuse priors for lognormal parameters
sigma ~ dgamma(0.0001, 0.0001)
tau <- pow(sigma, -2)
}
data
list(t=c(1258, 1388, 1022, 1989, 2024, 1638, 390, 4362, 2240, 1215, 2146, 655), N=12)

inits
list(mu=1, sigma=1)
list(mu=10, sigma=0.5)

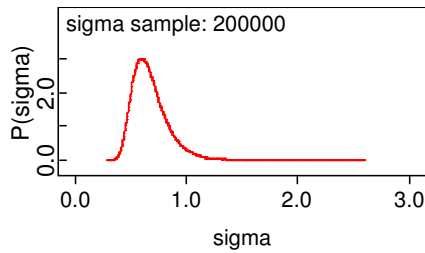
```

- Script 15. WinBUGS script for Bayesian inference of random durations using lognormal likelihood.

Running the script in the usual way produces the posterior distributions for μ and σ shown below. As stated above, it is not possible to interpret these distributions in terms of evidence for or against an exponential model. For reference the MLEs of μ and σ are 7.27 and 0.6, close to the posterior mean values of 7.27 and 0.67. The Bayesian p-value for the lognormal model is 0.50, again suggesting a good model for the observed times.



- Figure 57 Posterior distribution of μ in Example 14.



• Figure 58 Posterior distribution of σ in Example 14

5.6.2 Choosing Among Alternative Distributions – Deviance Information Criterion (DIC)

In the examples above we have developed four alternative aleatory models for the repair times given in Example 12: exponential, gamma, Weibull, and lognormal. By looking at the posterior distributions for the gamma and Weibull shape parameters, we have concluded that the exponential model may not be adequate, as it appears that the failure rate is an increasing function of operating time. How can we select among these models? WinBUGS provides a quantitative measure called the *deviance information criterion* or DIC that can be used to make this selection. The DIC measures how well a model replicates the observed data, but penalizes models with more parameters, to avoid over-fitting, which can produce a model that is poor at extrapolating beyond the observed data.

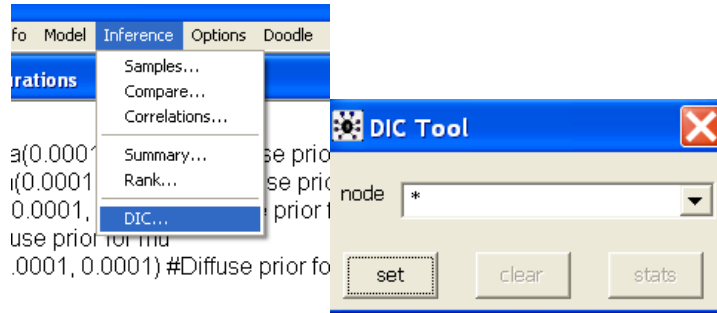
The model with the smallest DIC is the preferred model, although differences in DIC less than about five may not be significant. Note that DIC can be negative in some cases; the model with the smallest DIC is still preferred. As an example of this, if there were three models under consideration, with DICs of 10, -3, and -9, the model with a DIC of -9 would be the preferred model. Note that DIC is a measure of *relative* goodness of fit; the model with the smallest DIC may still not be an adequate model.

To avoid getting an invalid DIC, do not calculate DIC until parameter values have been estimated. This will ensure simulation has converged.

5.6.2.1 Calculating DIC in WinBUGS

First note that DIC must be calculated on the same data set. Also, DIC should not be estimated until the sampling has converged to the posterior distribution. Here are the steps for calculating DIC in WinBUGS:

1. Estimate parameter values in the usual way.
2. Set DIC as shown below
3. Run iterations to calculate DIC



Caution: the value for pD in the DIC results reported by WinBUGS must be positive in order for DIC to be used to compare models.

We illustrate by calculating DIC for each of the four models using the data from Example 12. The results are listed below. All three alternative models are significantly better than the exponential model, but among the three alternatives, all are equally good, and the Weibull model is probably the simplest to work with mathematically, because its cumulative distribution function can be written down in closed form. Of course, this is based on a small observed data set; with more data, DIC can more clearly distinguish among alternative models.

Exponential model:		Dbar	Dhat	DIC	pD
	time	203.5	202.4	204.5	1.029
	total	203.5	202.4	204.5	1.029
	Minimum deviance	202.4			
Gamma model:		Dbar	Dhat	DIC	pD
	time	198.0	195.9	200.1	2.088
	total	198.0	195.9	200.1	2.088
	Minimum deviance	195.9			
Weibull model:		Dbar	Dhat	DIC	pD
	time	198.7	196.7	200.7	2.023
	total	198.7	196.7	200.7	2.023
	Minimum deviance	196.6			
Lognormal model:		Dbar	Dhat	DIC	pD
	time	198.4	196.5	200.3	1.894
	total	198.4	196.5	200.3	1.894
	Minimum deviance	196.2			

5.7 Modeling Failure with Repair

Analyzing data from repairable systems can be complicated and may require expert consultation. Likewise, incorporating analysis results into a PSA may not be straightforward.

In earlier sections we analyzed times to occurrence of an event of interest. Such an event could be failure of a component or system. If the failure is not repaired, and the component or system is replaced following failure, then the earlier analysis methods are applicable. However, in this section, we consider the case in which the failed component or system is repaired and placed back into service. Analysis in this situation is a bit more complicated. The details will depend principally upon the nature of the system or component after repair. We will consider two cases:⁶

- Repair leaves the component or system the same as new,
- Repair leaves the component or system the same as old.

Both of these cases are modeling assumptions that an analyst must make, and they lead to different stochastic models for failure time. We will provide some qualitative guidance for when each assumption might be appropriate, and we will provide some qualitative and quantitative model checks that can be used to validate the assumption.

In all of the following discussion, we assume that we can ignore the time it takes to actually repair a component or system that has failed. This allows us to treat the failure process as a simple point process. The assumption is typically valid either because repair time is short with respect to operational time, or because we are only concerned with operational time, so time out for repair is accounted for through component or system maintenance unavailability estimates.

5.7.1 Repair Same As New – Renewal Process

In this case, repair leaves the failed component or system in the same state as a new component or system. The times between failures are thus independent and will be assumed to be identically distributed. If the times between failures are exponentially distributed, the methods of Sec. 3.3 can be applied. If the times between failures are not exponentially distributed (e.g., Weibull or lognormal), then the methods of Sec. 3.6 can be applied. Note in both cases that it is the *times between failures* that are analyzed, not the cumulative failure times.

Repair same as new may be plausible when a component is replaced with a new one or completely overhauled.

The assumption of repair same as new is plausible when the entire component or system is replaced or completely overhauled following failure. Examples would be replacement of a failed circuit board or rewinding a motor.

⁶Intermediate cases, in which repair leaves the component in a state in between new and old, can also be modeled, along with imperfect repair, which leaves a component worse than old. Such more general models are still an area of research and thus practical guidelines are difficult to give, so they are not included herein.

5.7.1.1 Qualitative Check for Time Dependence of Failure Rate in Renewal Process

If one assumes a renewal process, then a qualitative check on whether the failure rate is constant can be done using the times between failures. If the failure rate is constant in a renewal process, then the times between failures are exponentially distributed. If one plots the ranked times between failures on the x-axis, and $1/n_t$ on the y-axis, where n_t is the number of components still operating at time t , the result should be approximately a straight line if the failure rate is constant. If the slope is increasing (decreasing) with time, this suggests a renewal process whose failure rate is likewise increasing (decreasing) with time. Such a plot is referred to as a *cumulative hazard plot*.

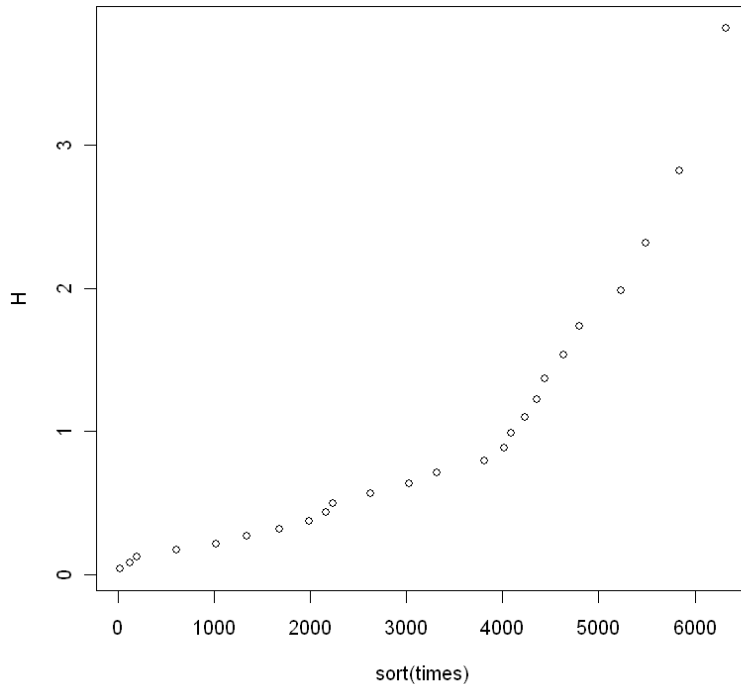
A cumulative hazard plot is only useful if repair is same as new, that is, failures are described by a renewal process.

• Example 15 Failure times from a renewal process

Consider the following 25 cumulative times of failure (in days) for a servo motor. Assume that the motor is replaced with a new one each time it fails, so the assumption of a renewal process seems reasonable. Use a cumulative hazard plot to decide if the failure rate appears to be increasing or decreasing with time.

Cumulative time (days)	Time between failures (days)	Sorted time between failures
127.4920	127.492	27.1964
154.6884	27.1964	74.28
330.4580	175.7696	76.22
739.9158	409.4578	87.037
1153.074	413.1582	123.803
1470.720	317.646	127.492
1809.616	338.896	141.964
2118.147	308.531	160.479
2289.570	171.423	171.423
2365.790	76.22	175.7696
2757.970	392.18	193.382
3154.409	396.439	201.649
3448.874	294.465	258.093
3941.777	492.903	294.465
4143.426	201.649	308.531
4217.706	74.28	317.646
4359.670	141.964	338.896
4483.473	123.803	352.491
4570.51	87.037	392.18
4763.892	193.382	396.439
4924.371	160.479	409.4578
5360.967	436.596	413.1582
5619.06	258.093	436.596
5971.551	352.491	476.511
6448.062	476.511	492.903

There are 25 times so the cumulative hazard plot increases by $1/25$ at 27.1964, by $1/24$ at 74.28, etc. The plot is shown below, and appears to indicate an increasing failure rate with time.



• Figure 59 Cumulative hazard plot for Example 15, suggesting increasing failure rate with operating time

Quantitative analysis of these times can be carried out using the methods in Secs. 3.3 and 3.6. Another quantitative check will be described below, after we have discussed the other extreme: repair same as old.

5.7.2 Repair Same As Old – Nonhomogeneous Poisson Process (NHPP)

In the case where repair only leaves the component in the condition it was in immediately preceding failure, then the times between failures may not be independent. For example, if the component is wearing out over time (aging), then later times between failures will tend to be shorter than earlier times, and conversely if the component is experiencing reliability growth. In these cases, the times between failures will also fail to meet the assumption of being identically distributed. Recall that to apply the methods of Secs 3.3 and 3.6, these two assumptions must be met.

If repair is not to the same as new, then the methods of Secs. 3.3 and 3.6 cannot be applied.

Repair to a state that is the same as old is a good default assumption for most components in a risk assessment, because a typical component is composed of subcomponents. When failure occurs, only a portion of the component (one or more subcomponents) is typically repaired, so the majority of the subcomponents are left in the condition they were in at the time of failure.

5.7.2.1 Qualitative Check for Trend in Rate of Occurrence of Failure When Repair is Same as Old

If the rate of occurrence of failure (ROCOF) is constant with time, then the times between failures will not tend to get shorter (aging) or longer (reliability growth) over time. If one plots the cumulative number of failures on the y-axis versus cumulative failure time on the x-axis, the resulting plot will be approximately a straight line if ROCOF is constant. If aging is occurring (increasing ROCOF), the slope will increase with time, as the times between failures get shorter. If reliability growth is occurring (decreasing ROCOF), the slope will decrease with time, as the times between failures get longer.

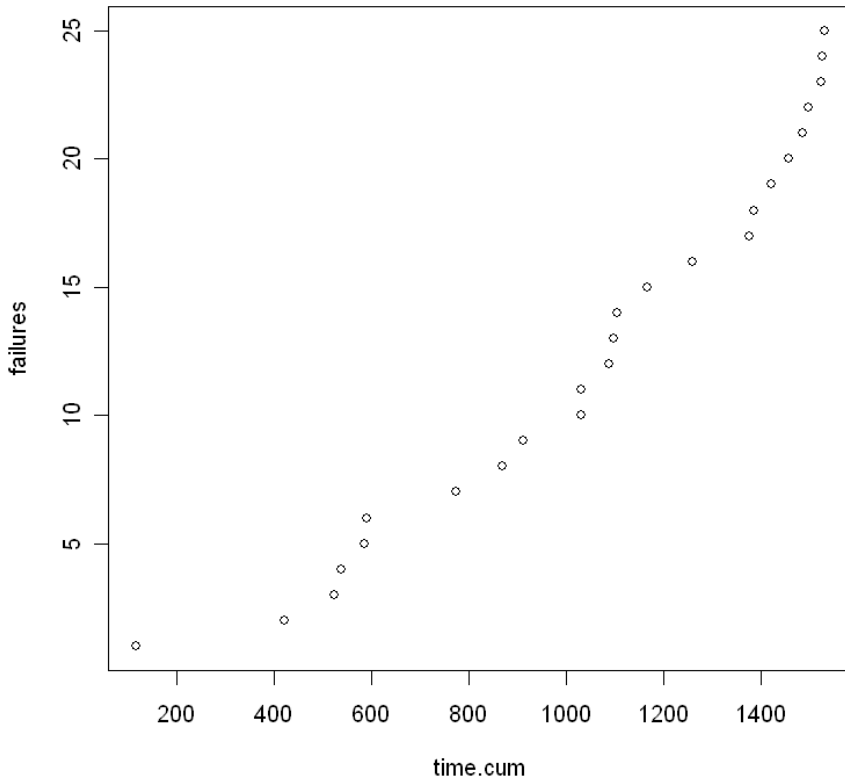
The cumulative failure plot is only useful if repair is same as old, that is, failures are not described by a renewal process.

• Example 16 Failure times from a non-renewal process

Consider the following 25 cumulative times in standby at which a cooling unit failed. Because the cooling unit consists of a large number of subcomponents, and only one or two of these were replaced at each failure, assume repair leaves the cooling unit in the state it was in immediately prior to failure. Plot cumulative number of failures versus cumulative failure time to check if there appears to be a time trend in the ROCOF for the cooling unit.

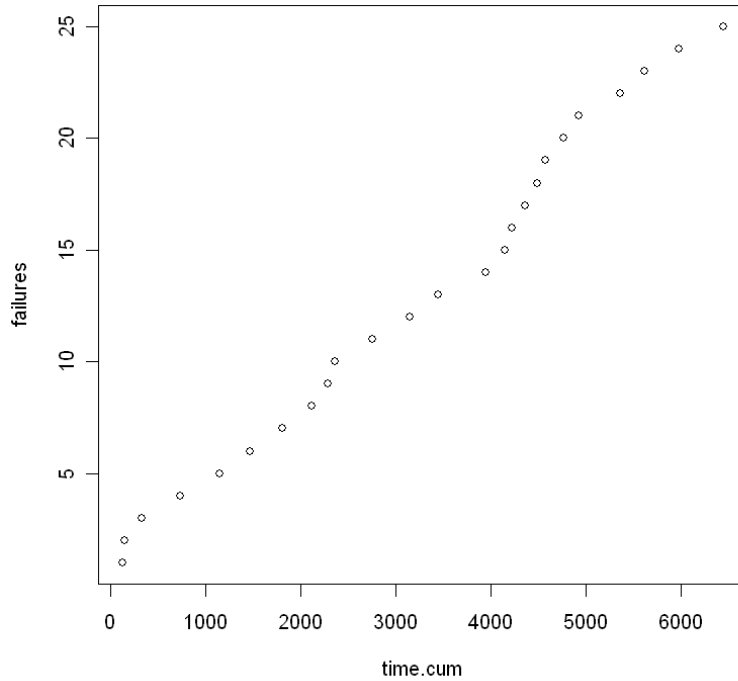
Cumulative time (days)
116.0454
420.8451
523.1398
538.3135
585.581
591.5301
772.365
868.7294
912.3777
1031.021
1031.133
1086.673
1096.476
1103.463
1165.640
1257.554
1375.917
1385.808
1421.459
1456.259
1484.755
1496.982
1523.915
1526.050
1530.836

The plot is shown below. The slope appears to be increasing with time, suggesting that the ROCOF for the cooling unit is increasing as a function of calendar time.



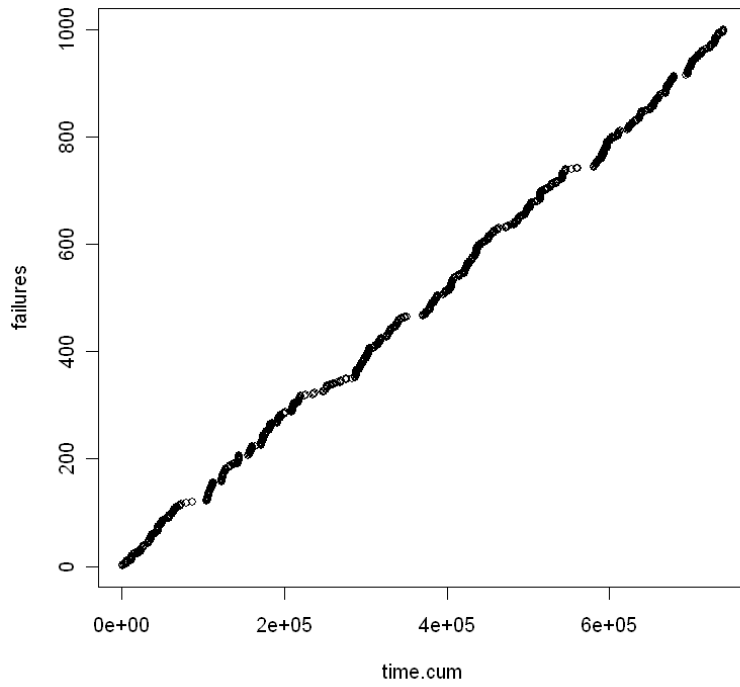
• Figure 60 Cumulative failure plot for Example 16, suggesting increasing ROCOF over time

An interesting exercise is to construct a cumulative failure plot for the data in Example 15, where the failure rate appeared to be an increasing function of operating time under the assumption that repair was same as new, a renewal process. The plot, shown below, does not suggest an increasing ROCOF under the assumption of repair same as old. This illustrates a subtle point in analyzing repairable systems. If repair is same as new after each failure, then times between failures will not exhibit a trend over calendar time. Aging or reliability growth only occurs over the time between one failure and the next, because the system returns to new, and the clock is reset, after each failure. On the other hand, when repair is same as old after each failure, then aging or reliability growth occurs over calendar time, and one can then expect to see a trend in the slope of cumulative failures versus time. Therefore, absence of a trend in the cumulative failure plot may suggest no aging or reliability growth under the assumption of repair same as old, but there still may be aging or reliability growth between each failure under the assumption of repair same as new. The cumulative hazard plot shown earlier can be used to check for that possibility.

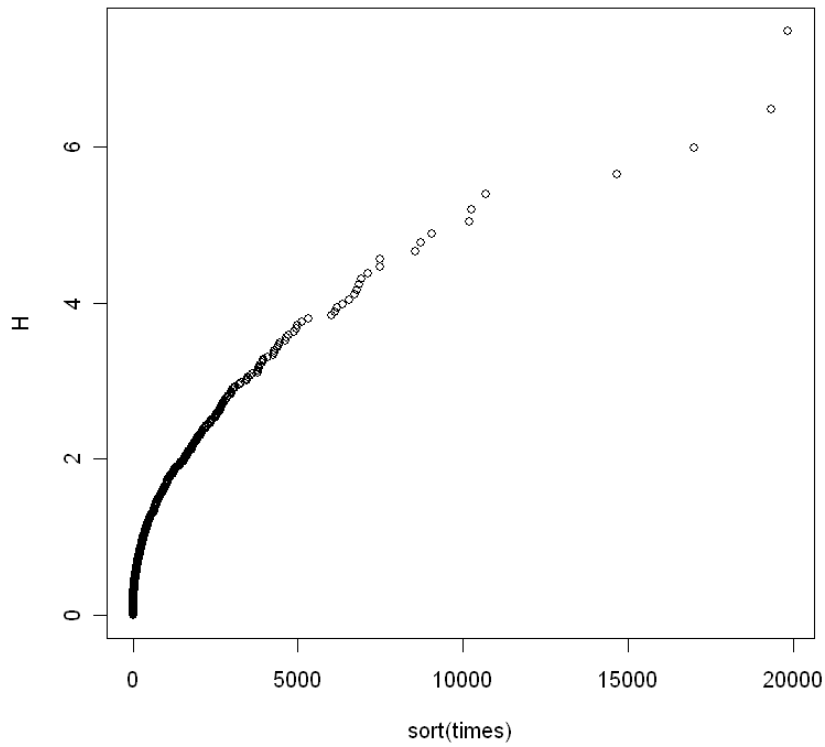


• Figure 61 Cumulative failure plot for data in Example 15, showing lack of trend in slope over calendar time

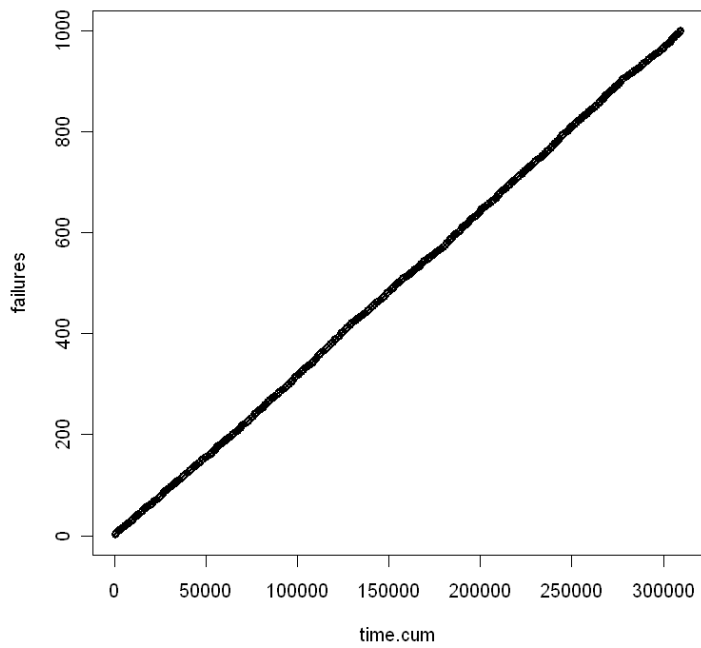
The plots below show cumulative failures versus cumulative time for 1,000 simulated failure times from two different renewal processes, one in which failure rate is decreasing with increasing operating time, the other where failure rate is increasing with operating time. Note in both cases that the cumulative failure plot produces a straight line, reinforcing the conclusion that this plot is useful for checking for aging or reliability growth under the same-as-old assumption for repair, but it cannot detect a time-dependent failure rate under the same-as-new repair assumption. The corresponding cumulative hazard plots shown below are useful for this purpose when repair is same as new.



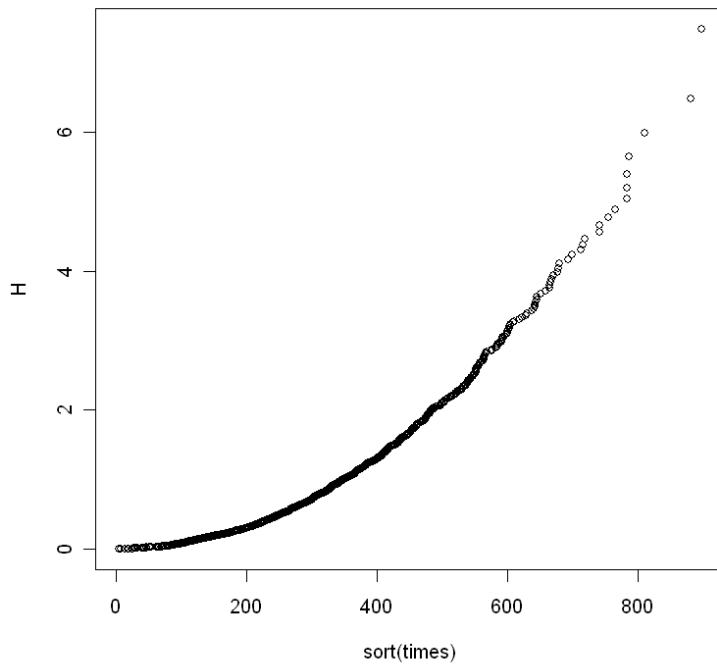
• Figure 62 Cumulative failure plot for 1,000 simulated failure times from renewal process with decreasing failure rate



• Figure 63 Cumulative hazard plot for 1,000 simulated failure times from renewal process with decreasing failure rate



• Figure 64 Cumulative failure plot for 1,000 simulated failure times from renewal process with increasing failure rate



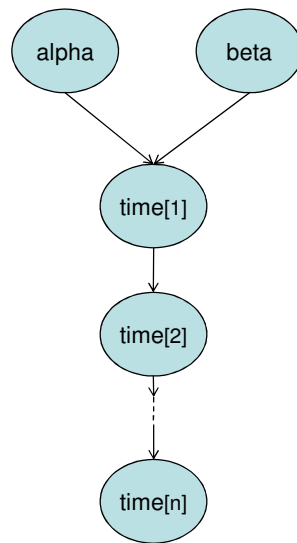
• Figure 65 Cumulative hazard plot for 1,000 simulated failure times from renewal process with increasing failure rate

5.7.2.2 Quantitative Analysis under Same-as-Old Repair Assumption

As stated above, if there is an increasing or decreasing trend in the ROCOF over time, then the times between failures will not be independently and identically distributed, and thus the methods of Secs. 4.3 and 4.6 cannot be applied. In particular, one cannot simply fit a Weibull, gamma, etc. distribution to the cumulative failure times or the times between failures. Instead, the likelihood function must be constructed using the fact that each failure time, after the first, is dependent upon the preceding failure time. One must also specify a functional form for the ROCOF. We will assume a power-law form for our analysis here. In this form, there are two parameters to estimate, which we denote as α and β . β determines how the ROCOF changes over time, and α sets the units with which time is measured. If β is less than one, reliability growth is occurring, if it is greater than one, aging is taking place, and if β equals one, there is no trend over time. The WinBUGS script shown below is used to estimate α and β . It also provides a Bayesian p-value to check the validity of the model.⁷ The underlying DAG model is shown in Figure 66.

The power-law process is also referred to in the reliability community as the Crow-AMSAA model.

⁷ In earlier sections, we used α for the shape parameter of the Weibull distribution. We use β in this section to conform to standard notation in the literature.



• Figure 66 DAG for modeling failure with repair under same-as-old repair assumption

Modeling NHPP (repair "as bad as old")

Using "zeros trick" to model NHPP

```

model {
for(i in 1:N) {
  zeros[i] <- 0
  zeros[i] ~ dpois(phi[i])
  #phi[i] = -log(likelihood)
}
#Power-law model (failure-truncated)
for(j in 1:N) {
  phi[j] <- -log(beta) + beta*log(alpha) - (beta-1)*log(t[j]) + pow(t[M]/alpha, beta)/N
}
#Model validation section
for(j in 1:N) {
  z.obs[j] <- pow(t[j]/alpha, beta)
}
z.inc.obs[1] <- z.obs[1]
for(k in 2:N) {
  z.inc.obs[k] <- z.obs[k] - z.obs[k-1]
}
for(j in 1:N) {
  z.inc.rep[j] ~ dexp(1)
  z.rep.ranked[j] <- ranked(z.inc.rep[], j)
  z.obs.ranked[j] <- ranked(z.inc.obs[], j)
  F.obs[j] <- 1 - exp(-z.obs.ranked[j])
  F.rep[j] <- 1 - exp(-z.rep.ranked[j])
  diff.obs[j] <- pow(F.obs[j] - (2*j-1)/(2*N), 2)
  diff.rep[j] <- pow(F.rep[j] - (2*j-1)/(2*N), 2)
}
CVM.obs <- sum(diff.obs[])
CVM.rep <- sum(diff.rep[])
p.value <- step(CVM.rep - CVM.obs) #Should be near 0.5
alpha ~ dgamma(0.0001, 0.0001)
beta ~ dgamma(0.0001, 0.0001)
}

```

- Script 16 WinBUGS script for analyzing data under same-as-old repair assumption (power-law process)

Let us apply this script to the data given in Example 16. Recall that Figure 60 suggested an increasing trend in ROCOF with time, corresponding to beta greater than one. We will run two chains, one starting with an initial value of beta less than one, the other with an initial beta greater than one. The initial values of the scale

parameter (α) are not crucial; we will vary α across the chains. Convergence to the joint posterior distribution appears to occur within the first 1,000 samples, so we discard the first 1,000 samples for burn in. We run another 10,000 samples to estimate parameter values, obtaining a posterior mean for β of 1.94 with a 90% credible interval of (1.345, 2.65). The posterior probability that β is greater than one is near unity, suggesting a ROCOF that is increasing with time, corresponding to aging. The Bayesian p-value is 0.57, suggesting a model that is good at replicating the observed data.⁸

The MLEs for α and β are surprisingly simple. The MLE for β is given by

$$\hat{\beta} = \frac{N}{\sum_{i=1}^N \log\left(\frac{t_i}{t_N}\right)} \quad (27)$$

The MLE is biased; an unbiased estimate is given by

$$\tilde{\beta} = \frac{N-2}{N} \hat{\beta} \quad (28)$$

Script 16 analyzes a process that is *failure-truncated*, that is, observation stops after the last failure. It is also possible to have a *time-truncated* process, in which observation continues after the last failure, up to time τ . To handle this case, only one line in Script 16 needs to be changed. The line defining the log-likelihood changes to the following:

```
for(j in 1:M) {
  phi[j] <- -log(beta) + beta*log(alpha) - (beta-1)*log(t[j]) + pow(tau/alpha, beta)/M
}
```

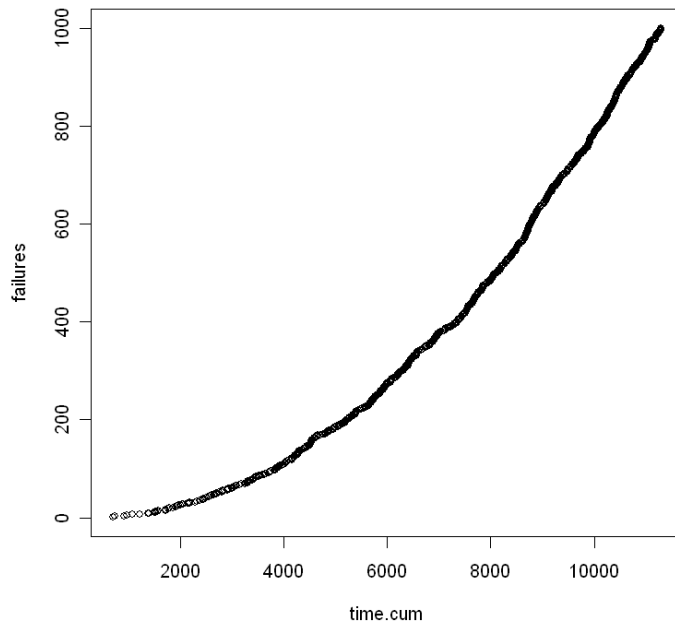
The stopping time (τ) is loaded as part of the observed data.

5.7.3 Impact of Assumption Regarding Repair

The assumption made regarding repair (same as old versus same as new) is crucial to the analysis. Consider times to failure being produced by a process with increasing ROCOF, corresponding to aging with repair same as old. As time progresses, times between failure will tend to decrease, and there will be a preponderance of short times between failures in a sample. If the process is assumed to be a renewal process, with times between failures described by a Weibull distribution, as in Sec. 4.6, the preponderance of short times between failures will cause the Weibull shape parameter to be less than one, corresponding to an apparent decreasing failure rate, the opposite of what is actually happening.

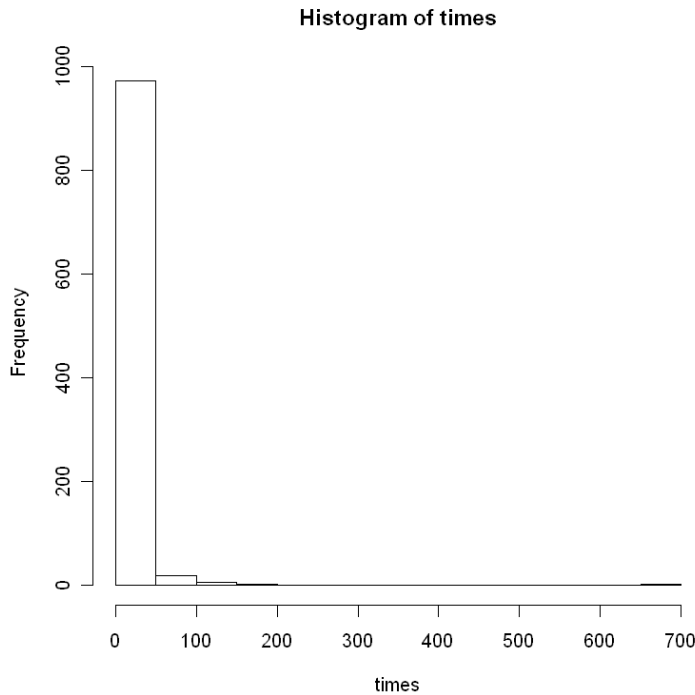
This can be illustrated by simulation. We generated 1,000 cumulative failure times for a system whose repair is same-as-old, described by a power-law process with shape parameter of 2 and scale parameter of 350. The cumulative failure plot below shows the increasing trend in ROCOF with time.

⁸ For reference, the data in Example 16 were simulated from a power-law process with $\beta = 2$ and $\alpha = 350$.



• Figure 67 Cumulative failure plot for 1,000 times simulated from power-law process with shape parameter of 2, illustrating increasing ROCOF

The histogram below of the times between failures shows the preponderance of short times between failures caused by the increasing ROCOF.



• Figure 68 Histogram of times between failures for simulated failure times from power-law process with increasing ROCOF

Assuming the repair is same-as-new and fitting a Weibull distribution to these times between failures using the techniques of Sec. 4.6, one estimates a Weibull shape parameter of about 0.8, which would suggest a failure rate that is decreasing with time. This result is caused by the fact that times between failures are tending to become shorter due to aging; treating the times between failures as independent and identically distributed leads to an erroneous conclusion about the process. Unfortunately, the Bayesian p-value may not help much in deciding which model is better, because both models can replicate the observed data quite well.

If the repair is same-as-new, and the failure rate increases with operating time or time in standby (whichever is being modeled), an assumption of same-as-old repair will, as suggested by Figure 64, lead to an estimate near one for the shape parameter of the power-law process. In this case, the Bayesian p-value can be helpful, as a power-law process with shape parameter near one cannot replicate data from a renewal process with increasing failure rate very well.

There is a non-parametric statistical test that can be helpful in distinguishing between a renewal process (same-as-new repair) and an NHPP (same-as-old repair). It is based upon the Laplace test, and it is simple to implement in a spreadsheet or other software. It uses the *cumulative* times to failure. The formula depends on whether the observation period is for a fixed period of time, or terminates at the time of the last failure. If it is for a fixed period of time, τ , and there are n observed failure times, one first computes the Laplace statistic, U , according to the following formula:

$$U = \frac{\sum_{i=1}^n t_i / \tau - n/2}{\sqrt{n/12}} \quad (29)$$

If the observation period is only up until the last observed cumulative failure time, t_n , U is computed by

$$U = \frac{\sum_{i=1}^{n-1} t_i / t_n - (n-1) / 2}{\sqrt{\frac{n-1}{12}}} \quad (30)$$

Next calculate the sample mean and standard deviation, using all n times if the observation was for a fixed period of time, and the first n-1 times if observation was terminated at the time of the last failure. Then take the ratio of the sample standard deviation to the sample mean (i.e., the sample coefficient of variation), and divide U by this ratio. This statistic quickly approaches a standard normal distribution under the null hypothesis of a renewal process. If the process is actually NHPP with increasing ROCOF, too many of the failures will occur after the midpoint of the observation period, and U will be too large. Conversely, if the process is NHPP with decreasing ROCOF, too many failures will occur before the midpoint of the observation period, and U will be too small. At a 5% significance level, we reject the null hypothesis of a renewal process if U is larger than 2 or smaller than -2.

For the data of Example 15, assuming that observation stopped at the time of the last failure, we find a value of U of -0.25, so we cannot reject a renewal process as describing these failures. For the data of Example 16, again assuming observation ended with the last failure, we calculate a value for U of 2.21, leading to rejection of the renewal process in favor of the alternative NHPP. Because U is positive, this suggests an NHPP with increasing ROCOF, as we found when we did quantitative analysis for this example using WinBUGS above.

5.7.4 Incorporating Results into PSA

How quantitative results are incorporated depends again on the process assumed to describe the failures. If the repair is assumed to be same-as-new, leading to a renewal process, then the stochastic model for failure time in the PSA is the renewal distribution (e.g., Weibull). Most PSA software packages can only model an exponential distribution for failure time, so this would appear to be a problem. A way around this problem for operating equipment with a Weibull renewal distribution is as follows. Assume the mission time for the operating equipment is t_m , and assume we have determined that failures of the operating equipment can be described by a renewal process with a Weibull distribution for times between failures, with shape beta and scale lambda. That is, the density function for times between failures is given by

$$f(t) = \beta \lambda t^{\beta-1} \exp(-\lambda t^\beta) \quad (31)$$

This is the stochastic model for failure, which most PSA software does not include. The event of interest is the probability that the equipment fails before the end of the mission time. To have the PSA software calculate this probability, one can input an exponential distribution with rate λ as the stochastic model in the PSA, but replace the mission time by $(t_m)^\beta$. This is adequate for a point estimate, but may not allow epistemic uncertainty in β and λ to be propagated through the PSA model. If the renewal distribution is gamma or lognormal, there is no such simple work-around. If DIC does not show a clear preference for one model over the other (i.e., lognormal or gamma clearly better than Weibull), it is best to use a Weibull renewal distribution, as it will be easiest to incorporate into the PSA.

If the failures are assumed to be described by an NHPP (repair same as old), the process for incorporating the results into the PSA is a bit trickier. If the component at hand is new, and its failures are assumed to be described by an NHPP with parameters estimated from past data, then the time to first failure of the component is of interest, and this time will be Weibull-distributed with shape parameter β and scale parameter λ , where λ is given by $\alpha^{-\beta}$. See above for how to “trick” the PSA software into using a Weibull stochastic model for time to first failure.

If the parameters have been estimated from a component's failure history, and this history continues into the future, then one is interested in the probability that the next failure will occur before the end of the mission. The distribution of the next cumulative failure time, t_i , is *not* a simple Weibull distribution. WinBUGS can be used to predict the next cumulative failure time and calculate the probability that $T_i \leq t_{i-1} + t_m$. The script below shows how this is implemented.

```

model {
  for(i in 1:N) {
    zeros[i] <- 0
    zeros[i] ~ dpois(phi[i])
    #phi[i] = -log(likelihood)
  }
  #Power-law model (failure-truncated)
  phi[1] <- -log(beta) + beta*log(alpha) - (beta-1)*log(t[1]) + pow(t[1]/alpha, beta)
  t[N] ~ dgamma(0.0001, 0.0001)|t[M-1,] #Monitor node t[M] for predicted time of next failure
  for(j in 2:N) {
    phi[j] <- -log(beta) + beta*log(alpha) - (beta-1)*log(t[j]) + pow(t[j]/alpha, beta) - pow(t[j-1]/alpha,
beta)
  }
  t.miss <- 24 #PSA mission time
  t.window <- t[N-1] + t.miss
  prob.fail <- step(t.window - t[N]) #Monitor mean of this node
  alpha ~ dgamma(0.0001, 0.0001)
  beta ~ dgamma(0.0001, 0.0001)
}

```

- Script 17 WinBUGS script for predicting next failure time in NHPP and estimating failure probability

The data loaded has an extra entry that is coded as NA, and initial values must be given for each chain. Running this script with the data from Example 16, we find the following results for the predicted cumulative time for the next failure.

Node statistics								
	mean	sd	MC_error	val5.0pc	median	val95.0pc	start	sample
t[26]	1563.0	34.63	0.3318	1532.0	1553.0	1632.0	1001	40000

The probability of failure in the 24-hour PSA mission time is estimated to be 0.54.

It is also possible to use the cumulative distribution function directly, as this is a left-truncated Weibull distribution. The probability that the next failure will occur before (cumulative) time t , given that the last failure occurred at (cumulative) time T , is given by

$$F(t) = 1 - \exp\{-\lambda[(T+t)^\beta - T^\beta]\} \quad (32)$$

where $\lambda = \alpha^\beta$. This has the advantage of giving an uncertainty distribution for the probability of failure. The WinBUGS script below encodes this approach. Running the script in the usual way, we find a mean probability of failure of 0.53, with a 90% interval of (0.36, 0.69).

```

model {
  for(i in 1:M) {
    zeros[i] <- 0
    zeros[i] ~ dpois(phi[i])
    #phi[i] = -log(likelihood)
  }
  #Power-law model (failure-truncated)
  phi[1] <- -log(beta) + beta*log(alpha) - (beta-1)*log(t[1]) + pow(t[1]/alpha, beta)
  for(j in 2:M) {
    phi[j] <- -log(beta) + beta*log(alpha) - (beta-1)*log(t[j]) + pow(t[j]/alpha, beta) - pow(t[j-1]/alpha,
beta)
  }
  t.miss <- 24
  t.window <- t[M] + t.miss
  lambda <- pow(alpha, -beta)
  prob.fail <- 1 - exp(-lambda*(pow(t.window, beta) - pow(t[M], beta)))
  alpha ~ dgamma(0.0001, 0.0001)
  beta ~ dgamma(0.0001, 0.0001)
}

```

- Script 18 WinBUGS script for estimating failure probability for NHPP using cumulative distribution function

One item to note is that the parameters of the power-law process, α and β , are highly correlated; the rank correlation between them is about 0.95 for the analysis above. This correlation must be taken into account in the calculation of the failure probability using Eq. (32) failure to take it into account leads to an over-estimate of the uncertainty. With no correlation between α and β , the 90% interval for the failure probability is estimated to be (0.16, 0.99), much wider than the interval obtained using Script 18, which takes the correlation into account automatically.

5.8 Treatment of Censoring

So far we have only analyzed cases where the observed data were known with certainty and completeness. Reality is messier in a number of ways with respect to observed data. When observing times at which failures occur (i.e., random durations), various types of *censoring* can occur. For example, a number of components may be placed in test, but the test is terminated before all the components have failed. This produces a set of observed data consisting of the recorded failure times for those components that have failed. For the components that did not fail before the test was terminated, all we know is that the observed failure time was longer than the duration of the test. As another example, in recording failure times, the exact time of failure may not be known; in some cases, all that may be available is an interval estimate (e.g., between 10 and 20 hours).

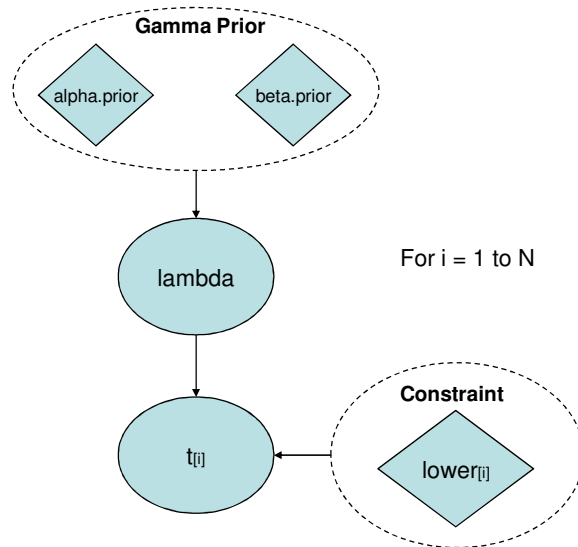
5.8.1 Censored Data for Random Durations

Consider the following example. A programmable logic controller (PLC) is being tested. Ten PLCs are placed in test, and each test is to be run for 1,000 hours. If a PLC fails before the end of the test, its failure time is recorded. Assume that two of the PLCs failed during the test, at 395 and 982 hours. The other eight PLCs were still operating when the test was terminated at 1,000 hours (this is referred to in the literature as *Type I censoring*). How can we use this information to carry out Bayesian inference for the PLC failure rate, λ , assuming that the time to failure can be described by an exponential distribution?

Frequentist estimation for censored data can be quite complicated, especially for nonexponential renewal distributions. Therefore, we cover only the Bayesian approach in this guideline. Because of WinBUGS, the Bayesian approach is very straightforward and only slightly more complicated than the complete-sample cases considered earlier. For more details on frequentist estimation, we refer the reader to (Bain and Engelhardt, 1991).

The key to dealing with censored times, for any likelihood distribution, is the **C(lower, upper)** construct in WinBUGS. In cases where a failure time was not recorded, this tells WinBUGS to impute a failure time from the specified distribution, between the bounds specified by lower and upper in the construct. For our example above, lower would be equal to 1,000 hours, and upper would be omitted, so we would have the following line in the WinBUGS script: `t[i] ~ dexp(lambda)C(1000,)`.

To incorporate censored data into WinBUGS, in the data portion of the script, a value of NA is entered for each censored failure time (time not recorded). Because “lower” is a vector, the failure time is entered if a failure time is observed, otherwise NA is entered. The script to analyze this example is shown below. The underlying DAG is shown in Figure 69.



• Figure 69 DAG model for Type I censoring in the exponential distribution

```

model {
  for(i in 1:N) {
    #Define likelihood function with Type I censoring
    t[i] ~ dexp(lambda)C(lower[i],)
  }
  lambda ~ dgamma(0.0001, 0.0001)      #Jeffreys prior for lambda
}

data
list(t=c(982,394.7,NA,NA,NA,NA,NA,NA,NA,NA),
lower=c(982,394.7,1000,1000,1000,1000,1000,1000,1000,1000), N=10)

inits
list(lambda=0.001)
list(lambda=0.0001)

```

- Script 19 WinBUGS script for modeling censored random durations.(Type I censoring)

Running this script in the usual way gives a posterior mean for lambda of 2.1×10^{-4} /hour and a 90% credible interval of $(3.8 \times 10^{-5}, 5.1 \times 10^{-4})$.

Another way in which censoring could occur is when observation is terminated after a certain number of components have failed, thus failure times are not observed for the remaining components. In this case, the total observation period is random, and the censoring is referred to as *Type II censoring*.

- Example 17 Modeling failure times with Type II censoring

The following example is taken from (Bain and Engelhardt, 1991). We have observed 30 components and recorded the first 20 times to failure, in days, as follows: 1, 3, 5, 7, 11, 11, 11, 12, 14, 14, 14, 16, 16, 20, 21, 23, 42, 47, 52, 62. Assuming these failures are exponentially distributed, find the mean and posterior 90% interval for the failure rate.

Solution

We have observed 20 times of failure, but there were 30 components under observation. It is a mistake to use only the observed failure times to estimate λ , because we know that 10 of the components survived longer than 62 days, and this is substantial information about λ . The WinBUGS script below shows how to incorporate this information using the I(.) construct, with a Jeffreys prior distribution for λ . The DAG is similar to the one shown in Figure 69.

- Script 20 WinBUGS script for Example 18 – Type II censoring of failure times

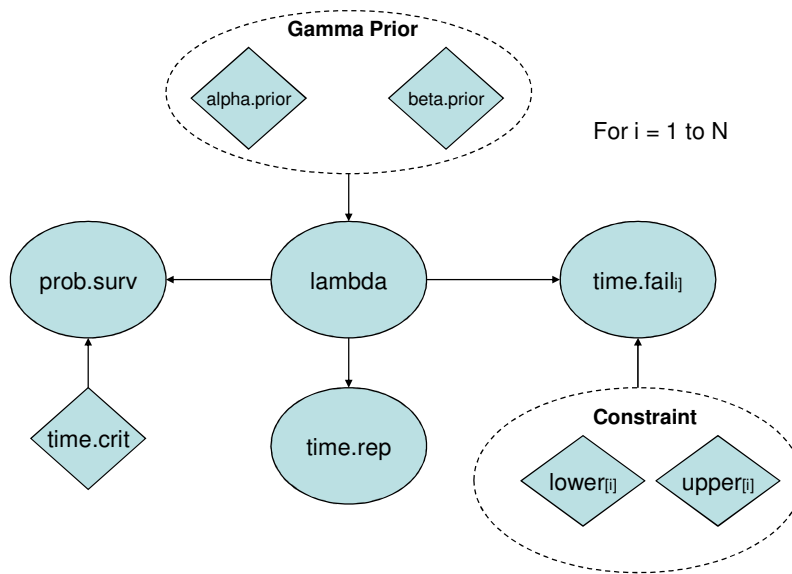
Running in this script in the usual way we find a posterior mean for λ of 0.02/day and a 90% credible interval of (0.01, 0.03). Had we ignored the 10 components that survived longer than 62 days, that is, had we estimated λ treating the 20 observed failure times as a complete sample, we would have found a posterior mean of 0.05, with a 90% interval of (0.03, 0.07), which could be a significant over-estimate.

- Example 18 Modeling interval-censored failure times

Consider the following data (in days) on time between failures: < 1, 5, < 10, 15, 4, 20, 30, 3, 30-60, 25. Find the probability that a component will operate longer than 20 days.

Solution

We start by specifying a stochastic model for the failure. The simplest model is the exponential distribution, so we will start with that, and examine alternative models later. We will use a Jeffreys prior distribution for λ , the failure rate in the exponential distribution. The WinBUGS script below illustrates how to encode the censored data for failure time. The underlying DAG model is shown in Figure 70 . Node time.rep estimates the failure time, unconditional upon λ , and node prob.surv estimates the probability that the component will operate for at least 20 days.



• Figure 70 DAG model for interval censoring in the exponential distribution

```

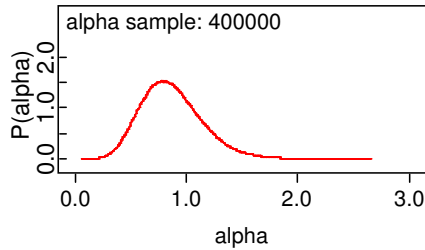
model {
  for (i in 1 : N) {
    time.fail[i] ~ dexp(lambda)C(lower[i], upper[i]) #Exponential distribution for failure times
  }
  time.rep ~ dexp(lambda)
  lambda ~ dgamma(0.0001,0.0001) #Diffuse prior for exponential parameter
  prob.surv <- exp(-lambda*time.crit) #Probability of survival beyond time.crit
  time.crit <- 20
}
data
list(time.fail=c(NA, 5, NA, 15, 4, NA, 3, NA, 25), N=9)
list(lower=c(0, 5, 0, 15, 4, 20, 3, 30, 25), upper=c(1, 5, 10, 15, 4, 30, 3, 60, 25))
inits
list(lambda = 0.1) #initial values
    
```

• Script 21 WinBUGS script for Example 18, interval-censored failure times with exponential likelihood.

Running this script in the usual way gives a mean survival probability at 20 days of 0.25, with a 90% interval of 0.09, 0.46). For later comparison with an alternative model, the DIC was found to be 51.3.

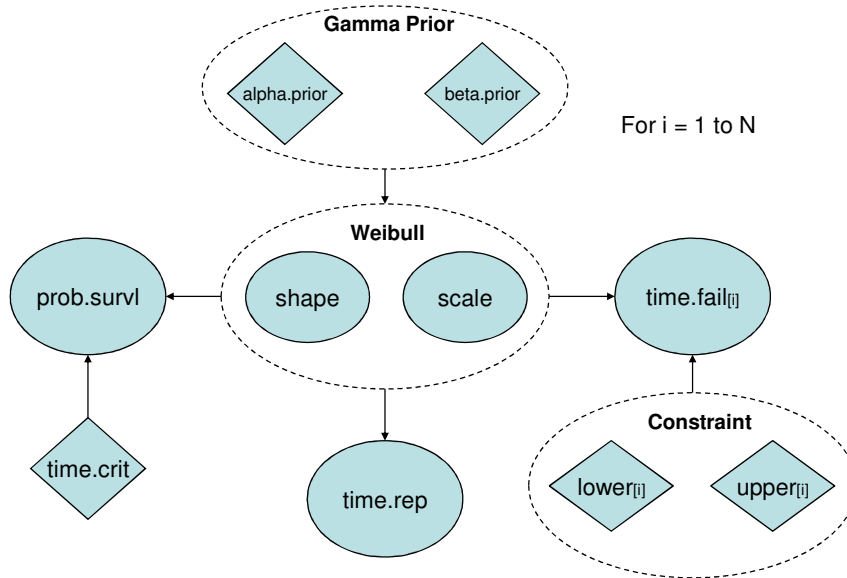
The Weibull distribution is an alternative stochastic model to the exponential distribution, as described in Sec. 5.6.1.2. The script below shows how to encode a Weibull likelihood function and estimate the probability of

survival beyond 20 days. The underlying DAG model is shown in Figure 72. Recall that if the shape parameter, α , equals one, then the Weibull distribution reduces to the exponential distribution. Running the script below in the usual way, gives the following posterior distribution for the Weibull shape parameter.



- Figure 71. Posterior distribution for shape parameter in Weibull distribution.

As this distribution is centered about one, the exponential distribution is probably adequate. The probability of survival with the Weibull model has a posterior mean of 0.26 and a 90% credible interval of (0.09, 0.47). These are essentially the same results as obtained with the simpler exponential model above. Finally, the DIC for the Weibull model is 52.47, slightly larger than the exponential model, allowing us to conclude that the exponential is a better model than the Weibull distribution.



• Figure 72 DAG model for interval-censoring in the Weibull distribution

```

model
{
  for (i in 1 : N) {
    time.fail[i] ~ dweib(alpha, scale)C(lower[i], upper[i])    #Weibull distribution for failure times
  }
  time.rep ~ dweib(alpha, scale)
  alpha ~ dgamma(0.0001,0.0001)                               #Diffuse priors for Weibull parameters
  scale ~ dgamma(0.0001,0.0001)
  prob.surv <- exp(-scale*pow(time.crit, alpha))               #Probability of survival beyond time.crit
  time.crit <- 20
}
data
list(time.fail=c(NA, 5, NA, 15, 4, NA, 3, NA, 25), N=9)
list(lower=c(0, 5, 0, 15, 4, 20, 3, 30, 25), upper=c(1, 5, 10, 15, 4, 30, 3, 60, 25))
inits
list(alpha = 1, scale=10) #initial values
list(alpha=0.5, scale=15)

```

- Script 22. WinBUGS script for Example 18, interval-censored failure times with Weibull likelihood.

5.8.2 Left-Censoring of Failure Times for NHPP

In this case, a component has been in operation since its commissioning date, denoted as time zero. However, times of failures, if they occurred, were not recorded prior to a certain date. Thus, failures may have occurred before recording started, but these times of failure are unknown. This will be referred to as *censoring on the left*. This is a straightforward situation to model in WinBUGS. The number of failures prior to the start of recording is modeled with a Poisson distribution (possibly with a time-dependent rate), and failures are imputed and combined with the usual likelihood function for the recorded times.

- Example 19 Example 16 revisited with left censoring at 100 days

Consider again the data from Example 16, which were cumulative times of failure for a component where repair was same-as-old, that is a non-renewal process. In Example 16, we used a power-law process as our stochastic model. Let us reconsider this model, under the condition that no failure times were collected prior to 100 days. That is, failures may have occurred during the first 100 days, but there was no record kept of whether a failure occurred or its time of occurrence. How does this affect our estimate of the shape parameter?

The WinBUGS script below is essentially the same as was used for Example 16. We have inserted a line that models failures prior to 100 days as a Poisson distribution, with time-dependent intensity. Note that x (the variable representing failures before 100 days) is given a value of "NA" in the data statement to indicate that x has not been observed (the value is missing). Running this script in the usual way, we find a posterior mean for β of 1.95, with a 90% credible interval of (1.33, 2.66). The results are very close to those obtained under the assumption of no censoring, so the effect of the censoring was minor in this example.

```

model {
for(i in 1:N) {
    zeros[i] <- 0
    zeros[i] ~ dpois(phi[i])
    #phi[i] = -log(likelihood)
}
#Treat left-censoring as missing data
x ~ dpois(mu)
mu <- pow(t.cens/alpha, beta)
#Power-law model (failure-truncated)
    for(j in 1:N) {
        phi[j] <- -log(beta) + beta*log(alpha) - (beta-1)*log(t[j]) + pow(t[M]/alpha, beta)/N
    }
alpha ~ dgamma(0.0001, 0.0001)
beta ~ dgamma(0.0001, 0.0001)
}

data
list(x=NA, t.cens=100)

```

• Script 23 WinBUGS script to analyze data from NHPP (repair same as old) with left censoring

5.9 Case Studies

In this section we describe several case studies that have been carried out by one or more groups of analysts.

The first case study analyzed virtual failure data similar to those from operating French and German reactors. The data were binned Poisson and binomial data obtained from a population of nominally similar components. Thus, “aging” would constitute an increasing trend in either failure rate (λ) or failure probability (p) across the bins. The original analysis of this data was frequentist in nature, and the approach used to estimate statistical model parameters was minimization of Fisher’s asymptotic chi-square statistic. This is an uncommonly used approach to parameter estimation in comparison with either maximum likelihood estimation (MLE, frequentist) or Bayesian approaches. It can, but does not necessarily lead to identical estimates as either MLE or Bayesian estimation with diffuse priors. Estimates from this approach can be larger or smaller than the estimates from the more standard approaches in some cases. In the case of a multivariate Gaussian likelihood, maximizing the likelihood is equivalent to minimizing the chi-square statistic.

We analyzed three representative data sets from this cast study: Cases #3 and 7.1 (binned Poisson data) and Case #DEF (binned binomial data). The results of the calculations are described below.

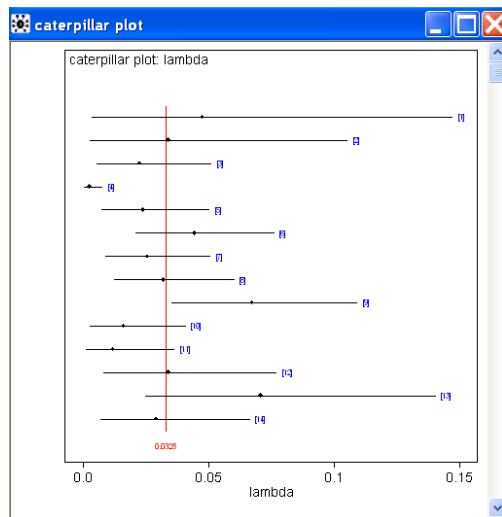
5.9.1 Case #3

The data analyzed are shown in the table below.

• Table 13 Data for Case #3

T	X	Exp. Time
0.5	1	31.64
1.5	1	44.135
3	3	156.407
6	1	608.38
8.5	4	189.545
9.5	9	214.63
10.5	5	216.355
11.5	6	204.091
12.5	12	185.926
13.5	2	157.536
14.5	1	127.608
15.5	3	104.105
16.5	5	77.86
20	3	121.167

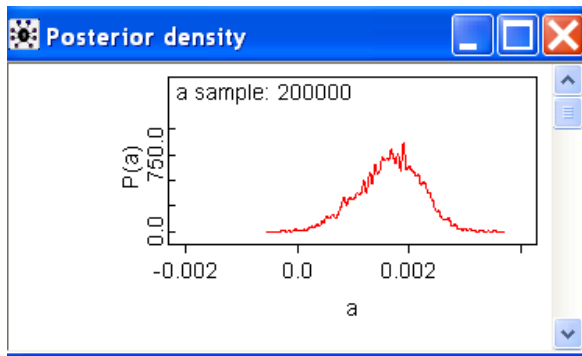
We first developed a caterpillar plot of side-by-side posterior 95% credible intervals for each bin, using WinBUGS with a Jeffreys prior in each bin. These are the Bayesian analog of 95% confidence intervals. The plot, shown below, does not suggest a trend over time, but does suggest significant bin-to-bin variability.



• Figure 73 Plot of 95% credible intervals shows no significant time trend in λ

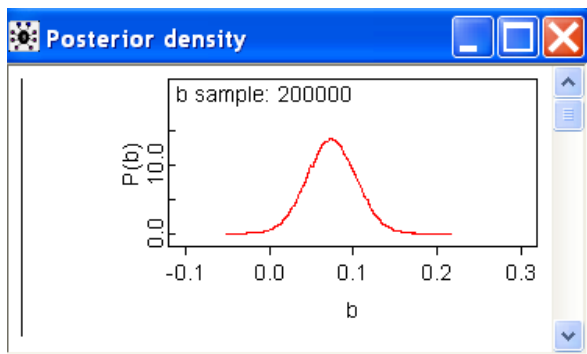
We next performed a Bayesian analysis of a model with no time trend (constant λ), again using a Jeffreys prior for λ . The resulting posterior mean was 0.023, with a 90% interval of (0.018, 0.028). To check the validity of this model, we used the posterior predictive distribution for the number of failures in each bin to compare observed and replicated chi-square statistics. The Bayesian p-value was only 0.004, indicating a poorly fitting model, in the sense that a model with constant λ is highly unlikely to be able to reproduce the variability in the observed data. Note that the estimate of λ in the original analysis of this data was 0.0297, and the p-value (from Fisher's asymptotic chi-square test) was 0.002. The MLE for the pooled data is 0.0197, with an (at least) 90% confidence interval of (0.015, 0.025).

We next examined a linear model for λ over time: $\lambda(t) = \lambda_0 + at$. Laplace priors over the real axis were used for both λ_0 and a , obtaining the following results. For λ_0 , the posterior mean was found to be 0.007, with a 90% credible interval of $(5.3 \times 10^{-4}, 0.017)$. This compares with the point estimate of 0.012 found in the original analysis. The posterior mean of “ a ” was 0.002, with a 90% credible interval of $(6.7 \times 10^{-4}, 0.002)$. The probability that “ a ” is positive, indicating an increasing trend in λ over time, is greater than 0.95. The marginal posterior density of “ a ” is shown below. The point estimate of “ a ” in the original analysis was 0.002. The Bayesian p-value was 0.01, compared with the frequentist p-value of 0.006. These results are again indicative of a poorly fitting model.



• Figure 74 Marginal posterior density for aging coefficient in linear model

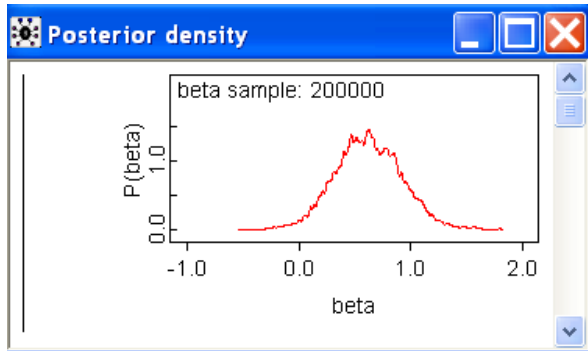
The next model examined was a loglinear model: $\log(\lambda) = a + bt$. Note that this is referred to as an exponential model in the original analysis, and the intercept parameter in that analysis is e^a . As in the linear model, Laplace priors over the real axis were used for a and b , giving the following results. The posterior mean of “ a ” was -4.579, with a 90% interval of $(-5.182, -4.005)$. This corresponds to an exponentiated value of 0.010, which can be compared with the point estimate of 0.015 in the original analysis. The posterior mean of b was 0.07, compared with a point estimate of 0.06 in the original analysis. The 90% credible interval for b was (0.122) . The marginal posterior density of b is shown below. The Bayesian p-value for the loglinear model was 0.01, compared with the frequentist p-value of 0.0055. These results are again indicative of a poorly fitting model.



• Figure 75 Marginal posterior density for aging coefficient in loglinear (exponential) model

Finally, a power-law model was examined, in which $\lambda(t) = \lambda_0 t^\beta$. Note that this is referred to as a Weibull model in the original analysis. With diffuse priors on λ_0 and β over the positive real axis, the following results were obtained with WinBUGS. The posterior mean and 90% credible interval for λ_0 were 0.007 and $(1.46 \times 10^{-4}, 0.017)$. The point estimate in the original analysis was 0.013. The posterior mean and 90% credible interval for β were 0.62 and $(0.16, 1.11)$, compared with a point estimate of 0.35 in the original analysis. The marginal

posterior density of β is shown below. Because of the rather large discrepancy between the posterior mean and the point estimate in the original analysis, the MLE for β was obtained. The MLE was 0.56, suggesting that this is one model for which the MLE and minimum chi-square estimates may differ significantly. The Bayesian p-value for the power-law model was 0.007, compared with the frequentist p-value of 0.003. Again, this is not a valid model.



• Figure 76 Marginal posterior density for shape parameter in power-law model

None of the potential trend models examined is very good at replicating the observed data, in agreement with the qualitative insight from Figure 73 that there is no apparent trend in λ over time. Because Figure 73 does suggest significant bin-to-bin variability in λ , a hierarchical Bayes model was used to develop a distribution that describes this variability. The functional form of the distribution was taken to be a gamma distribution, and diffuse hyperpriors were placed on the parameters of this gamma distribution. WinBUGS was used to carry out the analysis, giving a posterior mean of 0.028, with a 90% credible interval of (0.005, 0.07), a significantly wider interval than that obtained by pooling the data. The Bayesian p-value for the hierarchical model was 0.52, strong evidence of bin-to-bin variability in the observed data.

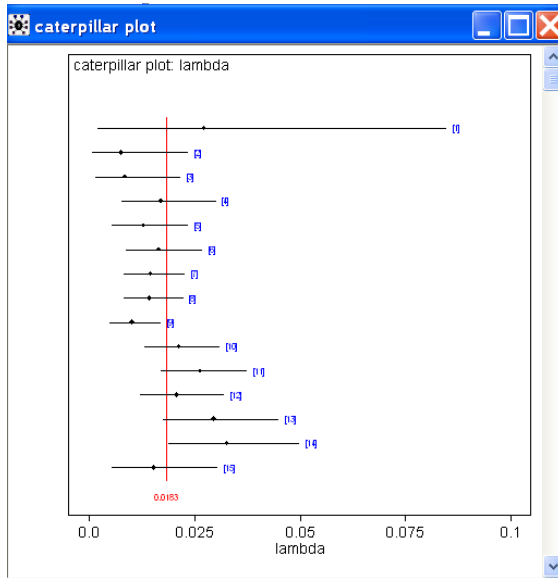
5.9.2 Case #7.1

The same analyses as described above were carried out for the data of Case 7.1. The data for this case are shown in the table below.

• Table 14 Data for Case 7.1

T	X	Exp. Time
2	1	54.873
4	1	198.003
5.5	2	295.425
6.5	8	500.097
7.5	7	585
8.5	12	759.681
9.5	15	1062.44
10.5	15	1082.37
11.5	10	1046.68
12.5	21	1014
13.5	25	971.88
14.5	16	796.419
15.5	17	591.708
16.5	16	507
18	5	361.374

The caterpillar plot below suggests a possible increasing trend in λ over time.



• Figure 77 Plot of 95% credible intervals suggests possible increasing time trend in λ

The table below shows the results of the quantitative analyses.

• Table 15 Results for Case 7.1

Model	Point Est.	Post. Mean	90% int.	Bayesian p-value	Frequentist p-value
Constant	0.019	0.017	(0.015, 0.020)	0.046	0.04
Linear	a: 0.004 b: 0.001	a: 0.004 b: 0.001	a: (4.1E-5, 0.01) b: (7.0E-4, 0.002)	0.41	0.43
Loglinear	a: 0.007 b: 0.079	a: 0.007 b: 0.079	a: (0.004, 0.01) b: (0.04, 0.12)	0.47	0.49
Power-law	λ_0: 0.003 β: 0.725	λ_0: 0.003 β: 0.814	λ_0: (7.8E-4, 0.007) β: (0.41, 1.27)	0.33	0.365

5.9.3 Case DEF

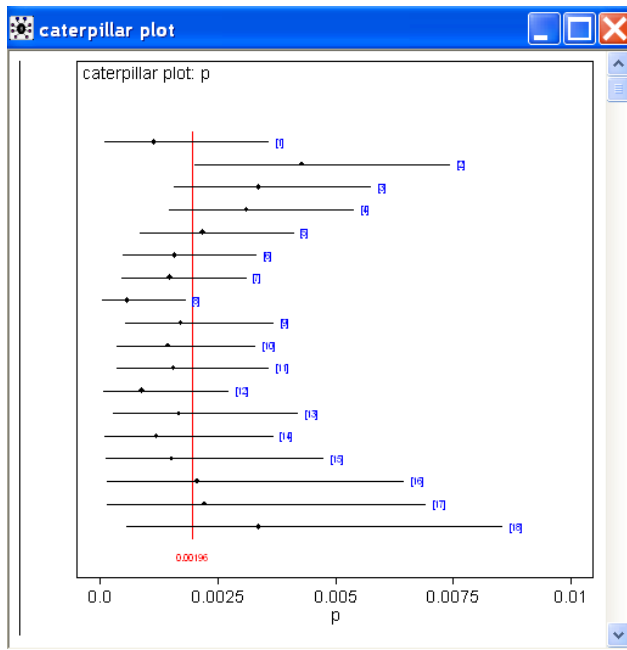
This case analyzes binned binomial data, where the unknown parameter of interest is p , the probability of failure on demand for standby equipment. The bins are set up with demands corresponding to “time” in the Poisson model above. The data analyzed are shown below.

• Table 16 Data for Case DEF

T	X	Demands
37.5	1	1297

112.5	9	2204
187.5	9	2826
262.5	9	3049
337.5	6	2991
412.5	4	2830
487.5	4	3055
562.5	1	2641
637.5	4	2614
712.5	3	2426
787.5	3	2229
862.5	1	1716
937.5	2	1507
1012.5	1	1254
1087.5	1	992
1162.5	1	736
1237.5	1	675
1387.5	2	734

The plot below shows the 95% credible intervals for p in each bin, based on a Jeffreys prior. This plot suggests a slight increasing trend in p over “time.”



• Figure 78 Caterpillar plot for p suggests slight increasing trend over time (demands)

A model with p constant over “time” was fit to this data, using a Jeffreys prior for p . The posterior mean of p was 0.0017, with a 90% credible interval of (0.0014, 0.0021). The point estimate of p in the original analysis was 0.0020. The Bayesian p -value was 0.26, suggesting that a model with constant p is adequate, but not outstanding at replicating the observed data. The p -value from the original analysis was 0.32.

A logit model, $\text{logit}(p) = a + b \cdot \text{time}$, was fit to this data, using Laplace priors for both parameters. The posterior mean and 90% interval for the intercept parameter were -5.86 and (-6.25, -5.48), compared with a

point estimate of -5.82 in the original analysis. For the aging parameter, b , the posterior mean and 90% interval were -0.001 and (-0.0017, -3.25E-4), compared with a point estimate of -8.5E-4 in the original analysis. The Bayesian p -value was 0.51, and the p -value in the original analysis was 0.66. These results suggest that a model with p increasing with "time" is statistically significant, but the small value of b suggests that the trend may be of little practical significance.

5.9.4 Case Study for Random Durations

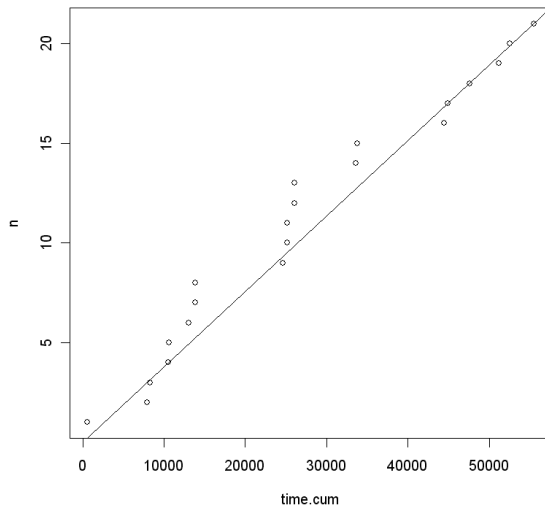
Random durations could take the form of times to failure of operating equipment or time in standby before failure for standby equipment. The data for this case study are times between failure, in days, for a single standby component: 590, 738, 55, 372, 9, 412, 132, 2, 1796, 94, 1, 136, 1, 1259, 41, 1778, 74, 452, 592, 219, 495.

The first question to decide is whether the data are from a renewal process (repair same as new after each failure) or a nonhomogeneous Poisson process (NHPP, repair same as old after each failure). We suggest that this be decided qualitatively, based on engineering judgment. Some guidelines for this judgment are as follows. For components for which failure requires replacement of the entire component (e.g., circuit board), an assumption of a renewal process seems justified. For complex components, consisting of many subcomponents, for which repair at failure typically involves replacement of a subcomponent, leaving the majority of the component as it was at the time of failure, the assumption of NHPP seems appropriate. Note that both the renewal process and NHPP include the HPP as subcases, so we still need to decide if a simple exponential distribution for time between failures is adequate.

The Laplace test is, as described in (Ascher and Feingold, 1984), a test of a homogeneous Poisson process (HPP) against a monotonic trend, that is, HPP vs. NHPP. More appropriate tests for deciding if the data are from a renewal process or NHPP are the Mann reversal test or the Lewis-Robinson test described in (Ascher and Feingold, 1984). The Lewis-Robinson test uses actual numerical values instead of the ranks, overcoming limitations of the Mann test. It is also a slight modification of the Laplace test: the value of U from the Laplace test is divided by the estimated coefficient of variation of the inter-arrival times. If the null hypothesis of HPP is true, the coefficient of variation is 1.0. Note that both of these tests are nonparametric, so their power can be limited.

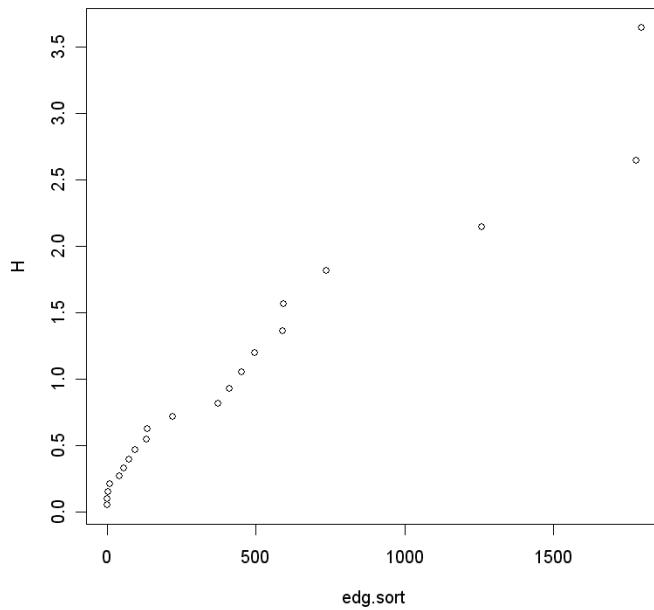
For the example times listed above, the value of the Laplace U -statistic is -0.53, and the p -value (two-sided) is 0.6, so the null hypothesis of HPP cannot be rejected. For comparing the null hypothesis of a renewal process against NHPP, the value of the Lewis-Robinson statistic is -0.43, and the two-sided p -value is 0.67, so the null hypothesis cannot be rejected. Therefore, it appears that we do not have data from an NHPP process; perhaps the data are from a renewal process with time-dependent failure rate.

A useful tool for exploratory data analysis is a plot of cumulative failures versus time. The slope of such a plot is an estimate of λ in an NHPP; if a renewal process is in effect, the slope will be approximately one. The figure below shows such a plot for the times above, with the slope of a model with constant λ shown. Thus, we have additional evidence against an NHPP.



• Figure 79 Cumulative failure plot for example times

Another tool that can be helpful is a cumulative hazard plot. This plot is generated by ordering the times between failures from smallest to largest, and then estimating the cumulative hazard function by a step function. The jumps in the step function occur at the observed times, and the height of each jump is $1/n_t$, where n_t is the number of components that have not failed by time t , when the jump occurs. If the observed failure times are exponential, the plot should be approximately a straight line. A decreasing slope with time implies a renewal process with decreasing hazard rate and vice versa. The figure below shows the plot for the listed times. The slope appears to be decreasing with time, suggesting a renewal process with decreasing hazard rate.



• Figure 80 Cumulative hazard plot, suggesting decreasing hazard rate

Analyzing these data quantitatively, using WinBUGS with diffuse priors, the Bayesian p-value for exponential inter-arrival times is 0.21, suggesting that the exponential model is adequate. However, for a Weibull model (renewal process), the Bayesian p-value is 0.53, much closer to 0.5, suggesting a significantly better fit. The posterior mean of the Weibull shape parameter was 0.6, suggesting a decreasing hazard rate. For comparison, the MLE of the shape parameter was 0.63, and is about 3.3 standard deviations below 1.0, strong evidence against the null hypothesis, and consistent with the decreasing slope suggested by Figure 80.

Alternatively, if repair is thought to be same as old (i.e., not a renewal process), then the NHPP model is more appropriate. Assuming a power-law process (Weibull time to first failure), with diffuse priors, Bayesian analysis gives a posterior mean for the shape parameter of 1.01. The MLE (corrected for bias) is 0.96, suggesting that the rate of occurrence of failures (ROCOF) is constant with time, and the NHPP can be approximated well by an HPP, as suggested by the approximately constant slope of Figure 79. However, the Bayesian p-value for this model is only 0.19.

The conclusion for this case study is that the data appear to be described best by a renewal process with decreasing hazard rate. A potential explanation for this result is that there is a preponderance of early failures, making the hazard rate appear to decrease with time in standby.

References

Rodionov A., 2008, *Overview of Reliability Data Collection with Regards to Ageing Reliability Analysis Applications*. - In: EC Enlargement and Integration Workshop on Use of PSA for Evaluation of Impact of Ageing Effects on the Safety of NPPs, 15-16 November 2007, Budapest (Hungary). EUR 23078 EN. VEIKI (Publ.), Budapest.

Atwood C.L., LaChance J.L., Martz H.F., Anderson D.J., Englehardt M., Whitehead D, Wheeler T., 2003, *Handbook of parameter estimation for probabilistic risk assessment*. NUREG/CR-6823. Washington, DC: US Nuclear Regulatory Commission.

Atwood C, Cronval O, Patrik M, Rodionov A., 2007, *Models and data used for assessing the ageing of systems, structures and components (European network on use of probabilistic safety assessment (PSA) for evaluation of ageing effects to the safety of energy facilities)*. EUR 22483 EN. Petten: EC DG JRC Institute for Energy; 2007.

Ascher H. and Feingold H., 1984, *Repairable Systems Reliability: Modeling, Inference, Misconceptions and Their Causes*, Lecture Notes in Statistics, Vol. 7, Marcel-Dekker, New York – Basel.

Bain L. J. and Engelhardt M. E., 1991, *Statistical Analysis of Reliability and Life-Testing Models: Theory and Methods*, 2nd Edition, Marcel-Dekker, New York – Basel.

Fundamental Concepts from Probability and Statistics¹

A.1 Events

Any repeatable process for which the result is uncertain can be considered an experiment, such as counting failures over time or measuring time to failure of a specific item of interest. The result of one execution of the experiment is referred to as an outcome. Repetitions or trials of a defined experiment would not be expected to produce the same outcomes due to uncertainty associated with the process. The set of all possible outcomes of an experiment is defined as the sample space.

Sample spaces can contain discrete points (such as pass, fail) or points in a continuum (such as measurement of time to failure). An event E is a specified set of possible outcomes in a sample space S (denoted $E \subset S$, where \subset denotes subset).

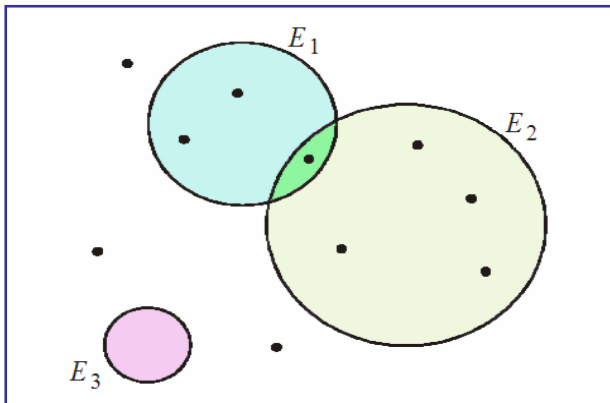


Figure B.1 Venn diagram with showing 3 events and 10 outcomes.

Most events of interest in practical situations are compound events, formed by some composition of two or more events. Composition of events can occur through the union, intersection, or complement of events, or through some combination of these.

For two events, E_1 and E_2 , in a sample space S , the union of E_1 and E_2 is defined to be the event containing all sample points in E_1 or E_2 or both, and is denoted by the symbol $(E_1 \cup E_2)$. Thus, a union is simply the event that either E_1 or E_2 or both E_1 and E_2 occur.

For two events, E_1 and E_2 , in a sample space S , the intersection of E_1 and E_2 is defined to be the event containing all sample points that are in both E_1 and E_2 , denoted by the symbol $(E_1 \cap E_2)$. The intersection is the event that both E_1 and E_2 occur.

Figure B.1 shows a picture, called a Venn diagram, of some outcomes and events. In this example, the event E_1 contains three outcomes, event E_2 contains five outcomes, the union contains seven outcomes, and the intersection contains one outcome.

The complement of an event E is the collection of all sample points in S and not in E . The complement of E is denoted by the symbol \bar{E} and is the event that all the outcomes in S that are not in E occur.

It is sometimes useful to speak of the empty or null set, a set containing no outcomes. In Figure B.1, the event E_3 is empty. It cannot occur.

Two events, E_1 and E_2 , are said to be mutually exclusive if the event $(E_1 \cap E_2)$ contains no outcomes in the sample space S . That is, the intersection of the two events is the null set. Mutually exclusive events are also referred to as disjoint events. Three or more events are called mutually exclusive, or disjoint, if each pair of events is mutually exclusive. In other words, no two events can happen together.

¹ The text in this Appendix is derived primarily from NUREG/CR-6823.

A.2 Probability Concepts

Each of the outcomes in a sample space has a probability associated with it. **Probabilities of outcomes** are seldom known; they are usually estimated via Bayesian methods. Once determined, the probabilities must satisfy two requirements:

- The probability of each outcome must be a number 0 and 1.
- The probabilities of all outcomes in a given sample space must sum to one.

Associated with any event E of a sample space S is the **probability of the event**, $\Pr(E)$. Probabilities are associated with each outcome in the sample space through a **probability model**. Probability models may be developed based on information derived from outcomes obtained from an experiment (where an “experiment” could be any operational event such as exercising a component or system).

A.3 Basic Rules and Principles of Probability

The definition of probability satisfies the following probability axioms:

- If $\Pr(E)$ is defined for a type of subset of the sample space S , and if $\Pr(E) \geq 0$, for every event E , and
- $\Pr(E_1 \cup E_2 \cup \dots) = \Pr(E_1) + \Pr(E_2) + \dots$, where the events E_1, E_2, \dots , are such that no two have a point in common, and
- $\Pr(S) = 1$,

then $\Pr(E)$ is called a **probability function**.

This probability function defines how the probability is distributed over various subsets E of a sample space S . From this definition, several rules of probability follow that provide additional properties of a probability function.

The probability of an impossible event (the empty or null set) is zero, written as:

$\Pr(\emptyset) = 0$, where \emptyset is the null set. The probability of the **complement** of E is given by:

$$\Pr(\bar{E}) = 1 - \Pr(E)$$

In general, the probability of the union of any two events is given by:

$$\Pr(E_1 \cup E_2) = \Pr(E_1) + \Pr(E_2) - \Pr(E_1 \cap E_2)$$

If E_1 and E_2 are mutually exclusive, then $\Pr(E_1 \cap E_2) = \Pr(\emptyset) = 0$, and

$$\Pr(E_1 \cup E_2) = \Pr(E_1) + \Pr(E_2)$$

which is a special case of the second axiom of probability stated above and is sometimes referred to as the addition rule for probabilities.

For three events,

$$\Pr(E_1 \cup E_2 \cup E_3) = \Pr(E_1) + \Pr(E_2) + \Pr(E_3) - \Pr(E_1 \cap E_3) - \Pr(E_2 \cap E_3) + \Pr(E_1 \cap E_2 \cap E_3).$$

This rule is also referred to as the **inclusion-exclusion principle** and can be generalized to n events.

The simultaneous occurrence of two or more events (the intersection of events) is called a joint event, and its probability is called a **joint probability**. Thus, the joint probability of both events E_1 and E_2 occurring simultaneously is denoted by $\Pr(E_1 \cap E_2)$.

The probability associated with one event, irrespective of the outcomes for the other events, can be obtained by summing all the joint probabilities associated with all the outcomes for the other events, and is referred to as a **marginal probability**. A marginal probability is therefore the *unconditional* probability of an event, unconditioned on the occurrence of any other event.

Two events E_1 and E_2 are often related in such a way that the probability of occurrence of one depends on whether the other has or has not occurred. The **conditional probability** of one event, given that the other has occurred, is equal to the joint probability of the two events divided by the marginal probability of the given event. Thus, the conditional probability of event E_2 , given event E_1 has occurred, denoted $\Pr(E_2|E_1)$, is defined as:

$$\Pr(E_2 | E_1) = \frac{\Pr(E_1 \cap E_2)}{\Pr(E_1)},$$

for $\Pr(E_1) > 0$. If $\Pr(E_1) = 0$, $\Pr(E_2|E_1)$ is undefined.

Rearranging this equation yields:

$$\Pr(E_1 \cap E_2) = \Pr(E_1) \Pr(E_2 | E_1) = \Pr(E_2) \Pr(E_1 | E_2).$$

Calculation of joint probability requires the concept of **statistical independence**. Two events E_1 and E_2 are statistically independent if the probability of one event does not change whenever the other event occurs or does not occur. Thus, E_2 is **independent** of E_1 if

$$\Pr(E_2 | E_1) = \Pr(E_2).$$

If E_2 is independent of E_1 , then E_1 is independent of E_2 . It follows that events E_1 and E_2 are independent if their joint probability is equal to the product of the *unconditional*, or *marginal*, probabilities of the events:

$$\Pr(E_1 \cap E_2) = \Pr(E_1) \Pr(E_2),$$

which is sometimes referred to as the multiplication rule for probabilities. If $\Pr(E_1)$ varies depending on whether or not event E_2 has occurred, then events E_1 and E_2 are said to be **statistically dependent**.

If E_1, E_2, \dots are mutually exclusive, and if the union of E_1, E_2, \dots equals the entire sample space, then the events E_1, E_2, \dots are said to form a **partition** of the sample space. Exactly one of the events must occur, not more than one but exactly one. In this case, the **law of total probability** says

$$\Pr(A) = \sum \Pr(A | E_i) \Pr(E_i) .$$

A special case can be written when there are only two sets. In this case, write E_1 as E and E_2 as $/E$.

Then the law of total probability simplifies to

$$\Pr(A) = \Pr(A | E) \Pr(E) + \Pr(A | /E) \Pr(/E)$$

for event A .

The concepts of mutually exclusive events and statistically independent events are often confused. If E_1 and E_2 are mutually exclusive events and $\Pr(E_1)$ and $\Pr(E_2)$ are nonzero, $\Pr(E_1 \cap E_2) = \Pr(\emptyset) = 0$. From the conditional probability equation: $\Pr(E_2 | E_1) = \frac{\Pr(E_1 \cap E_2)}{\Pr(E_1)} = 0$, which does not equal $\Pr(E_2)$. Thus, the two events are not statistically independent. Mutually exclusive events cannot be statistically independent and vice versa.

A.4 Random Variables and Probability Distributions

A.4.1 Random Variables

A **random variable** is any mathematical construct that associates real numbers with observable outcomes of an experiment. If the numbers associated with the outcomes of an experiment are all distinct and countable, the corresponding random variable is called a **discrete** random variable. An example of a discrete random variable would be the variable tracking the outcomes of a coin toss, either heads or tails.

If the sample space contains an infinite number of outcomes (like those contained in any interval), the random variable is **continuous**. Time T is a common continuous random variable, for example time to failure where the random variable T can assume any value over the range 0 to ∞ .

A.4.2 Probability Distributions

A **probability function** associates a probability with each possible value of a random variable and, thus, describes the distribution of probability for the random variable. For a discrete random variable, this function is referred to as a **discrete probability distribution function** (p.d.f.). A discrete p.d.f., commonly denoted by f , is also referred to as a **discrete distribution**, or **discrete probability mass function**, for example:

- The discrete random variable is X
- This random variable can take on a value x
- The probability distribution function is denoted $\Pr(x)$

The notation used here is that a random variable itself is given in *upper case* while an observed value (data) of the random variable is denoted in *lower case*.

Note that the sum of the probabilities over all the possible values of x must be 1.

Some discrete random variables have wide application and have therefore been defined and given specific names. The two most commonly used discrete random variables in PRA applications are the **binomial** and **Poisson** random variables. These are also called aleatory models.

A continuously distributed random variable has a **density function**, a nonnegative integral function, with the area between the graph of the function and the horizontal axis equal to 1. This density function is also referred to as the **continuous probability distribution function** (p.d.f.). If x denotes a value that the continuous random variable X can assume, the p.d.f. is often denoted as $f(x)$. The probability that X takes a value in a region A is the integral of $f(x)$ over A . In particular,

$$\Pr(a \leq X \leq b) = \int_a^b f(x) dx$$

and

$$\Pr(x \leq X \leq x + \Delta x) \approx f(x) \Delta x \tag{B.1}$$

for small Δx .

Some common continuous distributions in PRA are the **lognormal**, **exponential**, **gamma**, and **beta** distributions.

A.4.3 Cumulative Distribution Functions

Discrete probability distributions provide point probabilities for discrete random variables and continuous p.d.f.s provide point densities for continuous random variables. A related function is the **cumulative distribution function** (c.d.f.). This function is defined as the probability that the random variable assumes values less than or equal to the specific value x , and is denoted $F(x)$.

For a discrete random variable X_i with outcomes x_i , and the corresponding probabilities $\Pr(x_i)$, $F(x)$ is the sum of the probabilities of all $x_i \leq x$. That is,

$$F(x) = \Pr(X \leq x) = \sum_{x_i \leq x} \Pr(x_i).$$

For a continuous random variable X , $F(x)$ is the area beneath the p.d.f. $f(x)$ up to x . That is, $F(x)$ is the integral of $f(x)$:

$$F(x) = \Pr(X \leq x) = \int_{-\infty}^x f(y) dy.$$

Thus, $f(x)$ is the derivative of $F(x)$. If X takes on only positive values, the limits of integration are zero to x . Note that, because $F(x)$ is a probability, $0 \leq F(x) \leq 1$. If X ranges from $-\infty$ to $+\infty$, then $F(-\infty) = 0$ and $F(+\infty) = 1$.

If X has a restricted range, with a and b being the lower and upper limits of X respectively, $a < X < b$, then

$$F(a) = 0 \text{ and } F(b) = 1.$$

Also, $F(x)$ is a non-decreasing function of x , that is, if $x_2 > x_1$, $F(x_2) \geq F(x_1)$. Another important property of $F(x)$ is that

$$\Pr(x_1 < X \leq x_2) = F(x_2) - F(x_1)$$

for discrete random variables and

$$\Pr(x_1 < X \leq x_2) = F(x_2) - F(x_1)$$

for continuous random variables.

An example of a p.d.f. and the associated c.d.f. for a continuous distribution is shown in Figure B.2.

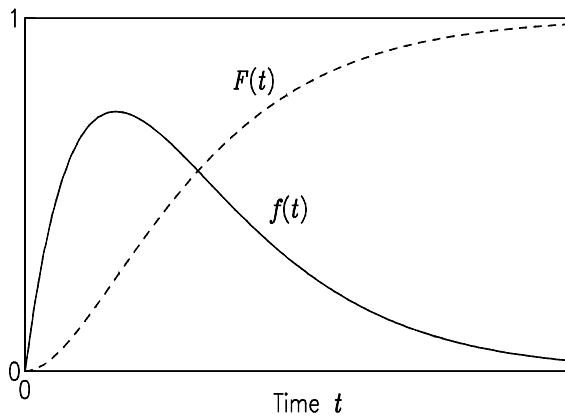


Figure B.2: Probability density function (p.d.f.) and cumulative distribution function (c.d.f.)

A.4.4 Reliability and Hazard Functions

A.4.4.1 Definitions

There are also characterizations that have special interpretations for time-to-failure distributions. Let T denote the random time to failure of a system. The **reliability function** of a system is defined as

$$R(t) = \Pr(T > t) .$$

Hence, $R(t)$, called the **reliability** at time t , is the probability that the system does not fail in the time interval $[0, t]$ or equivalently, the probability that the system is still operating at time t . (This discussion uses the notation (a, b) to mean the set of times $> a$ and $\leq b$, but the distinction between $<$ and \leq is a mathematical fine point, not important in practice.) The reliability function is also sometimes called the **survival function**. It is equal to $1 - F(t)$.

When used as a reliability criterion, it is common to state a time, say t_0 , called the **mission time**, and require for a system that the reliability at mission time t_0 be at least some prescribed level, say R_0 . For example, a

pump might be required to operate successfully for at least 12 hours with probability at least 0.95. The requirement in this case is $R_0 = 0.95$ and $t_0 = 12$. In terms of the reliability function, this would mean $R(12h) \geq 0.95$. One interpretation would be that such a pump would perform for the required mission time for 95% of the situations when it is called on to do so. Another interpretation is that 95% of all such pumps would perform as required.

Consider a system that operates for a particular mission time, unless it fails. If it fails, no immediate repairs are attempted, so some authors call the system **nonrepairable**. A common way to characterize this system's reliability is in terms of the **hazard function**. Suppose that the system is still operating at time t , and consider the probability that it will fail in a small interval of time $(t, t + \Delta t)$. This is the conditional probability $\Pr(t < T \leq t + \Delta t | T > t)$. The hazard function, h , is defined so that when Δt is small:

$$h(t)\Delta t \approx \Pr(t < T \leq t + \Delta t | T > t) \tag{B.2}$$

This function is also encountered, under the name of λ , in some treatments of Poisson processes. The hazard function gives, approximately,

$$h(t)\Delta t \approx \frac{\Pr(t < T \leq t + \Delta t)}{\Pr(T > t)} \approx \frac{f(t)\Delta t}{R(t)}$$

This is the basis for the formal definition of h :

$$h(t) = \frac{f(t)}{R(t)}$$

For details, see Bain and Engelhardt (1991, p. 541). Equation B.2 is analogous to Equation B.1, except that the probability in Equation B.2 is conditional on the system having survived until t , whereas Equation B.1 refers to all systems in the original population, either still surviving or not. Suppose a large number, say N , of identical systems are put into operation at time $t = 0$, and n is the number which fail in the interval $(t, t + \Delta t)$. It

follows that $f(t)\Delta t \approx \frac{n}{N}$, the observed relative frequency of systems failed in the interval $(t, t + \Delta t)$. On the other hand, if N_t denotes the number of the original N systems which are still in operation at time t , then

$$h(t)\Delta t \approx \frac{n}{N_t},$$

the observed relative frequency of *surviving* systems which fail in this same interval. Thus, $f(t)$

is a measure of the risk of failing at time t for any system in the original set, whereas $h(t)$ is a measure of the risk of failing at time t , but only for systems that have survived this long.

The hazard function is used as a measure of "aging" for systems in the population. If $h(t)$ is an increasing function, then systems are aging or wearing out with time. Of course, in general the hazard function can exhibit many types of behavior other than increasing with time, and other possible behaviors are discussed later in this handbook. In actuarial science the hazard function is called the **force of mortality**, and it is used as a measure of aging for individuals in a population. More generally, the hazard function gives an indication of "proneness to failure" of a system after time t has elapsed. Other terms which are also used instead of hazard function are **hazard rate** and **failure rate**. The term *failure rate* is often used in other ways in the literature of reliability (see Ascher and Feingold 1984, p. 19).

A.4.4.2 Relations among p.d.f., Reliability, and Hazard

Any one of the functions F , f , R , and h completely characterizes the distribution, and uniquely determines the other three functions. The definition

$$h(t) = \frac{f(t)}{R(t)}$$

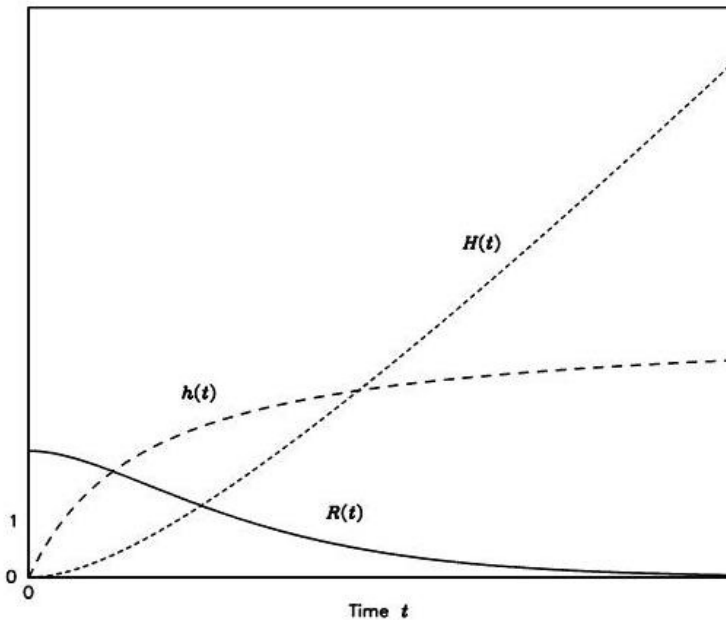
was given above. The right hand side can be written as the derivative of $-\ln[R(t)]$, leading to

$$R(t) = \exp\left(-\int_0^t h(u) du\right) = \exp(-H(t))$$

where the function $H(t)$ is called the **cumulative hazard function**. The reliability function, $R(t)$, and the c.d.f., $F(t) = 1 - R(t)$, are therefore uniquely determined by the hazard function, $h(t)$, and the p.d.f. can be expressed as

$$f(t) = h(t) \exp\left(-\int_0^t h(u) du\right).$$

Figure B.3 shows the reliability, hazard and the cumulative hazard function for the example of Figure B.2.



The hazard function in Figure B.3 is an increasing function of time. Therefore, it would be consistent with systems with a dominant wear-out effect for the entire life of the system. The lifetime of a system may be divided into three typical intervals: the **burn-in** or **infant** period, the **random** or **chance** failure period and the **wear-out** period. During the useful period, the dominant cause of failures is “random” failures. For example, systems might fail due to external causes such as power surges or other environmental factors rather than problems attributable to the defects or wear-out in the systems. This example is somewhat idealized because for many types of systems the hazard function will tend to increase slowly during the later stages of the chance failure period. This is particularly true of mechanical systems. On the other hand, for many electrical components such as transistors and other solid-state devices, the hazard function remains fairly flat once the burn-in failure period is over.

Figure B.3 The reliability function, hazard function and cumulative hazard function.

A.4.5 Joint, Marginal, and Conditional Distributions

Many statistical methods are based on selecting a sample of size n from a probability distribution $f(x)$. Such a sample is denoted by

$$(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) = (x_1, x_2, \dots, x_n),$$

where x_1, x_2, \dots, x_n are the actual values of the random variable X which has the distribution $f(x)$.

The concepts of simultaneous events and joint, marginal, and conditional probability, discussed in Section B.3, also pertain to random variables and probability distributions. Two random variables X_1 and X_2 (both continuous, both discrete, or one of each) can have a **joint distribution**, or joint p.d.f., denoted $f(x_1, x_2)$. The point (x_1, x_2) can be thought of as a point in two-dimensional Euclidean space. Similarly, n random variables have joint distribution $f(x_1, x_2, \dots, x_n)$. Also, the n random variables have joint cumulative distribution $F(x_1, x_2, \dots, x_n)$.

The **marginal distribution** of X_i is defined as the joint p.d.f. integrated (for continuous random variables) or summed (for discrete random variables) over the $n - 1$ other corresponding dimensions, resulting in a function of x_i alone. Thus, the marginal distribution of X_i is the unconditional p.d.f. of X_i , $f(x_i)$.

The **conditional distribution** of X_1 given X_2 , denoted $g(x_1 | x_2)$, is defined by

$$g(x_1 | x_2) = \frac{f(x_1, x_2)}{f_2(x_2)}$$

where $f_2(x_2) \neq 0$, and can be shown to satisfy the requirements of a probability function. Sampling from a conditional p.d.f. would produce only those values of X_1 that could occur for a given value of $X_2 = x_2$. The concept of a conditional distribution also extends to n random variables.

Two random variables X_1 and X_2 are independent if their joint p.d.f. is equal to the product of the two individual p.d.f.s. That is,

$$f(x_1, x_2) = f(x_1) f(x_2).$$

In general, X_1, X_2, \dots, X_n are **independent random variables** if

$$f(x_1, x_2, \dots, x_n) = f(x_1) \cdot f(x_2) \cdot \dots \cdot f(x_n).$$

A.4.6 Characterizing Random Variables and their Distributions

B.4.6.1 Distribution Characteristics

Probability distributions have many characteristics of interest, some of which are described by **distribution parameters**. The term parameter is used to refer to a fixed characteristic. In contrast to a statistic, which changes from sample to sample, a parameter for a particular distribution is a constant and does not change.

However, when a parameter's value is not known, sample statistics can be used to estimate the parameter value. Parameter estimation is discussed in Appendix B.10.

A very useful distribution characteristic is the parameter that serves as a measure of central tendency, which can be viewed as a measure of the middle of a distribution. When a change in the parameter slides the distribution sideways, as with the mean of a normal distribution, the parameter is referred to as the **location parameter**. It serves to locate the distribution along the horizontal axis. Sometimes, however, a change in the parameter squeezes or stretches the distribution toward or away from zero, as with the mean of the exponential distribution. In that case, the parameter is a **scale parameter**.

In any case, the most common measure of central tendency is the **mean**, μ , of the distribution, which is a weighted average of the outcomes, with the weights being probabilities of outcomes. For a discrete random variable X ,

$$\mu_X = \sum_i x_i \Pr(x_i).$$

For a continuous random variable X ,

$$\mu_X = \int_{-\infty}^{\infty} xf(x)dx.$$

Another distribution characteristic commonly used as a measure of central tendency, or location, is the **median**, which is the point along the horizontal axis for which 50% of the area under the p.d.f. lies to its left and the other 50% to its right. The median of a random variable, X , is commonly designated $\text{med}(X)$ or $x_{.50}$ and, for a continuous distribution, is the value for which $\Pr(X \leq x_{.50}) = .50$ and $\Pr(X \geq x_{.50}) = .50$. In terms of the cumulative distribution, $F(x_{.50}) = .50$. The median is a specific case of the general 100 α th **percentile**, x_{α} , for which $F(x_{\alpha}) = \alpha$. When the factor of 100 is dropped, x_{α} is called the α **quantile**. Along with the median as the 50th percentile (or equivalently, the 0.5 quantile), the 25th and 75th percentiles are referred to as **quartiles** of a distribution.

Figure B.4 shows the quartiles, $x_{0.25}$ and $x_{0.75}$, the median, $x_{0.50}$, and the mean. The quartiles and the median divide the area under the density curve into four pieces, each with the same area. Note that the mean is greater than the median in this example, which is the usual relation when the density has a long right tail, as this one does.

Figure B.5 shows the same quantities plotted with the c.d.f. By definition, the q quantile, x_q , satisfies $F(x_q) = q$.

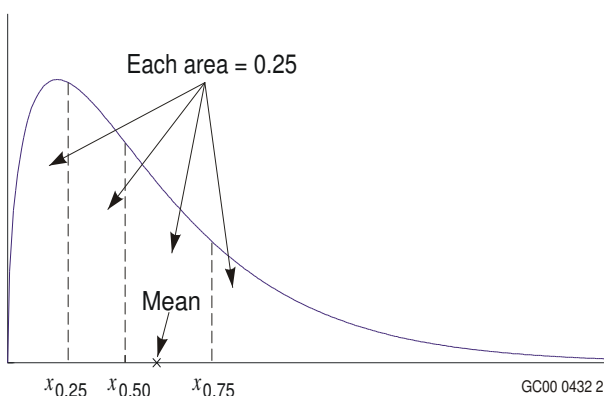


Figure B.4 Density, showing quartiles, median, and mean.

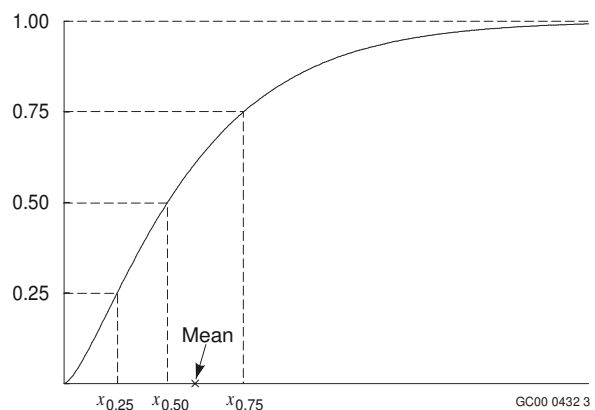


Figure B.5 Cumulative distribution function (c.d.f.) showing quartiles, median, and mean.

The mean and the median are used to measure the center or location of a distribution. Since the median is less affected by tail-area probabilities, it can be viewed as a better measure of location than the mean for highly-skewed distributions. For symmetric distributions, the mean and median are equivalent.

A different measure of center or location of a distribution is the **mode**, which indicates the most probable outcome of a distribution. The mode is the point along the horizontal axis where the “peak” or maximum of the p.d.f. is located. Note that the mode does not necessarily have to be near the middle of the distribution. It simply indicates the most likely value of a distribution. Note also that a peak does not have to exist and, in some cases, more than one peak can exist.

Another important characteristic of a distribution is its **variance**, denoted by σ^2 . The variance is the average of the squared deviations from the mean. The **standard deviation**, σ , of a distribution is the square root of its variance. Both the variance and standard deviation are measures of a distribution’s spread or dispersion. For a discrete random variable X ,

$$\sigma_X^2 = \sum_i (x_i - \mu)^2 \Pr(x_i).$$

For a continuous random variable X ,

$$\sigma_X^2 = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx .$$

Though less used than the mean and variance, the **skewness** is defined as

$$\frac{E(X - \mu)^3}{\sigma^3} .$$

It measures asymmetry. It is usually positive if the density has a longer right tail than left tail, and negative if the density has a longer left tail than right tail. For example, the density in Figure B.4 has positive skewness.

B.4.6.2 Mathematical Expectation

The definitions of distribution means and variances arise from **mathematical expectation** and **moments of a distribution**, which form an important method for calculating the parameters of a known p.d.f. In general, the **expectation (expected value** or mathematical expectation) of a function $g(X)$, denoted $E[g(X)]$, is

$$E[g(X)] = \sum_i g(x_i) \Pr(x_i) ,$$

when X is discrete, and

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f(x) dx ,$$

when X is continuous.

Because of their wide use, several expectations have special names. For $g(X) = X$, the expectation $E(X)$ becomes the mean of X . Thus, the mean is also commonly referred to as the expected value (or expectation) of the random variable X . In addition, for $g(X) = X$, the expectation $E(X)$ is known as the **first moment about the origin**.

The variance, σ_X^2 , also denoted by $\text{Var}(X)$, of a distribution is defined by mathematical expectation with $g(X) = (X - \mu_X)^2$. Thus,

$$\text{Var}(X) = \sigma_X^2 = E[(X - \mu_X)^2] = E(X^2) - [E(X)]^2,$$

which is known as the **second moment about the mean**.

Ordinary moments (moments about the origin) of a random variable X are defined as

$$M_r = E(X^r),$$

for $r = 1, 2, \dots$. Thus,

$$\text{Var}(X) = \sigma_X^2 = E(X^2) - [E(X)]^2 = M_2 - M_1^2.$$

Central moments (moments about the mean) of a random variable X are defined as being equal to $E[(X - \mu)^r]$ for $r = 2, 3, \dots$. The ordinary and central moments can be seen to define characteristics of distributions of random variables.

An important rule of expectation commonly used in PRA is that the expected value of a product of independent random variables is the product of their respective expected values. That is, $E(X_1 A X_2 A \dots A X_n) = E(X_1) A E(X_2) A \dots A E(X_n)$ when all X_i are independent. This rule also applies to conditionally independent random variables. If the random variables X_2, X_3, \dots, X_n are all conditionally independent given $X_1 = x_1$, then

$$f(x_2, x_3, \dots, x_n | x_1) = f(x_2 | x_1) A f(x_3 | x_1) A \dots A f(x_n | x_1)$$

It follows that

$$E(X_2 A X_3 A \dots A X_n | x_1) = E(X_2 | x_1) A E(X_3 | x_1) A \dots A E(X_n | x_1).$$

Thus,

$$E(X_1 A X_2 A \dots A X_n) = E[X_1 A E(X_2 | x_1) A E(X_3 | x_1) A \dots A E(X_n | x_1)].$$

B.4.6.4 Covariance and Correlation

For two random variables, X and Y , with means μ_X and μ_Y , the expected value $E[(X - \mu_X)(Y - \mu_Y)]$ is called the **covariance** of X and Y , denoted $\text{Cov}(X, Y)$. The covariance of X and Y divided by the product of the standard deviations of X and Y is called the **correlation coefficient** (or correlation) between X and Y , denoted $\text{Cor}(X, Y)$. That is,

$$\text{Cor}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}} = \frac{E(X - \mu_X)E(Y - \mu_Y)}{\sqrt{E[(X - \mu_X)^2]E[(Y - \mu_Y)^2]}} .$$

The correlation coefficient measures the degree of association between X and Y , that is, the strength of a linear relationship between X and Y .

A.4.7 Distribution of a Transformed Random Variable

This section considers the distribution of $Y = h(X)$, when X has a known distribution and h is a known function. The problem is straightforward when X has a discrete distribution. When X is continuous and h is monotone, either increasing or decreasing, the c.d.f.s are also related in the natural way, as follows. Let F be the c.d.f. of X and let G be the c.d.f. of Y . Then we have

$$G(y) = \Pr(Y \leq y) = \Pr[h(X) \leq y] .$$

If h is monotone increasing, this equals

$$\Pr[X \leq h^{-1}(y)] = F(x) ,$$

where x and y are related by $y = h(x)$, $x = h^{-1}(y)$. In summary, $G(y) = F(x)$.

If, instead, h is monotone decreasing, then a similar argument gives

$$G(y) = 1 - F(x) .$$

The surprise comes with the densities. Differentiate both sides of either of the above equations with respect to y , to obtain the density of y . This involves using the chain rule for differentiation. The result is

$$g(y) = f(x) \left| \frac{dx}{dy} \right| .$$

That is, the density of Y is not simply equal to the density of X with a different argument. There is also a multiplier, the absolute value of the derivative. If $Y = \exp(X)$, then

$$g(y) = f[\ln(y)](1/y)$$

If $Y = 1/X$, then

$$g(y) = f(1/y)(1/y^2) .$$

The formulas here are the basis for the densities of the lognormal distribution and the inverted gamma distribution.

A.5 Bayes' Theorem

It is frequently desired to calculate the probability of an event A given that another event B has occurred at some prior point in time. It can also be of interest to calculate the probability that a state of nature exists given that a certain sample is observed, for example, belonging to a certain population based on a sample measurement or observation. Conditional probability leads directly to **Bayes' Theorem**, which, along with subjective probability, forms the basis for Bayesian inference commonly used in PRA.

Bayes' Theorem states that: if A_1, A_2, \dots, A_n are a sequence of disjoint events and if B is any other event such that $\Pr(B) > 0$, then

$$\Pr(A_i | B) = \frac{\Pr(B | A_i) \Pr(A_i)}{\Pr(B)} \quad (\text{B.3})$$

Where

$$\Pr(B) = \sum_{j=1}^n \Pr(B | A_j) \Pr(A_j) .$$

Equation B.6 follows from the definition of conditional probability in Equation B.2:

$$\Pr(A_i | B) = \frac{\Pr(B \cap A_i)}{\Pr(B)} = \frac{\Pr(B | A_i)}{\Pr(B)}$$

The $\Pr(A_i | B)$ is the **posterior** (or a posteriori) probability for the event A_i , meaning the probability of A_i once B is known. The $\Pr(A_i)$ is the **prior** (or a priori) probability of the event A_i before experimentation or observation. The event B is the observation. The $\Pr(B | A_i)$ is the probability of the observation given that A_i is true. The denominator serves as a normalizing constant.

Calculating the posterior probabilities $\Pr(A_i|B)$ requires knowledge of the probabilities $\Pr(A_i)$ and $\Pr(B|A_i)$, $i = 1, 2, \dots, n$. The probability of an event can often be determined if the population is known, thus, the $\Pr(B|A_i)$ can be determined. However, the $\Pr(A_i)$, $i = 1, 2, \dots, n$, are the probabilities that certain states of nature exist and are either unknown or difficult to ascertain. These probabilities, $\Pr(A_i)$, are called prior probabilities for the events A_i because they specify the distribution of the states of nature prior to conducting the experiment.

Application of Bayes' Theorem utilizes the fact that $\Pr(B|A_i)$ is easier to calculate than $\Pr(A_i|B)$. If probability is viewed as degree of belief, then the prior belief is changed, by the test evidence, to a posterior degree of belief. In many situations, some knowledge of the prior probabilities for the events A_1, A_2, \dots, A_n exists. Using this prior information, inferring which of the set A_1, A_2, \dots, A_n , is the true population can be achieved by calculating the $\Pr(A_i|B)$ and selecting the population that produces the highest probability.

Equation B.6 pertains to disjoint discrete events and discrete probability distributions. Bayes' Theorem has analogous results for continuous p.d.f.'s. Suppose X is a continuous random variable, with p.d.f. depending on parameter θ , and with conditional p.d.f. of X , given θ , specified by $f(x|\theta)$. Consider θ to be a possible value of the random variable Θ (using the convention of denoting random variables with uppercase letters). If the prior p.d.f. of Θ is denoted $g(\theta)$, then for every x such that $f(x) > 0$ exists, the posterior p.d.f. of Θ , given $X = x$, is

$$g(\theta | x) = \frac{f(x | \theta)g(\theta)}{f(x)},$$

Where

$$f(x) = \int f(x | \theta)g(\theta)d\theta$$

is the marginal p.d.f. of X . Again, the prior and posterior p.d.f.'s can be used to represent the knowledge and beliefs about the likelihood of various values of a random variable Θ prior to and posterior to observing a value of another random variable X .

A.6 Discrete Random Variables

A.6.1 Binomial Distribution

The **binomial** distribution describes the number of failures X in n independent trials. The random variable X has a binomial distribution if:

1. The number of random trials is one or more and is known in advance.
2. Each trial results in one of two outcomes, usually called success and failure (or could be pass-fail, hit-miss, defective-nondefective, etc.).
3. The outcomes for each trial are statistically independent.
4. The probability of failure, p , is constant across trials.

Equal to the number of failures in the n trials, a binomial random variable X can take on any integer value from 0 to n . The probability associated with each of these possible outcomes, x , is defined by the binomial(n, p) p.d.f. as

$$\Pr(X = x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, \dots, n$$

Here

$$\binom{n}{x} = \frac{n!}{x!(n-x)!}$$

is the **binomial coefficient** and

$$n! = n(n-1)(n-2) \dots (2)(1)$$

denotes n factorial, with $0!$ defined to be equal to 1. This binomial coefficient provides the number of ways that exactly x failures can occur in n trials (number of combinations of n trials selected x at a time).

The binomial distribution has two parameters, n and p , of which n is known. (Although n may not always be known exactly, it is treated as known in this handbook.)

The mean and variance of a binomial(n, p) random variable X are

$$E(X) = np$$

and

$$\text{Var}(X) = np(1-p).$$

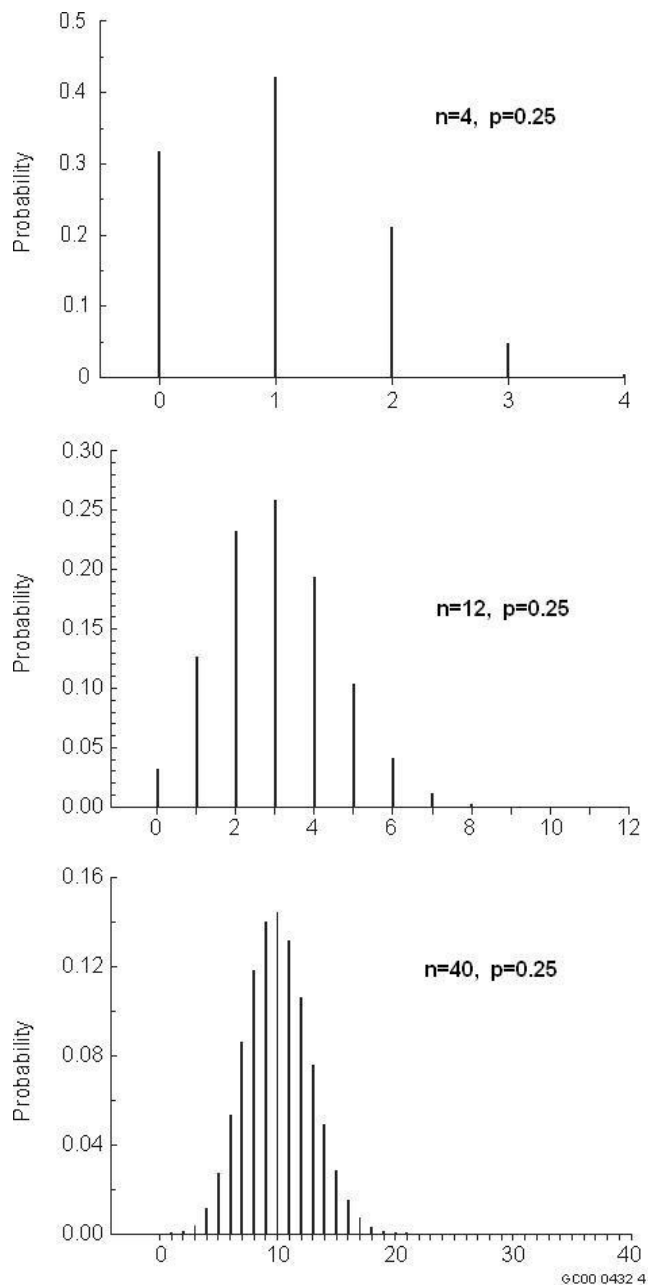


Figure B.6 Three binomial probability distribution functions.

Figure B.6 shows three binomial probability distribution functions, with parameter $p = 0.25$, and $n = 4, 12$, and 40. In each case, the mean is np . The means have been aligned in the three plots.

A.6.2 Poisson Distribution

The **Poisson** distribution provides a discrete probability model that is appropriate for many random phenomena that involve counts. Examples are counts per fixed time interval of the number of items that fail, the number of customers arriving for service, and the number of telephone calls occurring. A common use of the Poisson distribution is to describe the behavior of many rare event occurrences. The Poisson distribution is also frequently used in applications to describe the occurrence of system or component failures under steady-state conditions.

The count phenomena that occur as Poisson random variables are not necessarily restricted to occurring over a time interval. They could also be counts of things occurring in some region, such as defects on a surface or within a certain material. A process that leads to a Poisson random variable is said to be a **Poisson process**.

The Poisson distribution describes the total number of events occurring in some interval of time t (or space). The p.d.f. of a Poisson random variable X , with parameter $\mu = \lambda t$, is

$$\Pr(X = x) = \frac{e^{-\mu} \mu^x}{x!} = \frac{e^{-\lambda t} (\lambda t)^x}{x!} \quad (\text{B.4})$$

for $x = 0, 1, 2, \dots$, and $x! = x(x-1)(x-2) \dots (2)(1)$, as defined previously.

The Poisson distribution has a single parameter μ , denoted $\text{Poisson}(\mu)$. If X denotes the number of events that occur during some time period of length t , then X is often assumed to have a Poisson distribution with parameter $\mu = \lambda t$. In this case, X is considered to be a Poisson process with **intensity** $\lambda > 0$ (Martz and Waller 1991). The variable λ is also referred to as the **event rate** (or **failure rate** when the events are failures). Note that λ has units 1/time; thus, $\lambda t = \mu$ is dimensionless.

If only the total number of occurrences for a single time period t is of interest, the form of the p.d.f. in Equation B.7 using μ is simpler. If the event rate, λ , or various time periods, t , are of interest, the form of the p.d.f. in Equation B.7 using λt is more useful.

The expected number of events occurring in the interval 0 to t is $\mu = \lambda t$. Thus, the mean of the Poisson distribution is equal to the parameter of the distribution, which is why μ is often used to represent the parameter. The variance of the Poisson distribution is also equal to the parameter of the distribution. Therefore, for a $\text{Poisson}(\mu)$ random variable X ,

$$E(X) = \text{Var}(X) = \mu = \lambda t.$$

Figure B.7 shows three Poisson probability distribution functions, with means $\mu = 1.0, 3.0$, and 10.0 , respectively. The three means have been aligned in the graphs. Note the similarity between the Poisson distribution and the binomial distribution when $\mu = np$ and n is not too small.

Several conditions are assumed to hold for a Poisson process that produces a Poisson random variable:

For small intervals, the probability of exactly one occurrence is approximately proportional to the length of the interval (where λ , the event rate or intensity, is the constant of proportionality).

For small intervals, the probability of more than one occurrence is essentially equal to zero (see below).

The numbers of occurrences in two non-overlapping intervals are statistically independent.

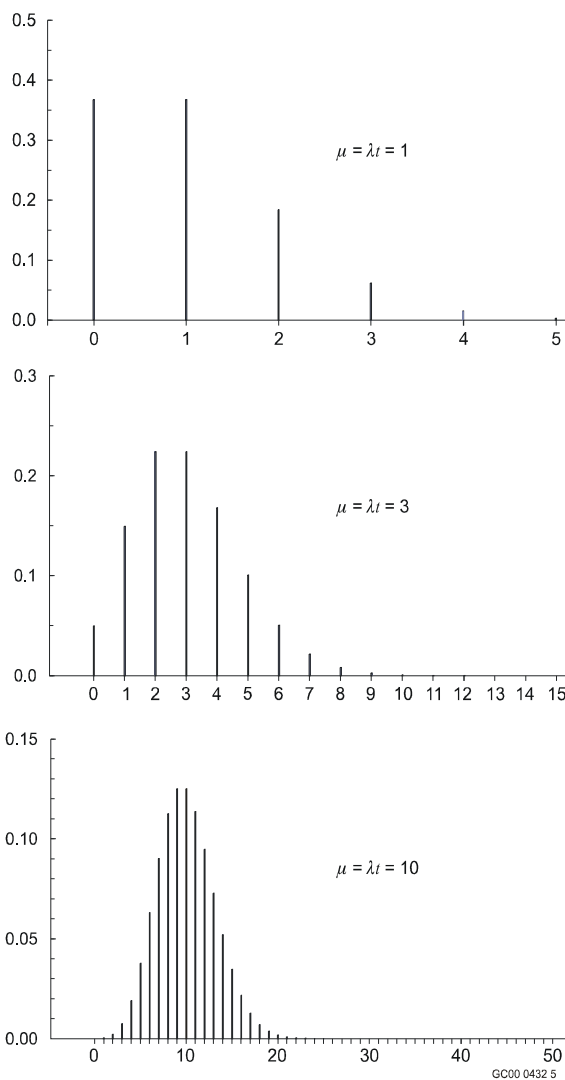


Figure B.7 Three Poisson probability distribution functions.

More precise versions of condition 2 are: (1) the probability of more than one event occurring in a very short time interval is negligible in comparison to the probability that only one event occurs (Meyer 1970), (2) the probability of more than one event occurring in a very short time interval goes to zero faster than the length of the interval (Pfeiffer and Schum 1973), and (3) simultaneous events occur only with probability zero (Çinlar 1975). All of these versions have the practical interpretation that common cause events do not occur. The phrase “do not occur” is used in this handbook, as it is in Thompson (1981).

The Poisson distribution also can serve as an **approximation to the binomial distribution**. Poisson random variables can be viewed as resulting from an experiment involving a large number of trials, n , that each have a small probability of occurrence, p , of an event. However, the rare occurrence is offset by the large number of trials. As stated above, the binomial distribution gives the probability that an occurrence will take place exactly x times in n trials. If $p = \mu/n$ (so that p is small for large n), and n is large, the binomial probability that the rare occurrence will take place exactly x times is closely approximated by the Poisson distribution with $\mu = np$. In general, the approximation is good for large n , small p , and moderate μ (say $\mu \leq 20$) (see Derman et al. 1973).

The Poisson distribution is important because it describes the behavior of many rare event occurrences, regardless of their underlying physical process. It also has many applications to describing the occurrences of system and component failures under steady-state conditions. These applications utilize the relationship between the Poisson and **exponential** (continuous random variable, see Section B.7.4) distributions: the times between successive events follow an exponential distribution.

A.7 Continuous Random Variables

A.7.1 Uniform Distribution

A **uniform** distribution, also referred to as a rectangular distribution, represents the situation where any value in a specified interval, say $[a, b]$, is equally likely. For a uniform random variable, X , because the outcomes are equally likely, $f(x)$ is equal to a constant. The p.d.f. of a uniform distribution with parameters a and b , denoted $\text{uniform}(a, b)$ is

$$f(x) = \frac{1}{b-a}$$

for $a \leq x \leq b$.

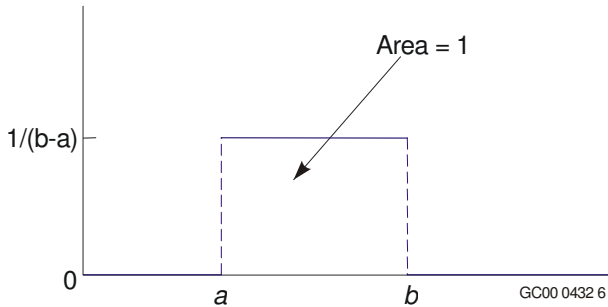


Figure B.8 Density of uniform(a, b) distribution.

Figure B.8 shows the density of the uniform(a, b) distribution.

The mean and variance of a uniform(a, b) distribution are

$$E(X) = \frac{b+a}{2}$$

And

$$Var(X) = \frac{(b-a)^2}{12}$$

A.7.2 Normal Distribution

One of the most widely encountered continuous probability distributions is the **normal** distribution, which has the familiar bell shape and is symmetrical about its mean value. The importance of the normal distribution is due to: (1) its applicability in describing a very large number of random variables that occur in nature and (2) the fact that certain useful functions of nonnormal random variables are approximately normal. Details on the derivation of the normal distribution can be found in many basic mathematical statistics textbooks (e.g., Hogg and Craig 1995).

The normal distribution is characterized by two parameters, μ and σ . For a random variable, X , that is normally distributed with parameters μ and σ , the p.d.f. of X is

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] \quad (\text{B.5})$$

for $-\infty < x < \infty$, $-\infty < \mu < \infty$, and $\sigma > 0$. Increasing μ moves the density curve to the right and increasing σ spreads the density curve out to the right and left while lowering the peak of the curve. The units of μ and σ are the same as for X .

The mean and variance of a normal distribution with parameters μ and σ are

$$E(X) = \mu$$

and

$$Var(X) = \sigma^2.$$

The normal distribution is denoted normal(μ, σ^2).

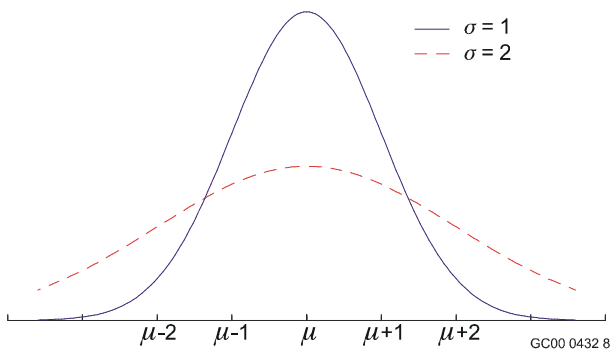


Figure B.9 Two normal densities.

Figure B.9 shows two normal(μ, σ^2) densities. The distribution is largest at μ and is more concentrated around μ when σ is small than when σ is large.

Note the similarity of the normal density to a binomial p.d.f. with large np or a Poisson p.d.f. with large μ . This illustrates the fact that a normal distribution can sometimes be used to approximate those distributions.

The normal(0, 1) distribution is called the **standard normal** distribution, which, from Equation B.5, has p.d.f.

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \quad (\text{B.6})$$

for $-\infty < x < \infty$.

The cumulative distribution of the standard normal distribution is denoted by Φ . Tables for the standard normal distribution are presented in Appendix C and in almost all books on statistics.

It can be shown that the transformed random variable $Z = (X - \mu)/\sigma$ is normal(0, 1). Thus, to calculate probabilities for a normal(μ, σ^2) random variable, X , when $\mu \neq 0$ and/or $\sigma^2 \neq 1$, the tables for the standard normal can be used. Specifically, for any number a ,

$$\Pr[X \leq a] = \Pr\left[\frac{(X - \mu)}{\sigma} \leq \frac{(a - \mu)}{\sigma}\right] = \Pr\left[Z \leq \frac{(a - \mu)}{\sigma}\right] = \Phi\left[\frac{(a - \mu)}{\sigma}\right]$$

Part of the importance of the normal distribution is that it is the distribution that sample sums and sample means tend to possess as n becomes sufficiently large. This result is known as the **central limit theorem**, which states that, if X_1, X_2, \dots, X_n are independent random variables, each with mean μ and variance σ^2 , the sum of these n random variables, $\sum_{i=1}^n X_i$, tends toward a normal($n\mu, n\sigma^2$) distribution for large enough n . Since the sample mean is a linear combination of this sum, the central limit theorem also applies. Thus, $\bar{X} = \sum_{i=1}^n X_i/n$ tends to a normal($\mu, \sigma^2/n$) distribution. The importance of the central limit theorem is it can be used to provide approximate probability information for the sample sums and sample means of random variables whose distributions are unknown. Further, because many natural phenomena consist of a sum of several random contributors, the normal distribution is used in many broad applications.

Because a binomial random variable is a sum, it tends to the normal distribution as n gets large. Thus, the normal distribution can be used as an **approximation to the binomial distribution**. One rule of thumb is that the approximation is adequate for $np \geq 5$.

A Poisson random variable also represents a sum and, as presented previously, can also be used as an approximation to the binomial distribution. It follows that the normal distribution can serve as an **approximation to the Poisson distribution** when $\mu = \lambda t$ is large. One rule of thumb is that the approximation is adequate for $\mu \geq 5$.

A.7.3 Lognormal Distribution

Use of the **lognormal distribution** has become increasingly widespread. It is commonly used as a distribution for failure time and in maintainability analysis (Martz and Waller 1991). It has also been widely used as a prior distribution for unknown positive parameters.

The lognormal distribution arises from the *product* of many independent random variables. If $Y = Y_1 \cdot Y_2 \cdot \dots \cdot Y_n = \prod_i Y_i$ is the product of n independent positive random variables that are (nearly) identically distributed, then $\ln(Y) = \ln\left(\prod_i Y_i\right) = \sum_i \ln(Y_i)$ is a sum that tends toward a normal distribution.

The distribution of Y is defined to be lognormal when the distribution of $\ln(Y)$ is normal. That is, when Y is lognormal, $\ln(Y)$ is normal(μ, σ^2). The parameters of the lognormal distribution are μ and σ , the parameters from the underlying normal distribution. For a random variable, Y , that is lognormally distributed with parameters μ and σ , denoted lognormal(μ, σ^2), the p.d.f. of Y is

$$f(y) = \frac{1}{\sigma y \sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma^2}(\ln(y) - \mu)^2\right]$$

for $0 < y < \infty$, $-\infty < \mu < \infty$, and $\sigma > 0$. Note the y in the denominator, for reasons explained in Section B.4.7. The mean and variance of a lognormal(μ, σ^2) distribution are

$$E(Y) = \exp\left[\frac{(\mu + \sigma^2)}{2}\right]$$

and

$$Var(Y) = \exp(2\mu + \sigma^2) [\exp(\sigma^2) - 1].$$

In addition, the median of a lognormal distribution is $\exp(\mu)$ and the mode is $\exp(\mu - \sigma^2)$. See Martz and Waller (1991) for more information on the lognormal distribution.

Sometimes the median of $Y = \exp(\mu)$ is used as a parameter. In addition, a parameter commonly used in PRA is the **error factor** (EF), where

$$EF = \exp(1.645\sigma),$$

and is defined as

$$\Pr\left[\frac{med(Y)}{EF} \leq Y \leq med(Y) * EF\right] = 0.90.$$

Figure B.10 shows three lognormal densities. The value $\mu = -7$ corresponds to a median of about 1.E-3. [More exactly, it corresponds to $\exp(-7) = 9.E-4$.] The value $\mu = -6.5$ corresponds to a median of about 1.5E-3. The value $\sigma = 0.67$ corresponds to an error factor $EF = 3$, and $\sigma = 1.4$ corresponds to an error factor $EF = 10$.

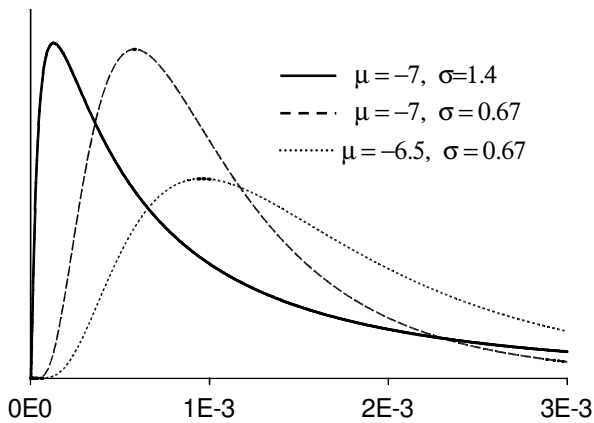


Figure B.10 Three lognormal densities.

The two distributions with $\sigma = 0.67$ and different values of μ have essentially the same shape, but with different scales. The larger μ corresponds to spreading the distribution out more from zero. The distribution with $\sigma = 1.4$, and therefore $EF = 10$, has a very skewed distribution.

To calculate probabilities for a lognormal(μ, σ^2) random variable, Y , the tables for the standard normal can be used. Specifically, for any number b ,

$$\begin{aligned} \Pr[Y \leq b] &= \Pr[\ln(Y) \leq \ln(b)] = \Pr[X \leq \ln(b)] \\ &= \Phi\left[\frac{\ln(b) - \mu}{\sigma}\right] \end{aligned}$$

where $X = \ln(Y)$ is normal(μ, σ^2).

A.7.4 Exponential Distribution

The **exponential distribution** is widely used for modeling time to failure and is inherently associated with the Poisson process (see Martz and Waller 1991). For a Poisson random variable X defining the number of failures in a time interval t and for a random variable T defining the time to failure, it can be shown that T has the exponential p.d.f.

$$f(t) = \lambda e^{-\lambda t},$$

for $t > 0$. Thus, the time to first failure and the times between successive failures follow an exponential distribution and the number of failures in a fixed time interval follows a Poisson distribution.

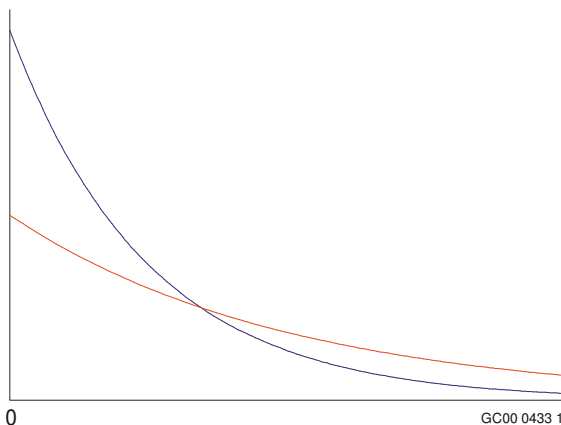


Figure B.11 Two exponential densities.

Figure B.11 shows two exponential densities, for two values of λ . The intercept (height of the curve when $t = 0$) equals λ . Thus, the figure shows that the distribution is more concentrated near zero if λ is large. This agrees with the interpretation of λ as a frequency of failures and t as time to first failure.

The exponential distribution parameter, λ , corresponds to the λt parameterization of the Poisson p.d.f. in Equation B.4. and is referred to as the **failure rate** if the component or system is repaired and restarted immediately after each failure. It is called the **hazard rate** if the component or system can only fail once and cannot be repaired. Section 4.6.2 discusses modeling duration times with different distributions and defines the hazard rate as $h(t) = \frac{f(t)}{[1 - F(t)]}$. For

the exponential distribution, the hazard rate is constant, λ .

The c.d.f. of the exponential distribution is

$$F(t) = 1 - e^{-\lambda t}.$$

The exponential distribution with parameter λ is denoted exponential(λ). The mean and variance of an exponential(λ) distribution are

$$E(T) = \frac{1}{\lambda}$$

and

$$\text{Var}(T) = \frac{1}{\lambda^2}.$$

The relationship of the exponential distribution to the Poisson process can be seen by observing that the probability of no failures before time t can be viewed in two ways. First, the number of failures, X , can be counted. The probability that the count is equal to 0 is given by Equation B.4 as

$$\Pr(X = 0) = e^{-\lambda t} \frac{(\lambda t)^0}{0!} = e^{-\lambda t}.$$

Alternatively, the probability that first failure time, T , is greater than t is

$$\Pr(T > t) = 1 - \Pr(T \leq t) = 1 - F(t) = 1 - [1 - e^{-\lambda t}] = e^{-\lambda t}.$$

Thus, the two approaches give the same expression for the probability of no failures before time t .

The assumptions of a Poisson process require a constant failure rate, λ , which can be interpreted to mean that the failure process has no memory (Martz and Waller 1991). Thus, if a device is still functioning at time t , it remains as good as new and its remaining life has the same exponential(λ) distribution. This constant failure rate corresponds to the flat part of the common "bathtub" failure curve (number of failures plotted against time) and does not pertain to initial (burn-in) failures and wear-out failures.

A different, sometimes useful, parameterization uses $\mu = 1/\lambda = E(T)$. For example, if T represents a time to failure, μ is called the mean time to failure. If T is the time to repair, or to fire suppression, or to some other event, the name for μ is the mean time to repair, or other appropriate name. The exponential(μ) distribution for T has density

$$f(t) = \left(\frac{1}{\mu}\right) \exp\left(-\frac{t}{\mu}\right), \text{ for } t \geq 0$$

and c.d.f.

$$F(t) = 1 - \exp\left(-\frac{t}{\mu}\right), \text{ for } t \geq 0.$$

The units of μ are the same as the units of t , minutes or hours or whatever the data have. The mean and variance are

$$E(T) = \mu$$

and

$$Var(T) = \mu^2 .$$

A.7.5 Weibull Distribution

The **Weibull distribution** is widely used in reliability and PRA and generalizes the exponential distribution to include nonconstant failure or hazard rates (Martz and Waller 1991). Different Weibull distributions have been successfully used to describe initial failures and wear-out failures. The Weibull distribution is appropriate when a system is composed of a number of components, and system failure is due to any one of the components failing. It, therefore, is commonly referred to as a distribution corresponding to failure of the weakest link.

For a random variable, T , that has a Weibull distribution, the p.d.f. is

$$f(t) = \frac{\beta}{\alpha} \left(\frac{t - \theta}{\alpha} \right)^{\beta-1} \exp \left[- \left(\frac{t - \theta}{\alpha} \right)^\beta \right],$$

for $t \geq \theta \geq 0$ and parameters $\alpha > 0$ and $\beta > 0$. The parameter θ is a location parameter and corresponds to a period of guaranteed life that is not present in many applications (Martz and Waller 1991). Thus, θ is usually set to zero. The c.d.f. for T is

$$F(t) = 1 - \exp \left[- \left(\frac{t - \theta}{\alpha} \right)^\beta \right],$$

for $t \geq \theta$ and $\alpha > 0$ and $\beta > 0$.

The α parameter is a scale parameter that expands or contracts the density along the horizontal axis. The β parameter is a shape parameter that allows for a wide variety of distribution shapes [see Martz and Waller (1991) for further discussion and examples]. When $\beta = 1$, the distribution reduces to the exponential distribution. Therefore, the Weibull family of distributions includes the exponential family of distributions as a special case.

A Weibull distribution with parameters α , β , and θ is referred to as Weibull(α , β , θ) and, when $\theta = 0$, Weibull(α, β). The mean and variance of the Weibull distribution are given by Martz and Waller (1991) as

$$\theta + \alpha \Gamma \left(1 + \frac{1}{\beta} \right)$$

and

$$\alpha^2 \left[\Gamma\left(1 + \frac{2}{\beta}\right) - \Gamma^2\left(1 + \frac{1}{\beta}\right) \right].$$

Here, Γ is the gamma function, defined in Sec. B.7.6.

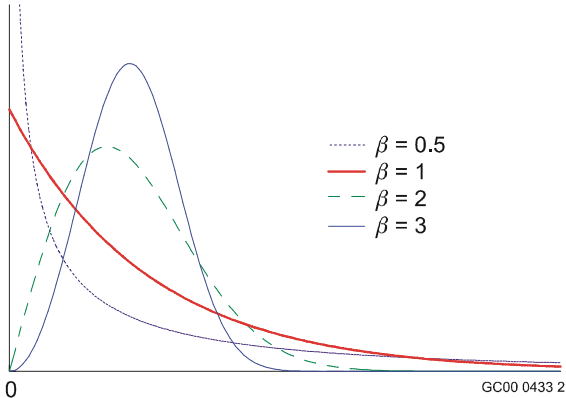


Figure B.12 shows four Weibull densities, all with the same scale parameter, α , and all with location parameter $\theta = 0$. The shape parameter, β , varies. When $\beta < 1$, the density becomes infinite at the origin. When $\beta = 1$, the distribution is identical to the exponential distribution. Surprisingly, the distribution is *not* asymptotically normal as β becomes large, although it is approximately normal when β is near 3.

Figure B.12 Four Weibull densities, all having $\theta = 0$ and all having the same α .

A.7.6 Gamma and Chi-Squared Distributions

The **gamma distribution** is an extension of the exponential distribution and is sometimes used as a failure time model (Martz and Waller, 1991). It is also often used as a prior distribution in Bayesian estimation (see Appendix B) of the failure rate parameter λ from Poisson(λt) or exponential(λ) data. The **chi-squared distribution** is a re-expression of a special case of the gamma distribution.

The gamma distribution arises in many ways. The distribution of the sum of independent exponential(λ) random variables is gamma, which forms the basis for a confidence interval for λ from exponential(λ) data. Because the sum of n independent exponentially distributed random variables has a gamma distribution, the gamma distribution is often used as the distribution of the time, or waiting time, to the n th event in a Poisson process. In addition, the chi-squared distribution is the distribution for a sum of squares of independent, identically distributed normal random variables, which forms the basis for a confidence interval for the variance of a normal distribution. The gamma distribution is also often used as a distribution for a positive random variable, similar to the lognormal and Weibull distributions. In PRA work, it is often used as a Bayesian distribution for an uncertain positive parameter.

Two parameterizations of the gamma distribution are common, with various letters used for the parameters. The parameterization given here is most useful for Bayesian updates, the primary use of the gamma distribution in this handbook. For a random variable, T , that has a gamma distribution, the p.d.f. is

$$f(t) = \frac{\beta^\alpha}{\Gamma(\alpha)} t^{\alpha-1} \exp(-t\beta) ,$$

for t , α , and $\beta > 0$. Here

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx$$

is the **gamma function** evaluated at α . If α is a positive integer, $\Gamma(\alpha) = (\alpha - 1)!$.

A gamma distribution with parameters α and β is referred to as $\text{gamma}(\alpha, \beta)$. The mean and variance of the $\text{gamma}(\alpha, \beta)$ random variable, T , are:

$$E(T) = \frac{\alpha}{\beta}$$

and

$$\text{Var}(T) = \frac{\alpha}{\beta^2} .$$

The parameters α and β are referred to as the shape and scale parameters. The shape parameter α allows the density to have many forms. If α is near zero, the distribution is highly skewed. For $\alpha = 1$, the gamma distribution reduces to an exponential (β^{-1}) distribution. Also, the $\text{gamma}(\alpha = n/2, \beta = 1/2)$ distribution is known as the **chi-squared distribution** with n degrees of freedom, denoted $\chi^2(n)$. The p.d.f. for the $\chi^2(n)$ distribution is found by substituting these values into the above formula for the gamma p.d.f. It also can be found in many statistics texts (e.g., Hogg and Craig 1995, Chapter 4).

In addition, if T has a $\text{gamma}(\alpha, \beta)$ distribution, then $2\beta T$ has a $\chi^2(2\alpha)$ distribution, which forms the defining relationship between the two distributions. The gamma and chi-squared distributions can, therefore, be viewed as two ways of expressing one distribution. Since the chi-squared distribution usually is only allowed to have integer degrees of freedom, the gamma distribution can be thought of as an interpolation of the chi-squared distribution.

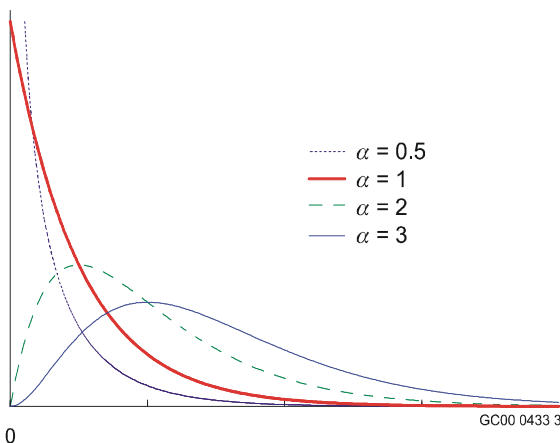


Figure B.13 Gamma densities with four shape parameters.

Percentiles of the chi-squared distribution are tabulated in Appendix C. These tables can be used as follows to find the percentiles of any gamma distribution. The $100 \times p$ percentile of a $\text{gamma}(\alpha, \beta)$ distribution is $\chi^2_p(2\alpha)/(2\beta)$, where $\chi^2_p(2\alpha)$ denotes the $100 \times p$ percentile of the chi-squared distribution with 2α degrees of freedom.

Figure B.13 shows gamma densities with four shape parameters, α . When $\alpha < 1$, the density becomes infinite at 0. When $\alpha = 1$, the density is identical to an exponential density. When α is large, the distribution is approximately a normal distribution.

As stated previously, the sum of exponential lifetimes or waiting times has a gamma distribution, with the shape parameter α equal to the number of exponential lifetimes. Thus, when α is large, the **gamma distribution is approximately normal**.

An alternative parameterization of the gamma distribution uses the scale parameter, say $\tau = \beta^{-1}$. If T has a $\text{gamma}(\alpha, \tau)$ distribution, its p.d.f. is

$$f(t) = \frac{1}{\tau^\alpha \Gamma(\alpha)} t^{\alpha-1} \exp\left(-\frac{t}{\tau}\right)$$

for t , α , and $\tau > 0$. The mean and variance of the gamma(α , τ) random variable, T , are:

$$E(T) = \alpha\tau$$

and

$$\text{Var}(T) = \alpha\tau^2.$$

A.7.7 Inverted Gamma and Inverted Chi-Squared Distributions

The **inverted gamma distribution** is often used as a prior distribution for Bayesian estimation of the time to failure of an exponential(λ) distribution (Martz and Waller 1991). It is also used as a prior and posterior distribution for σ^2 when the data have a normal distribution with variance σ^2 (Box and Tiao 1973, Lee 1997).

For a gamma(α , β) random variable, T , $W = 1/T$ has an inverted gamma distribution with p.d.f.

$$f(w) = \frac{\beta^\alpha}{\Gamma(\alpha)} \left(\frac{1}{w}\right)^{\alpha+1} \exp\left(-\frac{\beta}{w}\right),$$

for w , α , and $\beta > 0$. The parameters here are the same as for the gamma distribution. For example, if T has units of time then w and β both have units 1/time. A comparison of this density with the gamma density shows that this density has an extra w^2 in the denominator, for reasons explained in Section B.4.7.

The parameters of the inverted gamma distribution are α and β and this distribution is denoted inverted gamma(α , β). Similar to the gamma(α , β) distribution, α is the shape parameter and β is the scale parameter. The distribution can also be parameterized in terms of $\tau = \beta^{-1}$.

The mean and variance of an inverted gamma(α , β) random variable, W , are

$$E(W) = \frac{\beta}{\alpha-1}, \quad \alpha > 1,$$

and

$$\text{Var}(W) = \frac{\beta^2}{(\alpha-1)^2(\alpha-2)}, \quad \alpha > 2.$$

Note that, for $\alpha \leq 1$, the mean and higher moments do not exist and, for $1 < \alpha \leq 2$, the mean exists but the variance does not exist (Martz and Waller, 1991).

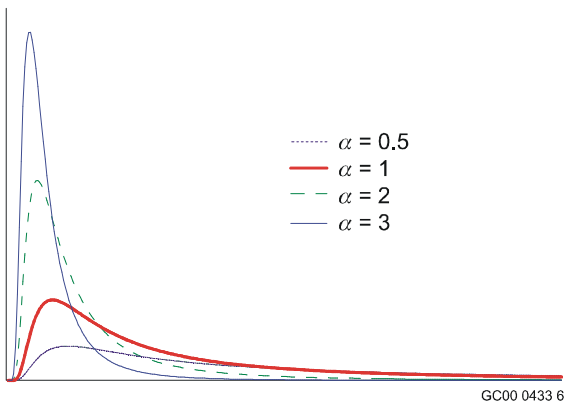


Figure B.14 Four inverted gamma densities, having the same scale parameter, β , and various

Figure B.14 shows four inverted gamma distributions, all having the same scale parameter, β , and having various shape parameters, α .

In the special case with $\alpha = n/2$ and $\beta = 1/2$, the distribution is called the **inverted chi-squared** distribution with n degrees of freedom. Values from this distribution are sometimes denoted $\chi^{-2}(n)$. This form of the distribution is often used in connection with a prior for σ^2 when the data are normally distributed.

A.7.8 Beta Distribution

Many continuous quantitative phenomena take on values that are bounded by known numbers a and b . Examples are percentages, proportions, ratios, and distance to failure points on items under stress. The **beta distribution** is a versatile family of distributions that is useful for modeling phenomena that can range from 0 to 1 and, through a transformation, from a to b .

The beta distribution family includes the uniform distribution and density shapes that range from decreasing to uni-modal right-skewed to symmetric to U-shaped to uni-modal left-skewed to increasing (Martz and Waller 1991). It can serve as a model for a reliability variable that represents the probability that a system or component lasts at least t units of time. The beta distribution is also widely used in Bayesian estimation and reliability analysis as a prior distribution for the binomial distribution parameter p that represents a reliability or failure probability.

The p.d.f. of a beta random variable, Y , is

$$f(y) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} y^{\alpha-1}(1-y)^{\beta-1},$$

for $0 \leq y \leq 1$, with the parameters $\alpha, \beta > 0$, and is denoted $\text{beta}(\alpha, \beta)$. The gamma functions at the front of the p.d.f. form a normalizing constant so that the density integrates to 1.

The mean and variance of the $\text{beta}(\alpha, \beta)$ random variable, Y , are

$$E(Y) = \frac{\alpha}{\alpha + \beta}$$

and

$$\text{Var}(Y) = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}.$$

Various beta distributions are shown in Figures B.15 and B.16. Figure B.15 shows beta densities with $\alpha = \beta$, and therefore with mean 0.5. When $\alpha < 1$, the density becomes infinite at 0.0, and when $\beta < 1$, the density becomes infinite at 1.0. When $\alpha = \beta = 1$, the density is uniform. When α and β are large, the density is approximately normal.

Figure B.16 shows densities with mean 0.1. Again, when $\alpha < 1$, the density becomes infinite at 0.0, and when $\alpha > 1$, the density is zero at 0.0. As the parameters α and β become large, the density approaches a normal distribution.

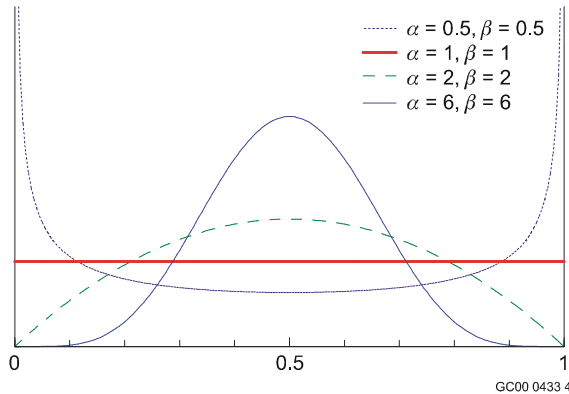


Figure B.15 Beta distributions with mean = 0.5.

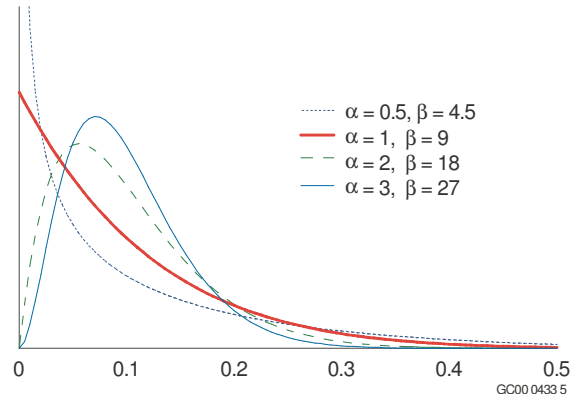


Figure B.16 Four beta distributions with mean 0.1.

Another parameterization of the beta distribution uses the parameters $x_0 = \alpha$ and $n_0 = \alpha + \beta$. This parameterization is used by Martz and Waller (1991) because it simplifies Bayes formulas and Bayesian estimation. The p.d.f. of a $\text{beta}(x_0, n_0)$ is

$$f(y) = \frac{\Gamma(n_0)}{\Gamma(x_0)\Gamma(n_0 - x_0)} y^{x_0-1}(1-y)^{n_0-x_0-1},$$

for $0 \leq y \leq 1$, with the parameters x_0 and n_0 satisfying

$$n_0 > x_0 > 0.$$

The mean and variance of the $\text{beta}(x_0, n_0)$ random variable, Y , are

$$E(Y) = \frac{x_0}{n_0}$$

and

$$\text{Var}(Y) = \frac{x_0(n_0 - x_0)}{n_0^2(n_0 + 1)}.$$

Percentiles of the beta distribution occur in the formula for a Bayes credible interval for p when a conjugate prior is used. Many software packages, including some commonly used spreadsheets, can calculate these percentiles.

A.7.9 Logistic-Normal Distribution

While not widely used in PRA, this distribution is commonly used for Bayesian inference in other fields of application, especially as a prior for the binomial parameter p when p could plausibly be fairly large. X has a

logistic-normal distribution if $\ln\left[\frac{X}{(1-X)}\right]$ is normally distributed with some mean μ and variance σ^2 . The

function $\ln\left[\frac{X}{(1-X)}\right]$ may appear strange, but it is common enough in some areas of application to have a name, the **logit** function. Therefore, the above statements could be rewritten to say that X has a logistic-normal distribution if $\text{logit}(X)$ is normally distributed.

Properties of the logistic-normal distribution are summarized here.

- Let $y = \ln[x/(1-x)]$. Then $x = e^y / (1 + e^y)$. This implies that x must be between 0 and 1.
- As x increases from 0 to 1, $y = \ln[x/(1-x)]$ increases monotonically from $-\infty$ to $+\infty$. Thus, y can be generated from a normal distribution with no problem of going outside the possible range.
- The monotonic relation between x and y means that the percentiles match. For example, the 95th percentile of Y is $\mu + 1.645\sigma$. Denote this by $y_{0.95}$. Therefore, the 95th percentile of X is

$$x_{0.95} = \exp\left[\frac{y_{0.95}}{[1 + \exp(y_{0.95})]}\right].$$

Alternatively, this can be written as

$$y_{0.95} = \ln\left[\frac{x_{0.95}}{(1-x_{0.95})}\right].$$

- If X is close to 0 with high probability, so that $X/(1-X)$ is close to X with high probability, then the logistic-normal and lognormal distributions are nearly the same.

The third bullet shows how to find the percentiles of a logistic-normal distribution. Unfortunately there is no equally easy way to find the moments, such as the mean or variance. Moments must be found using numerical integration.

Figure B.17 shows several logistic normal distributions that all have median 0.5. These correspond to a normally distributed y with mean $\mu = 0$ and with various values of σ . Figure B.18 shows several logistic normal distributions that all have median 0.1. These correspond to a normally distributed y with mean $\mu = -2.2 = \ln[0.1/(1-0.1)]$.

Note the general similarities to the beta distributions in Figures B.15 and B.16. Note also the differences: Logistic-normal distributions are characterized most easily by percentiles, whereas beta distributions are characterized most easily by moments. Also, the beta densities can be J-shaped or U-shaped, but the logistic-normal densities always drop to zero at the ends of the range.

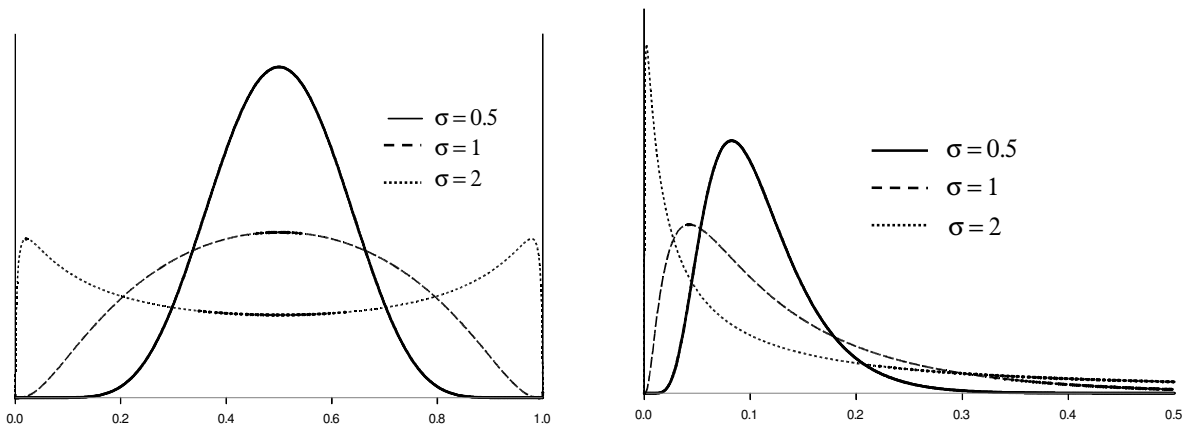


Figure B.17 Three logistic-normal densities with median = 0.5.

A.7.10 Student's *t* Distribution

The **Student's *t*** distribution is not used in a central way in PRA. However, it appears in a peripheral way in places in this handbook, when dealing with the parameters of a normal or lognormal distribution, or in large-sample situations when a distribution is approximated as normal or lognormal. Therefore, the basic facts are summarized here.

If *Z* has a standard normal distribution, *X* has a chi-squared distribution with *d* degrees of freedom, and *Z* and *X* are statistically independent, then

$$T = \frac{Z}{\sqrt{\frac{X}{d}}}$$

has a Student's *t* distribution with *d* degrees of freedom. Therefore, *T* has a distribution that is symmetrical about 0, and it can take values in the entire real line. If *d* is large, the denominator is close to 1 with high probability, and *T* has approximately a standard normal distribution. If *d* is smaller, the denominator adds extra variability, and the extreme percentiles of *T* are farther out than are the corresponding normal percentiles. Tables of the distribution are given in Appendix C.

Although not needed for ordinary work, the p.d.f. and first two moments of *T* are given here. (See many standard texts, such DeGroot 1975.) The p.d.f. is

$$f(t) = \frac{\Gamma\left[\frac{(d+1)}{2}\right]}{(d\pi)^{1/2}\Gamma\left(\frac{d}{2}\right)} \left[1 + \left(\frac{t^2}{2}\right)\right]^{-\frac{(d+1)}{2}}$$

Figure B.18 Three logistic-normal densities with median = 0.1.

If $d > 1$ the mean is 0. If $d > 2$ the variance is $\frac{d}{(d-2)}$. If $d \leq 2$ the variance does not exist. If $d = 1$, even the mean does not exist; in this case the distribution is called a **Cauchy distribution**.

A.8 Random Samples

When sampling from a distribution (or population), it is usually assumed that the n observations are taken at random, in the following sense. It is assumed that the n random variables X_1, X_2, \dots, X_n are independent. That is, the sample X_1, X_2, \dots, X_n taken from a distribution $f(x)$ has the joint p.d.f. h satisfying

$$h(x_1, x_2, \dots, x_n) = f(x_1) \cdot f(x_2) \cdot \dots \cdot f(x_n).$$

This follows the definition of independent random variables given in Section B.4.5. A sample taken in this way is called a **random sample**. (As elsewhere in this handbook, upper case letters denote random variables and lower case letters denote particular values, number.)

The random variables X_1, X_2, \dots, X_n forming such a random sample are referred to as being independent and identically distributed. If n is large enough, the sampled values will represent the distribution well enough to permit inference about the true distribution.

A.9 Sample Moments

Mathematical expectation and moments provide characteristics of distributions of random variables. These ideas can also be used with observations from a random sample from a distribution to provide **estimates** of the parameters that characterize that distribution.

A **statistic** is a function of one or more random variables that does not depend on any unknown parameters. A function of random variables that can be computed from the collected data sample is thus a statistic. Note that a function of random variables is also a random variable that has its own probability distribution and associated characteristics.

If X_1, X_2, \dots, X_n denote a random sample of size n from a distribution $f(x)$, the statistic

$$\bar{X} = \sum_{i=1}^n \frac{X_i}{n}$$

is the **mean of the random sample**, or the **sample mean** and the statistic

$$S^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}$$

is the **variance of the random sample**. Note that $n - 1$ is used as the denominator in the S^2 statistic to make the statistic an *unbiased* estimator of the population variance, σ^2 .

Similarly, the statistics defined by

$$m_r = \sum_{i=1}^n \frac{X_i^r}{n},$$

for $r = 1, 2, \dots$, are called the **sample moments**.

One of the common uses of statistics is estimating the unknown parameters of the distribution from which the sample was generated. The sample mean, or average, \bar{X} , is used to estimate the distribution mean, or population mean, μ , the sample variance, S^2 , is used to estimate the population variance, σ^2 , and so forth.

A.10 Statistical Inference

Since values of the parameters of a distribution are rarely known, the distribution of a random variable is rarely completely known. However, with some assumptions and information based on a random sample of observations from the distribution or population, values of the unknown parameters can often be estimated. Probabilities can then be calculated from the corresponding distribution using these parameter estimates.

Statistical inference is the area of statistics concerned with using sample data to answer questions and make statements about the distribution of a random variable from which the sample data were obtained. **Parameter estimators** are functions of sample data that are used to estimate the distribution parameters. Statements about parameter values are inferred from the specific sample to the general distribution of the random variable or population. This inference cannot be perfect; all inference techniques involve uncertainty. Understanding the performance properties of various estimators has received much attention in the statistics field.

For the purposes of this handbook, statistical inference procedures can be classified as follows:

- parameter estimation
 - estimation by a point value
 - estimation by an interval
- hypothesis testing
 - tests concerning parameter values
 - goodness-of-fit tests and other model-validation tests.

Parametric statistical inference assumes that the sample data come from a particular, specified family of distributions, with only the parameter values unknown. However, not all statistical inference is based on parametric families. In many cases, in addition to not knowing the distribution parameter values, the form of the parametric family of distributions is unknown. **Distribution-free**, also called **nonparametric**, techniques are applicable no matter what form the distribution may have. **Goodness-of-fit tests** are an important type of nonparametric tests that can be used to test whether a data set follows a hypothesized distribution.

A.11 Distribution Summary

Probability Density Function	Mean	Variance [†]	Excel Formulation	SAPHIRE Formulation
Beta				
$\pi(p) = \frac{p^{\alpha-1}(1-p)^{\beta-1}}{B(\alpha, \beta)}$ <p>where $B(\alpha, \beta) = \frac{\Gamma(\beta)\Gamma(\alpha)}{\Gamma(\alpha + \beta)}$</p>	$\frac{\alpha}{\beta + \alpha}$	$\frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}$	Density Not available Cumulative BETADIST(p, α, β) Inverse cumulative BETAINV(% , α, β)	1 st parameter Mean 2 nd parameter β
Gamma				
$\pi(\lambda) = \frac{\beta^\alpha \lambda^{\alpha-1} e^{-\beta\lambda}}{\Gamma(\alpha)}$	$\frac{\alpha}{\beta}$	$\frac{\alpha}{\beta^2}$	Density Gammadist(x, α, 1/β, false) Cumulative Gammadist(x, α, 1/β, true) Inverse cumulative Gammadist(% , α, 1/β)	1 st parameter Mean 2 nd parameter α (called "r")
Lognormal				
$\pi(\lambda) = \frac{1}{\sqrt{2\pi}\sigma\lambda} \exp\left[-\frac{(\ln \lambda - \mu)^2}{2\sigma^2}\right]$	$\exp\left(\mu + \frac{\sigma^2}{2}\right)$	$(\text{mean})^2 [\exp(\sigma^2) - 1]$		1 st parameter Mean 2 nd parameter Error factor
Uniform				
Same as the beta distribution with α=1 and β=1				

[†] Recall that the variance is equal to the standard deviation squared, or $\text{Var}(x) = \sigma^2$.

A.12 REFERENCES FOR APPENDIX B

- Ascher, H. and H. Feingold, 1984, *Repairable Systems Reliability – Modeling, Inference, Misconceptions and Their Causes*, Marcel Dekker, Inc., New York.
- Bain, L. J., and M. Engelhardt, 1991, *Introduction to Probability and Mathematical Statistics, 2nd Ed.*, PWS, Boston.
- Çinlar, E., 1975, *Introduction to Stochastic Processes*. Prentice-Hall, Englewood Cliffs, NJ.
- DeGroot, Morris H., 1975, *Probability and Statistics*, Addison-Wesley, Reading, MA.
- Derman, C., L. J. Gleser and I. Olkin, 1973, *A Guide to Probability Theory and Application*. Holt, Rinehart and Winston, NY.
- Esary, J. D. and F. Proschan, 1970, A Reliability Bound for Systems of Maintained, Interdependent Components in *Journal of the American Statistical Association*, Vol. 65, pp. 329-338.
- Esary, J. D. and F. Proschan, 1963, Coherent Structures of Non-Identical Components, *Technometrics*, **6**, pp. 191-209.
- Hogg, R. V. and A. T. Craig, 1995, *Introduction to Mathematical Statistics*. Macmillan, NY.
- Lee, Peter M., 1997, *Bayesian Statistics: An Introduction, Second Edition*, Arnold, a member of the Hodder Headline Group, London.
- Martz, H. F. and R. A. Waller, 1991, *Bayesian Reliability Analysis*. R. E. Krieger Publishing Co., Malabar, FL.
- Meyer, P. L., 1970, *Introductory Probability and Statistical Applications*. Addison-Wesley Pub. Co., Reading, MA.
- Pfeiffer, P. E. and D. A. Schum, 1973, *Introduction to Applied Probability*. Academic Press, NY.
- Snedecor, G. W. and W. G. Cochran, 1989, *Statistical Methods*. Iowa State University Press, Ames, IA.
- Thompson, W. A., Jr., 1981, On the Foundations of Reliability, in *Technometrics*, Vol. 23, No. 1, pp. 1-13.
- Vesely, W. E., F. F. Goldberg, N. H. Roberts and D. F. Haasl, 1981, *Fault Tree Handbook*. NUREG-0492. U.S. Nuclear Regulatory Commission, Washington, D.C.

B.1 WinBUGS

BUGS refers to **B**ayesian inference **U**sing **G**ibbs **S**ampling. WinBUGS is freely available software for implementing Markov chain Monte Carlo (MCMC) sampling. It is available in a Windows version from the BUGS Project website <http://www.mrc-bsu.cam.ac.uk/bugs/welcome.shtml> which also has links to the open source code (OpenBUGS) directly available at <http://mathstat.helsinki.fi/openbugs/Home.html> . Both programs are commonly referred to as WinBUGS or just BUGS.

WinBUGS is commonly used independently but can be called from the open-source statistical program R through the R2WinBUGS or BRUGS packages for R. For more information, see the R Project homepage, www.r-project.org.

B.1.1 Distributions Supported in WinBUGS

Over 20 distributions are supported in WinBUGS. Discrete and continuous univariate and multivariate distributions are supported. Common distributions used in PRA that are supported include:

- Binomial: `dbin(p,n)`
- Poisson: `dpois(mu)`
 - Users will often have $\mu = \lambda t$
- Exponential: `dexp(lambda)`
- Weibull: `dweib(v,lambda)`
- Gamma: `dgamma(r,mu)`
- Uniform: `dunif(a,b)`
- Beta: `dbeta(a,b)`
- Lognormal: `dlnorm(mu,tau)`

- $\tau = \frac{1}{\sigma^2}$

- $\sigma = \frac{\ln(EF)}{1.645}$

It is also possible to analyze user-defined distributions in WinBUGS. See the WinBUGS User Manual for information on how to do this.

B.1.2 WinBUGS Script

WinBUGS is a scripting language with a menu-driven interface. There are three parts to a WinBUGS script: the model description, data section, and initial values. A sample script is presented in Figure C.1.

```

Script to update frequency
model {
  #Model defined between {} symbols
  events ~ dpois(mu) #Poisson distribution for number of events
  mu <- lambda*time #Parameter of Poisson distribution
  lambda ~ dgamma(1.6,44) #Prior distribution for lambda
}

Data #Observed data
list(events=1,time=44)

```

The model description includes the likelihood function, prior distribution, and any derived quantities (e.g. system reliability).

The data can be written within the WinBUGS script or input from a separate text file. The initial values can be written within the WinBUGS script or from a separate text file.

Figure B.1: Sample WinBUGS Script

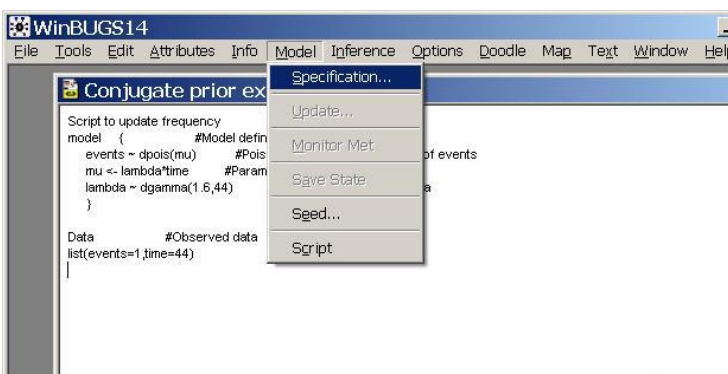
The use of the # character is for comments. It is highly encouraged to comment scripts.

B.1.3 Demonstration of WinBUGS via an Example Analysis

B.1.3.1 Conjugate Prior Example

Assume the frequency of an event (λ) has a $\text{gamma}(1.6, 44 \text{ yr})$ prior distribution. The likelihood function is a Poisson distribution with observed data consisting of 1 event in a 44-year period. The posterior distribution is $\text{gamma}(1.6+1, 44+44 \text{ yr})$ and the posterior mean of λ is $2.6/88 \text{ yr} = 0.03 \text{ per year}$.

The script in Figure C.1 is used to update the prior distribution for λ with the observed data. This can be written either in the WinBUGS new project screen or cut and pasted from a text editor. The **first** step is to check the model's syntax by highlighting the word "model" and selecting "Model – Specification" from the toolbar as in Figure C.2.



The Specification Tool will appear on the screen, after which the "Check Model" button should be selected. A status message will be displayed at the bottom left of the WinBUGS palette. If all is well, the message is returned, "model is syntactically correct".

Figure B.2: Checking the WinBUGS model for syntax.

Leave the Specification Tool on the screen (Figure C.3) and double-click (highlight) the word “list” in the data portion of the script, then select “load data”. A status message will be displayed at the bottom left of the WinBUGS palette that states “data loaded”. Next, select “compile” and WinBUGS reports “model compiled”. Note that models using multiple chains can be run. The number of chains in this example is one.

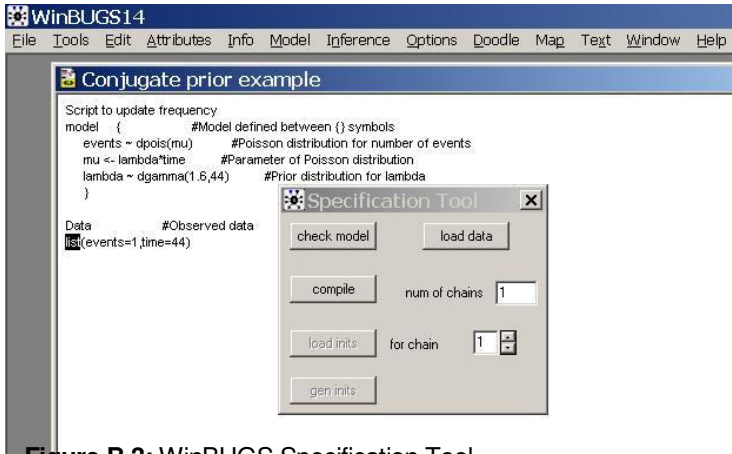


Figure B.3: WinBUGS Specification Tool.

The last task to do with the Specification Tool is to load the initial values. For the conjugate prior example, the model is such a simple one that we can let WinBUGS generate initial values. Select “gen inits” to have WinBUGS generate the initial values and a message of “initial values generated, model initialized” should appear in the status bar.

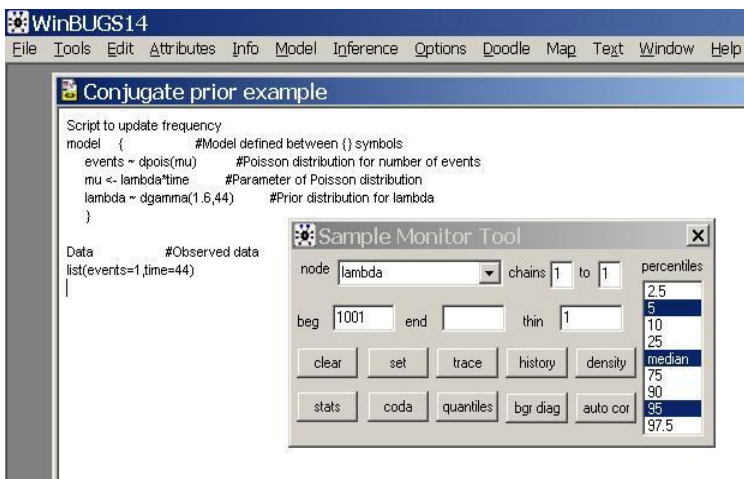


Figure B.4: WinBUGS Sample Monitor Tool.

The next step in setting up the analysis is to select the variables to monitor. For the conjugate prior example, we are interested in the frequency, or rate of occurrence (lambda). Close the Specification Tool and select “Inference” and then “Samples” from the toolbar. The Sample Monitor Tool (Figure C.4) will pop up. Type the variable name “lambda” in the “node” box. Enter 1001 in the “beg” box as the iteration at which to start monitoring lambda. This will allow this particular model enough iterations to converge. A discussion on convergence will follow. Select the default of 10,000,000 as the “end” value. Save the setting by clicking the “set” button. Multiple variables and chains can be monitored, although we will only monitor lambda for this example.

Close the Sample Monitor Tool and select “Model” and “Update” from the toolbar. The Update Tool appears on the screen (Figure C.5). Enter in the number of updates to perform, in this case 10,000 is the quantity desired.

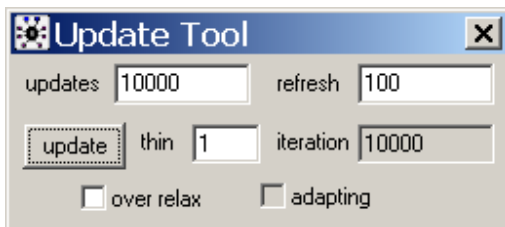


Figure B.5: WinBUGS Update Tool.

The refresh rate will control the display in the iteration box during an update. A lower number for the refresh rate will be informative, but will also slow the sampling time. Select “update”. WinBUGS will report that the “model is updating” in the status bar and the iteration window will display the iterations by increments specified in the refresh rate. When the model update has completed a message will appear in the status bar that indicates “updates took X s”.

Close the Update Tool when the sampling is complete and re-open the Sample Monitor Tool by selecting “Inference → Samples” from the toolbar. Select “lambda” from the drop-down list of monitored nodes. Highlight the percentiles 5, median (50%), and 95 by holding down the Ctrl key while selecting with the left mouse button. To display a graph of the posterior density (Figure C.6) select “density”.

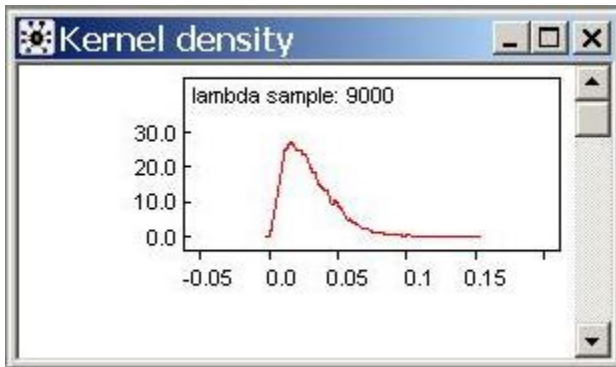


Figure B.6: Posterior distribution density.

Next select the “stats” button to get the posterior mean and percentiles selected. These results shown in Figure C.7 can be compared to hand/spreadsheet calculation results of:

- Posterior mean = 0.029/yr
- 5th percentile = 0.007/yr
- 95th percentile = 0.065/yr

Note that the results, both the density plots and the statistics, may be selected and copied (via the CTRL+C key combination in Windows) for pasting into other programs.

node	mean	sd	MC error	5.0%	median	95.0%	start	sample
lambda	0.02962	0.01835	1.916E-4	0.007385	0.0259	0.06418	1001	9000

Figure B.7: Posterior statistics for lambda.

B.1.4 Monitoring Convergence in WinBUGS

WinBUGS uses Markov chain Monte Carlo (MCMC) sampling to generate values directly from the target posterior distribution. However, it takes some time for the sampling to converge to the posterior distribution; any values sampled prior to convergence should not be used to estimate parameter values. For simple problems involving one parameter, such as p in the binomial distribution or λ in the Poisson distribution, 1,000 iterations will be more than sufficient for convergence. In more complicated problems, which usually involve inference for more than one parameter, this may not be the case, and the user will have to check for convergence. This section discusses practical checks that the user can do, using features built into WinBUGS.

B.1.4.1 Running Multiple Chains

For problems with more than one parameter, at least two chains should be run, with starting values that are dispersed around the estimated node of the posterior distribution. Usually the analysis will not be very sensitive to the initial values selected for the chains, but this is not always the case. An expert analyst should be consulted in such cases, which arise especially often when modeling population variability, as there are sophisticated approaches that can be used to derive initial values.

After user-specified initial values are loaded, it may still be necessary to have WinBUGS generate remaining initial values.

After the model has been compiled (remember to specify the number of chains before compiling the model) and initial values have been loaded, specify the nodes to be monitored. All parameters should be monitored and convergence should be checked for each of these monitored nodes. Now run 1,000 samples and select History from the Inference menu to generate a trace of the first 1,000 samples for each monitored node. A plot like the one shown below, in which the two chains are well mixed, is indicative of convergence.

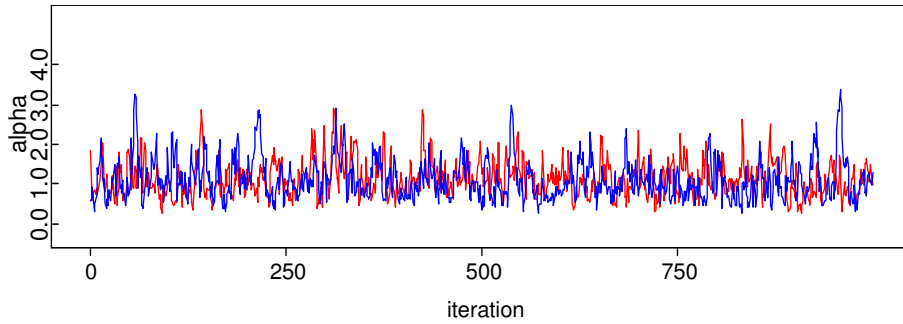


Figure B.8: History plot showing two well-mixed chains, indicative of convergence.

In contrast, the figure below shows a case in which the chain are not well mixed, which indicates that more iterations must be run to reach convergence.

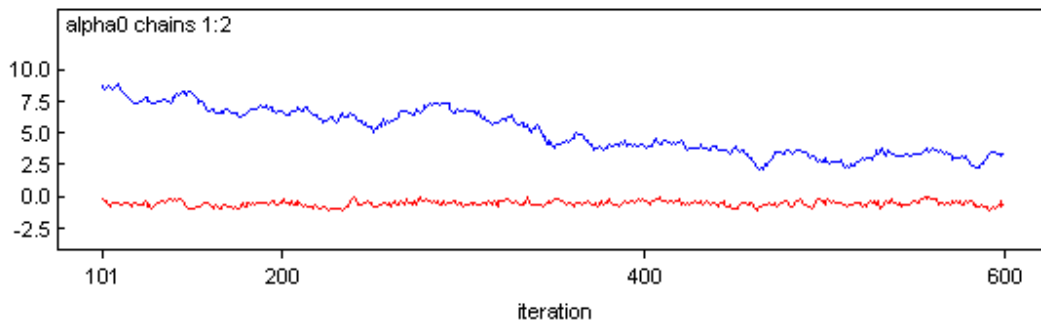


Figure B.9: History plot showing two poorly-mixed chains, indicative of failure to converge.

B.1.4.2 Brooks-Gelman-Rubin (BGR) Convergence Diagnostic

WinBUGS has a built-in convergence diagnostic that can be used in conjunction with the history plots shown above to help the user decide when burn-in samples have been taken. It requires multiple chains, and is based on an analysis of the variance within- and between-chains. If the chains have converged, all chains should have approximately the same within-chain variance, and this should be approximately equal to the between-chain variance estimate. The BGR diagnostic in WinBUGS looks at a ratio of these estimates, which is normalized to equal one when the chains have converged.

To implement the BGR diagnostic, run at least two chains, as described above, then select “bgr diag” from the Inference menu. The resulting plot will have three curves. The estimate of the within-chain variance is shown in blue, the between-chain estimate is in green, and the BGR ratio is shown in red. The BGR ratio is expected to start out greater than one. The heuristic is that this ratio should be less than about 1.2 for convergence. However, the between-chain and within-chain estimates should be stable. Double-clicking with the left mouse button on the BGR graph, followed by <Ctrl>-left click will bring up a table of the values over the history.

The first graph below shows a typical BGR plot for a problem that has converged. The second shows a plot where convergence may have been reached just at the end of the 1,000 samples, but more burn-in iterations should be taken to ensure convergence before estimating parameter values.

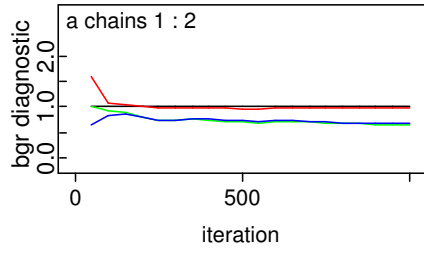


Figure B.10: BGR plot illustrating convergence.

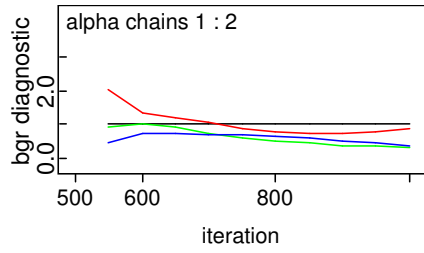


Figure B.11: BGR plot illustrating need for more burn-in samples.

Contents

Introduction	2
Inference for a Single Parameter (3.2).....	2
Binomial Distribution with Conjugate Prior (3.2.1).....	2
Poisson Distribution with Conjugate Prior (3.2.2).....	2
Exponential Distribution with Conjugate Prior (3.2.3).....	3
Binomial Distribution with Nonconjugate Prior (3.2.1).....	3
Poisson Distribution with Nonconjugate Prior (3.2.2).....	4
Exponential Distribution with Nonconjugate Prior (3.2.3).....	5
Noninformative Priors (3.2.1 – 3.2.3)	5
Model Validation (3.3).....	5
Use of Summary Statistics.....	5
Deviance Information Criterion	7
Time-Trend Models for p and λ (3.4).....	8
Hierarchical Bayesian Modeling of Population Variability (3.5).....	9
More Complex Models for Time to Failure – Nonrepairable Failure (3.6).....	11
Gamma Distribution	11
Weibull Distribution	11
Lognormal Distribution.....	12
Modeling Failure with Repair (3.7).....	13
Treatment of Censoring (3.8)	14
References	15

Introduction

This appendix provides details of the salient mathematics behind the WinBUGS scripts found in the main body of the guidebook. The scripts in the guidebook are intended to be the primary tools to help an analyst think through a Bayesian inference problem. However, some analysts may gain additional benefit from seeing the supporting equations, at least in those cases where it is possible to write them down. Furthermore, some problems in Bayesian inference are amenable to solution by numerical tools other than the Markov chain Monte Carlo (MCMC) sampling employed by WinBUGS, and analysts may feel more comfortable using such tools. For example, inference for a single parameter with a nonconjugate prior can be handled via widely available numerical integration routines, such as trapezoidal integration, Simpson's Rule, or quadrature. This appendix presents the relevant equations that an analyst would need to solve for these types of problems, but does not cover alternative numerical solution techniques (e.g., quadrature); an analyst who wishes to employ such tools must take responsibility for their proper use.

Inference for a Single Parameter (3.2)

This section describes inference for p in the binomial distribution, typically used for failure on demand; λ in the Poisson distribution, used for initiating events and operating failures; and λ in the exponential distribution, used for random durations, such as time to recover failed equipment.

Binomial Distribution with Conjugate Prior (3.2.1)

The likelihood function is the binomial distribution, given by

$$f(x | p) = \binom{n}{x} p^x (1-p)^{n-x} \quad (1)$$

The conjugate prior is a beta distribution, with density function

$$\pi_0(p) = \frac{p^{\alpha-1} (1-p)^{\beta-1}}{B(\alpha, \beta)} \quad (2)$$

$B(\alpha, \beta)$ is the beta function, defined as follows for $\alpha, \beta > 0$:

$$B(\alpha, \beta) = \int_0^1 x^{\alpha-1} (1-x)^{\beta-1} dx \quad (3)$$

If x failures are observed in n demands, the posterior distribution is easily shown to be a beta distribution with parameters $\alpha + x$ and $\beta + n - x$. The mean and variance can be found in closed form, but credible intervals must be found numerically, using a spreadsheet tool, as described in the guidebook, or another routine.

Poisson Distribution with Conjugate Prior (3.2.2)

The likelihood function is the Poisson distribution, given by

$$f(x | \lambda) = \frac{(\lambda t)^x e^{-\lambda t}}{x!} \quad (4)$$

The conjugate prior is a gamma distribution, with density function

$$\pi_0(\lambda) = \frac{\beta^\alpha \lambda^{\alpha-1} e^{-\beta\lambda}}{\Gamma(\alpha)} \quad (5)$$

$\Gamma(\alpha)$ is the gamma function, defined for $\alpha > 0$ by

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx \quad (6)$$

If x events are observed in time t , the posterior is easily shown to be a gamma distribution with parameters $\alpha + x$ and $\beta + t$. As above, the mean and variance can be written in closed form but credible intervals must be found numerically, using a spreadsheet tool, as described in the guidebook, or another routine.

Exponential Distribution with Conjugate Prior (3.2.3)

Inference in this case is essentially identical to inference for the Poisson distribution, but with a likelihood function that changes as follows. We are now observing times to occurrence. We assume that each time, T_i , $i = 1, 2, \dots, n$, is independently and identically distributed as an exponential random variable with rate λ . Thus, the density function for each T_i can be written

$$f(t_i | \lambda) = \lambda e^{-\lambda t_i} \quad (7)$$

Because of the independence assumption, the joint distribution for the sample factors, allowing us to write the likelihood function as the product of n exponential density functions:

$$f(t_1, t_2, \dots, t_n | \lambda) = \prod_{i=1}^n f(t_i | \lambda) = \lambda^n \exp(-\lambda \sum_{i=1}^n t_i) \quad (8)$$

As with the Poisson likelihood above, the conjugate prior for λ is again a gamma distribution, with density given by Eq. (5). The posterior is easily shown to be a gamma distribution with parameters $\alpha + n$ and $\beta + \sum t_i$. As above, the mean and variance can be written in closed form but credible intervals must be found numerically, using a spreadsheet tool, as described in the guidebook, or another routine.

Binomial Distribution with Nonconjugate Prior (3.2.1)

The lognormal prior for p in the binomial distribution is parameterized in terms of the logarithmic mean, μ , and logarithmic precision, τ , defined as $\tau = 1/\sigma^2$. The lognormal density function in this parameterization is given by

$$\pi_o(p) = \sqrt{\frac{\tau}{2\pi}} \frac{1}{p} \exp\left[-\frac{\tau}{2}(\log(p) - \mu)^2\right] \quad (9)$$

The likelihood function is the binomial distribution, given by Eq. (1) above. The posterior distribution is obtained from Bayes' Theorem:

$$\pi_1(p|x) = \frac{f(x|p)\pi_o(p)}{\int_0^1 f(x|p)\pi_o(p)dp} \quad (10)$$

The denominator of this equation does not have an analytic solution and must be evaluated numerically. As will be true for all of the cases utilizing a nonconjugate prior, the posterior distribution does not have a closed form, and is not a member of any particular functional family (e.g., lognormal). However, it can be approximated with a particular functional distribution, if desired.

The posterior mean of p is given by

$$\int_0^1 p\pi_1(p|x)dp \quad (11)$$

Again, this equation does not have an analytic solution and must be evaluated numerically.

Percentiles can be found by numerically solving equations such as the following:

$$\int_0^{p_\alpha} \pi_1(p|x)dp = \alpha \quad (12)$$

The logistic-normal prior for p is also nonconjugate. The density is given by

$$\pi_o(p) = \sqrt{\frac{\tau}{2\pi}} \frac{1}{p(1-p)} \exp\left[-\frac{\tau}{2}\left(\log\left(\frac{p}{1-p}\right)\right)^2\right] \quad (13)$$

where μ and τ are the mean and precision of the underlying normal distribution. The posterior mean and credible intervals are obtained by substituting the logistic-normal density into Eqs. (11) and (12) above.

Poisson Distribution with Nonconjugate Prior (3.2.2)

The unknown parameter is now λ , the intensity of the Poisson process. Simply replace p with λ in Eq. (9) above for the lognormal prior. The likelihood function is the Poisson distribution, given by Eq. (4). The posterior distribution is given by Bayes' Theorem:

$$\pi_1(\lambda|x) = \frac{f(x|\lambda)\pi_o(\lambda)}{\int_0^\infty f(x|\lambda)\pi_o(\lambda)d\lambda} \quad (14)$$

Again, this equation does not have an analytic solution and must be evaluated numerically. The posterior mean and credible intervals are found numerically as in Eqs. (11) and (12) above using this posterior distribution.

Exponential Distribution with Nonconjugate Prior (3.2.3)

As mentioned above, inference in this case is essentially the same as Poisson inference, but with a likelihood function given by Eq. (8). The posterior distribution is obtained by substituting this likelihood function for the Poisson likelihood function into Eq. (14) above.

Noninformative Priors (3.2.1 – 3.2.3)

The guidebook discusses only the Jeffreys prior, which is obtained by applying Jeffreys rule:

$$\pi_{NI}(\theta) \propto \sqrt{J(\theta)} \quad (15)$$

$J(\theta)$ depends upon the likelihood function:

$$J(\theta) = -E \left[\frac{d^2 (\log(f(x|\theta)))}{d\theta^2} \right] \quad (16)$$

For the binomial distribution, this leads to a beta(0.5, 0.5) distribution as the Jeffreys prior. For the Poisson distribution, the Jeffreys prior is an improper distribution proportional to $\lambda^{-0.5}$, which, for purposes of Bayesian updating, can be thought of as a (conjugate) gamma(0.5, 0) distribution, yielding a gamma posterior distribution with parameters $x + 0.5$ and t if x events are observed in time t . For the exponential distribution, the Jeffreys prior is an improper distribution proportional to $1/\lambda$, which can be thought of as a (conjugate) gamma(0, 0) distribution for Bayesian updating. Note that the Jeffreys prior is not often used as a noninformative prior in multiparameter inference problems. Instead, as illustrated in the guidebook, one usually uses independent diffuse priors, typically uniform distributions or gamma distributions with both parameters close to zero.

Model Validation (3.3)

The posterior predictive distribution is the posterior analog of the marginal distribution in the denominator of Bayes' Theorem. It is the distribution of future data, conditional upon the observed data. It is given by

$$f(x_{rep} | x) = \int_{\Theta} f(x_{rep} | \theta) \pi_1(\theta | x) d\theta \quad (17)$$

If the likelihood function and prior distribution form a conjugate pair, then the posterior predictive distribution can be written down in closed form. For checking the predictive validity of simple models, the tail area probability, $P(X_{rep} \geq x)$ can be used, with small tail area probability being indicative of a model with poor predictive validity.

Use of Summary Statistics

One can also use summary statistics derived from the posterior predictive distribution. This is useful when examining situations more complicated than simple binomial or Poisson event counts. Example 4 presents valve leakage data for a nine-year period and asks whether the data appear to be reasonably homogeneous across time, allowing a pooled-data estimate of p , the probability of leakage during the weekly leak test.

We use a Bayesian analog of the commonly used frequentist chi-square test. The frequentist test estimates p under the null hypothesis that the data are homogeneous, then calculates the following statistic, which has approximately a chi-square distribution with $N - 1$ degrees of freedom if the null hypothesis is true:

$$\chi^2 = \sum_{i=1}^N \frac{(x_i - n_i \hat{p})^2}{n_i \hat{p} (1 - \hat{p})} \quad (18)$$

In this equation, the pooled estimate under the null hypothesis is given by

$$\hat{p} = \frac{\sum x_i}{\sum n_i} \quad (19)$$

If the observed value of χ^2 is too large, then the null hypothesis is rejected.

For the Bayesian analog we calculate two χ^2 summary statistics: one for the observed data and one for replicated data from the posterior predictive distribution. These statistics are defined below. The pooled estimate of p is the posterior distribution for p obtained by updating the Jeffreys prior, a beta(0.5, 0.5) distribution, with the pooled data, $\sum x_i$ and $\sum n_i$. This yields a beta distribution with parameters $\sum x_i + 0.5$ and $\sum n_i - \sum x_i + 0.5$.

$$\chi_{obs}^2 = \sum_{i=1}^N \frac{(x_i - n_i \tilde{p})^2}{n_i \tilde{p} (1 - \tilde{p})} \quad (20)$$

$$\chi_{rep}^2 = \sum_{i=1}^N \frac{(x_{rep,i} - n_i \tilde{p})^2}{n_i \tilde{p} (1 - \tilde{p})} \quad (21)$$

Note that both χ_{obs}^2 and χ_{rep}^2 are now random variables, each possessing a probability distribution function. If the model (likelihood + prior) is "perfect", then the replicated and observed data will be exactly the same, and the distributions for χ_{obs}^2 and χ_{rep}^2 will coincide. If this is the case, we will have $P(\chi_{rep}^2 \geq \chi_{obs}^2) = 0.5$. If this probability, which is estimated by Monte Carlo sampling in WinBUGS via the step() function, is too small or too large, then the model is judged to be poor at replicating the observed data. A plot of the replicated data versus the observed data can be helpful in diagnosing where the problems lie.

A similar summary statistic can be used to test the homogeneity of Poisson data across bins. The above equations are simply replaced by the following.

$$\chi_{obs}^2 = \sum_{i=1}^N \frac{(x_i - \tilde{\lambda} t_i)^2}{\tilde{\lambda} t_i} \quad (22)$$

$$\chi_{rep}^2 = \sum_{i=1}^N \frac{(x_{rep,i} - \tilde{\lambda} t_i)^2}{\tilde{\lambda} t_i} \quad (23)$$

As before, the pooled estimate of λ is the posterior distribution for λ obtained by updating the Jeffreys prior, an improper distribution proportional to $\lambda^{-0.5}$, with the pooled data, $\sum x_i$ and $\sum t_i$, yielding a gamma distribution with parameters $\sum x_i + 0.5$ and $\sum t_i$.

With an exponential likelihood function, the usual summary statistic is the sum of the observed times, $\sum t_i$. However, if a Jeffreys prior is used for λ (an improper prior proportional to λ^{-1}), the sum of the observed times, because it is a sufficient statistic for λ , will be at the 50th percentile of the distribution of the sum of the

replicated times from the posterior predictive distribution, so another summary statistic is needed. The Cramer-von Mises statistic uses the fact that the cumulative distribution function, when viewed as a random variable, has a uniform(0, 1) distribution. It uses the observed times, ranked from smallest to largest. The statistic is calculated as

$$CVM = \frac{1}{12n} + \sum_{i=1}^n \left[F(t_{i:n}) - \frac{2i-1}{2n} \right]^2 \quad (24)$$

where $F(\bullet)$ is the putative cumulative distribution function, $1 - e^{-\lambda t}$ for the exponential distribution. Again, the Bayesian analog computes an observed and replicated statistic (distribution), and compares the overlap, with values near 0 or 1 indicating a model that is poor at replicating the observed data. A Jeffreys prior is used for λ to remove the influence of the prior and focus on the likelihood portion of the model. In the WinBUGS scripts that implement this summary statistic, the factor $1/(12n)$ is dropped, as it appears in both the observed and replicated statistics, and cancels out when the difference is computed for each Monte Carlo iteration via the `step()` function in WinBUGS.

For more details on the use of summary statistics from the posterior predictive distribution, see (Martz & Picard, 1998) and (Gelman & al., 1996).

Deviance Information Criterion

Frequentist statistics has long made use of the deviance, defined as $\text{deviance} = -2\log(\text{likelihood})$, as a measure of how well a model fits observed data. If the data are normally distributed with known variance then the deviance can be written as

$$D \propto \sum \frac{(x_i - \mu)^2}{\sigma^2} \quad (25)$$

Thus, the deviance generalizes the chi-square statistic used earlier.

As discussed by (Gelman & al., 2004) the expected deviance, computed by averaging the deviance over the true sampling distribution, equals twice the Kullback-Leibler information (up to a fixed constant). As the sample size grows large, the model with the lowest Kullback-Leibler information, and therefore the lowest expected deviance, has the highest posterior probability.

Note that the discrepancy between the model and the observed data also depends on the unknown parameter θ (μ in the equation above). Therefore, one can average the deviance over the posterior distribution of θ , using draws from the simulated posterior distribution of θ to implement this in a numerical framework.

The deviance averaged over the posterior distribution of θ will be different than the deviance calculated for a single value of θ , such as the posterior mean. The difference represents the effect of model fitting, and can be thought of as the effective number of parameters in a Bayesian model. This difference represents the decrease in deviance (i.e., expected improvement in model fit) expected from estimating the parameters in the model.

One can also estimate the expected deviance in applying the fitted model to replicated data from the posterior predictive distribution, with the average carried out over this distribution. This expected predictive deviance is estimated by the deviance information criterion (DIC):

$$DIC = 2D_{avg} - D_{\hat{\theta}} \quad (26)$$

DIC is computed by WinBUGS. For more details, see (Spiegelhalter & al., 2002)

Time-Trend Models for p and λ (3.4)

These are special cases of generalized linear models (GLM), in which time is the explanatory variable. See (Gelman & al., 2004) for more details on Bayesian inference for GLMs. For p in the binomial distribution, we use the canonical link function, which is $\text{logit}(p)$. We then express $\text{logit}(p)$ as a linear function of time, yielding

$$\text{logit}(p) = \log\left(\frac{p}{1-p}\right) = a + bt \quad (27)$$

The likelihood function then becomes

$$f(x_1, x_2, \dots, x_n | a, b) = \prod_{i=1}^N \binom{n_i}{x_i} \left[\frac{\exp(a + bt_i)}{1 + \exp(a + bt_i)} \right]^{x_i} \left[\frac{1}{1 + \exp(a + bt_i)} \right]^{n_i - x_i} \quad (28)$$

The analyst must place a prior on the parameters, a and b . From Bayes' Theorem, we then have

$$\pi_1(a, b | x_1, x_2, \dots, x_n) = \frac{f(x_1, x_2, \dots, x_n | a, b) \pi_0(a, b)}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x_1, x_2, \dots, x_n | a, b) \pi_0(a, b) da db} \quad (29)$$

It is typical, although not necessary, to use independent, improper Laplace priors for a and b (there are flat priors over the real axis, implemented as `dflat()` in WinBUGS). In this instance, the posterior distribution reduces to

$$\pi_1(a, b | x_1, x_2, \dots, x_n) = \frac{f(x_1, x_2, \dots, x_n | a, b)}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x_1, x_2, \dots, x_n | a, b) da db} \quad (30)$$

In WinBUGS, initial values must be supplied for a and b , and these can be based on frequentist estimates, obtained by some standard procedure such as iteratively reweighted least-squares regression. Note that convergence to the joint posterior distribution is usually not a problem for these types of models, and the joint posterior distribution is proper, even with improper priors on a and b .

One can predict p in an upcoming year by using the posterior distribution for a and b :

$$p(t_{new}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{\exp(a + bt_{new})}{1 + \exp(a + bt_{new})} \pi_1(a, b | x_1, x_2, \dots, x_n) da db \quad (31)$$

For λ in the Poisson distribution, the canonical link function is the natural logarithm, yielding $\log(\lambda) = a + bt$. The likelihood function becomes

$$f(x_1, x_2, \dots, x_n | a, b) = \prod_{i=1}^n \frac{[\exp(a + bt_i)]^{x_i} \exp[-\exp(a + bt_i)]}{x_i!} \quad (32)$$

As for p above, the analyst must place a prior on the parameters, a and b . Again, it is typical, although not necessary, to use independent, improper Laplace priors for a and b . Initial values must be supplied for a and b , and these can again be based on frequentist estimates, obtained by some standard procedure such as iteratively reweighted least-squares regression. As for p above, convergence to the joint posterior distribution is usually not a problem for these types of models and the posterior distribution is proper.

One can use other link functions, if desired, although the differences in the numerical results should be small. Other link functions for p in the binomial distribution are the probit and the complementary log-log, defined as follows:

Probit: $\Phi^{-1}(p) = a + bt$, where Φ^{-1} is the inverse of the standard normal cumulative distribution function.

Complementary log-log: $\log(-\log(1 - p)) = a + bt$

For λ in the Poisson distribution, other link functions are the linear and power-law, defined as follows:

$$\text{Linear: } \lambda = a + bt \quad (33)$$

$$\text{Power-law: } \lambda = \lambda_0 t^\beta \quad (34)$$

Hierarchical Bayesian Modeling of Population Variability (3.5)

The Bayesian approach when there appears to be significant heterogeneity in the data is to utilize a hierarchical prior distribution that reflects this heterogeneity, or variability as it is sometimes called. As discussed by (Siu & Kelly, 1998), treatment of variability that can exist among sources of data is important if total uncertainty, including source-to-source variability, is to be properly represented by the resulting posterior distribution. (Siu & Kelly, 1998) discusses hierarchical Bayesian analysis briefly, but (due to software and computer limitations that existed at that time) focuses on two other methods: an older approach called two-stage Bayes (Kaplan, 1983), and (parametric) empirical Bayes. As pointed out by (Siu & Kelly, 1998), the two-stage Bayes approach can be viewed as a particular example of the more general hierarchical Bayes, and empirical Bayes can be viewed as an approximation to hierarchical Bayes. Thus, with the advent of MCMC software such as WinBUGS, the hierarchical Bayesian approach represents an achievable, fully Bayesian state of the art. Empirical Bayes, which is not fully Bayesian, still has a role to play, however, as will be discussed below.

Hierarchical Bayes is so-named because it utilizes hierarchical or multistage prior distributions. In the hierarchical Bayes framework, the prior distribution for the parameter of interest, denoted $\pi_1(\theta)$, is written as

$$\pi_0(\theta) = \int_{\Phi} \pi_1(\theta | \varphi) \pi_2(\varphi) d\varphi \quad (35)$$

In this equation, $\pi_1(\theta|\varphi)$ is the first-stage prior, representing the source-to-source variability in θ for a given value of φ (note that φ is typically a vector), and $\pi_2(\varphi)$, called the hyperprior, is the distribution representing the uncertainty in φ , whose components are called hyperparameters. The first-stage prior, $\pi_1(\theta|\varphi)$, is usually assumed to be of a particular parametric form, such as a gamma or lognormal distribution for λ or a beta or logistic-normal distribution for p . It is also typical, although not necessary, to use independent diffuse hyperpriors for the components of φ . Nothing limits the analysis to two stages, although the use of more than two stages has been rare in applications.

WinBUGS samples directly from the joint posterior distribution of θ and φ , denoted $\pi_1(\theta, \varphi|x)$. We are interested in the marginal posterior distribution of θ , $\pi_1(\theta|x)$, which is obtained by integrating over φ .

As a specific example, consider a Poisson process with intensity λ_i for each source, so $\theta = (\lambda_1, \lambda_2, \dots, \lambda_n)^\top$. To model variability in λ , we will use a gamma first-stage prior, with parameters α and β . We then have $\varphi = (\alpha, \beta)^\top$. We choose independent, diffuse hyperpriors for α and β , so $\pi_2(\varphi)$ becomes the product of these diffuse hyperpriors. The posterior distribution of λ_i , representing source-to-source variability, will be given by an average of the posterior distribution for λ_i , conditional upon α and β (a gamma distribution), weighted by the joint posterior distribution for α and β . The λ s are conditionally independent, given α and β . Thus, the posterior distribution for λ_i is a continuous mixture of gamma distributions:

$$\pi(\lambda_i | x, t) = \iint \pi_1(\lambda_i | x, t, \alpha, \beta) \pi_2(\alpha, \beta | x, t) d\alpha d\beta \quad (36)$$

The overall distribution describing source-to-source variability in λ , the posterior predictive distribution for λ , will be a similar mixture of gamma distributions.

With regard to carrying out this analysis with MCMC, convergence to the joint posterior distribution of $((\lambda_1, \lambda_2, \dots, \lambda_n, \alpha, \beta))$ can be problematic in these types of problems, especially when using nearly improper gamma hyperpriors for α and β . A complicating factor is the correlation between α and β that arises in the joint posterior distribution. It is sometimes helpful to reparameterize the problem to reduce this correlation. A practical parameterization here is to use the mean, equal to α/β , and the coefficient of variation, defined as the ratio of the standard deviation to the mean, which in this case is

$$\delta = \frac{1}{\sqrt{\alpha}} \quad (37)$$

For the mean, we use a gamma(10^{-4} , 10^{-4}) hyperprior. This is a very diffuse distribution which approximates an improper flat prior over the positive real axis. WinBUGS requires proper prior distributions, so a flat prior over the positive real axis cannot be used. For the coefficient of variation we use a diffuse gamma(10^{-4} , 10^{-4}) hyperprior, in general, or an exponential distribution with mean of 1. The exp(1) hyperprior is reasonable for the example in the guidebook because λ is expected to display significant uncertainty, meaning that the coefficient of variation should be ≈ 1 . As discussed by (Siu & Kelly, 1998), the exponential distribution is a maximum entropy distribution for the case where the variable of interest can take on any positive value, and a mean value is specified as a constraint. Such a distribution will not be appropriate for all problems. Choosing a suitably diffuse, yet proper hyperprior can be a delicate issue in some applications. See (Kelly & Atwood, 2008) for more details.

Multiple MCMC chains are used to facilitate monitoring convergence, both graphically and via the Brooks-Gelman-Rubin diagnostic (Brooks & Gelman, 1998), which requires at least two chains. Each of the chains must be given an initial value of the coefficient of variation and the mean of λ . As discussed by (Gelman, 1996), the starting values should be over-dispersed around the mode of the posterior distribution of the hyperparameters. Empirical Bayes, discussed at length by (Siu & Kelly, 1998), can provide a reasonable estimate for the location of the posterior mode in some cases, as it provides the maximum likelihood estimates of the marginal likelihood function. In this example, because the first-stage gamma prior is conjugate to the Poisson likelihood, the marginal likelihood can be written down in closed form:

$$L(\alpha, \beta) = \prod_i \frac{\Gamma(\alpha + x_i)}{x_i! \Gamma(\alpha)} \left(\frac{t_i}{\beta}\right)^x \left(1 + \frac{t_i}{\beta}\right)^{-(\alpha+x_i)} \quad (38)$$

The natural logarithm of this equation is maximized numerically to find estimates of α and β , and these estimates are used to approximate the mode of the posterior distribution for α and β . Initial values for the chains are then dispersed around this approximate mode.

More Complex Models for Time to Failure – Nonrepairable Failure (3.6)

In these models, the simple exponential distribution for a random duration is replaced by another distribution that allows for a time-dependent Poisson intensity. Three alternatives are described: gamma, Weibull, and lognormal. Each of these is a two-parameter distribution, so there are now two parameters for which inference must be carried out.

Gamma Distribution

In this case, each observed duration is treated as a random sample from a gamma distribution, with density

$$f(t_i | \alpha, \beta) = \frac{\beta^\alpha t_i^{\alpha-1} e^{-\beta t_i}}{\Gamma(\alpha)} \quad \alpha, \beta > 0 \quad (39)$$

In this equation, $\Gamma(\alpha)$ is the gamma function, defined for $\alpha > 0$ by Eq. (6). Note that, when $\alpha = 1$, the gamma distribution reduces to the exponential distribution. If $\alpha > 1$, then the intensity, defined as $f(t)/[1 - F(t)]$, is an increasing function of time. If $\alpha < 1$, then the intensity is a decreasing function of time.

If a random sample of n durations is observed, the likelihood function becomes

$$\begin{aligned} f(t_1, t_2, \dots, t_n | \alpha, \beta) &= \prod_{i=1}^n f(t_i | \alpha, \beta) \\ &= \frac{\beta^{n\alpha} \prod_{i=1}^n t_i^{\alpha-1} \exp(-\beta \sum_{i=1}^n t_i)}{[\Gamma(\alpha)]^n} \end{aligned} \quad (40)$$

The analyst must specify a prior distribution for α and β . In most of the scripts, we use independent, diffuse gamma(10^{-4} , 10^{-4}) distributions as priors, although informative priors can be used if desired. Because this is a multiple-parameter problem, we use at least two MCMC chains in WinBUGS and monitor convergence. Initial values can be chosen based on maximum likelihood estimates, which must be found numerically, or simpler method-of-moment estimates. In practice, convergence is not an issue for these problems and accurate initial values are not needed. Note that α and β will be highly correlated in the joint posterior distribution.

Weibull Distribution

In this case, each observed duration is treated as a random sample from a Weibull distribution, with density

$$f(t_i | \alpha, \lambda) = \alpha \lambda t_i^{\alpha-1} \exp(-\lambda t_i^\alpha) \quad \alpha, \lambda > 0 \quad (41)$$

Note that this is the parameterization used by WinBUGS. An alternative parameterization sometimes used is obtained by defining $\lambda = \beta^\alpha$, giving

$$f(t_i | \alpha, \beta) = \frac{\alpha}{\beta} \left(\frac{t_i}{\beta}\right)^{\alpha-1} \exp\left[-\left(\frac{t_i}{\beta}\right)^\alpha\right] \quad (42)$$

As above, if the shape parameter, α , equals one, then the Weibull distribution reduces to the exponential distribution. Likewise, for $\alpha > (<) 1$, the intensity increases (decreases) with time.

If n durations are observed, the likelihood function becomes

$$\begin{aligned} f(t_1, t_2, \dots, t_n | \alpha, \lambda) &= \prod_{i=1}^n f(t_i | \alpha, \lambda) \\ &= \alpha^n \lambda^n \prod_{i=1}^n t_i^{\alpha-1} \exp(-\lambda \sum_{i=1}^n t_i^\alpha) \end{aligned} \quad (43)$$

The analyst must specify a prior distribution for α and λ . In most of the scripts, we use independent, diffuse gamma(10^{-4} , 10^{-4}) distributions as priors, although informative priors can be used if desired. Because this is a multiple-parameter problem, we use at least two MCMC chains in WinBUGS and monitor convergence. Initial values can be chosen based on maximum likelihood estimates, which must be found numerically. In practice, convergence is not an issue for these problems and accurate initial values are not needed. Note that α and λ will be highly correlated in the posterior distribution.

Lognormal Distribution

In this case, each observed duration is treated as a random sample from a lognormal distribution, with density

$$f(t_i | \mu, \tau) = \sqrt{\frac{\tau}{2\pi}} \frac{1}{t_i} \exp\left[-\frac{\tau}{2} (\log(t_i) - \mu)^2\right] \quad -\infty < \mu < \infty, \tau > 0 \quad (44)$$

The lognormal distribution does not have a shape parameter and the intensity first increases rapidly, peaks, and then decreases monotonically as a function of increasing time. If n durations are observed, the likelihood function becomes

$$\begin{aligned} f(t_1, t_2, \dots, t_n | \mu, \tau) &= \prod_{i=1}^n f(t_i | \mu, \tau) \\ &= \left(\frac{\tau}{2\pi}\right)^{n/2} \prod_{i=1}^n \frac{1}{t_i} \exp\left[-\frac{\tau}{2} \sum_{i=1}^n (\log(t_i) - \mu)^2\right] \end{aligned} \quad (45)$$

The analyst must specify a prior for μ and τ . As before, it is common, though not necessary, to use independent diffuse priors. For μ , one can use an improper Laplace prior over the entire real axis. One could use a diffuse gamma prior for τ , or one may wish to reparameterize in terms of the logarithmic standard deviation, given by

$$\sigma = \frac{1}{\sqrt{\tau}} \quad (46)$$

To avoid numerical difficulties that are sometimes encountered with the diffuse gamma(10^{-4} , 10^{-4}) distribution in this type of problem, one can use a uniform prior for σ , as recommended by (Gelman, 2006). A value of 1.4 for σ corresponds to an error factor of 10, so a uniform(0, 5) prior for σ is quite diffuse. The analyst should examine the marginal posterior distribution for σ to ensure that the upper tail is not truncated, as truncation of the upper tail suggests a more diffuse prior should be used. See (Kelly & Atwood, 2008) for more details.

Modeling Failure with Repair (3.7)

The guidebook discusses two extremes of repair, which are felt to span the practical state of the art:

- Repair same as new (“reincarnation”)
- Repair same as old (“resuscitation”)

If the assumption of same-as-new repair is made (renewal process), then the observed times between failures are treated as a random sample from a common renewal density (e.g., Weibull), and the methods above (Sec. 3.6) can be applied.

If the assumption is that repair leaves the component in the state it was in at the time of failure (repair same as old), then we are dealing with a nonhomogeneous Poisson process (NHPP) in which the times between failures do not constitute a random sample from a common distribution. Instead, each cumulative failure time is conditional upon the preceding times. There are various ways in which to parameterize the rate of occurrence of failure (ROCOF) in this case; the guidebook considers only a power-law process, in which the ROCOF is given by¹

$$\lambda(t) = \frac{\beta}{\alpha} \left(\frac{t}{\alpha} \right)^{\beta-1} \quad \alpha, \beta > 0 \quad (47)$$

If $\beta > 1$, then we have wearout (aging), and if $\beta < 1$, we have reliability growth. If $\beta = 1$, then there is no trend in the ROCOF with time. If the process is failure-truncated, and t_i is the cumulative time until the i^{th} failure, then the likelihood function is given by

$$f(t_1, t_2, \dots, t_n | \alpha, \beta) = \frac{\beta^n}{\alpha^{n\beta}} \prod_{i=1}^n t_i^{\beta-1} \exp \left[- \left(\frac{t_n}{\alpha} \right)^\beta \right] \quad (48)$$

This equation can be derived from the fact that the time to first failure has a Weibull (β , α) distribution, and each succeeding cumulative failure time has a Weibull distribution, truncated on the left at the preceding failure time. The analyst must specify a prior for the parameters α and β , as before. Model validation in this case uses the Bayesian analog of the Cramer-von Mises statistic described earlier, and is based on the fact that, for the power-law process, if one defines $z_i = (t_i/\alpha)^\beta$, then $z_{\text{inc}, i} = z_i - z_{i-1}$ has an exponential distribution with parameter 1.

¹ In earlier sections we used α as the shape parameter; we use β as the shape parameter in this section to conform to the notation typically used in the literature for this type of problem.

Treatment of Censoring (3.8)

The guidebook treats two kinds of censoring for data in the form of random durations. In the first, called Type I censoring in the literature, n items are placed on test, and the test is terminated at time t , at which point the number of items that has failed is random. If r out of n items have failed by time t , the likelihood function becomes (assuming an exponential aleatory model for T)

$$f(t_1, t_2, \dots, t_n | \lambda) = \lambda^r \exp\left\{-\lambda \left[\sum_{i=1}^r t_i + (n-r)t \right]\right\} \quad (49)$$

With a Jeffreys prior for λ , the posterior distribution is given by

$$\pi_1(\lambda | t_1, t_2, \dots, t_n) = \frac{\lambda \exp\left\{-\lambda \left[\sum_{i=1}^r t_i + (n-r)t \right]\right\}}{\int_0^\infty \frac{\lambda^r \exp\left\{-\lambda \left[\sum_{i=1}^r t_i + (n-r)t \right]\right\}}{\lambda} d\lambda} \quad (50)$$

The second type of censoring treated by the guidebook is interval censoring, in which some random durations are observed precisely, but others are observed in a fuzzier manner, and are known to lie in an interval, $(t_{\text{lower}}, t_{\text{upper}})$. If n durations are observed, and r of these have precise observed values, t_i , with the remaining $n-r$ values observed in specified intervals, I_j the likelihood function becomes (assuming an exponential aleatory model for T)

$$f(t_1, t_2, \dots, t_n | \lambda) = \lambda^r \exp\left(-\lambda \sum_{i=1}^r t_i\right) \prod_{j=1}^{n-r} [\exp(-\lambda t_{\text{lower},j}) - \exp(-\lambda t_{\text{upper},j})] \quad (51)$$

With a Jeffreys prior for λ , the posterior distribution is given by

$$\pi_1(\lambda | t_1, t_2, \dots, t_n) = \frac{\lambda^r \exp\left(-\lambda \sum_{i=1}^r t_i\right) \prod_{j=1}^{n-r} [\exp(-\lambda t_{\text{lower},j}) - \exp(-\lambda t_{\text{upper},j})]}{\int_0^\infty \frac{\lambda^r \exp\left(-\lambda \sum_{i=1}^r t_i\right) \prod_{j=1}^{n-r} [\exp(-\lambda t_{\text{lower},j}) - \exp(-\lambda t_{\text{upper},j})]}{\lambda} d\lambda} \quad (52)$$

If T is time to failure, as in Example 12 in the guidebook, then the probability of operating longer than time t_{crit} is given by averaging the survival probability, which is conditional upon a value of λ , over the posterior distribution for λ :

$$P(T > t_{crit}) = \int_0^{\infty} e^{-\lambda t_{crit}} \pi_1(\lambda | t_1, t_2, \dots, t_n) d\lambda \quad (53)$$

References

- Berger, J. O. (1985). *Statistical Decision Theory and Bayesian Analysis, Second Edition*. Springer.
- Brooks, S. P., & Gelman, A. (1998). General Methods for Monitoring Convergence of Iterative Simulations. *Journal of Computational and Graphical Statistics*, 434-455.
- Gelman, A. (1996). Inference and Monitoring Convergence. In W. R. Gilks, & e. al., *Markov Chain Monte Carlo in Practice* (pp. 131-143). Chapman & Hall/CRC.
- Gelman, A. (2006). Prior Distributions for Variance Parameters in Hierarchical Models. *Bayesian Analysis*, 1 (3), 515-533.
- Gelman, A., & al., e. (2004). *Bayesian Data Analysis, Second Edition*. Chapman & Hall/CRC.
- Gelman, A., & al., e. (1996). Posterior Predictive Assessment of of Model Fitness via Realized Discrepancies. *Statistica Sinica*, 733-807.
- Kaplan, S. (1983). On a Two-Stage Bayesian Procedure for Determining Failure Rates. *IEEE Transactions on Power Apparatus and Systems*, 195-262.
- Kelly, D. L., & Atwood, C. L. (2008). Bayesian Modeling of Population Variability: Practical Guidance and Pitfalls. *Ninth International Conference on Probabilistic Safety Assessment and Management*. Hong Kong.
- Lofgren, E. V., & al., e. (2005). *Space Shuttle Probabilistic Risk Assessment Functional Failure Data Report*. NASA.
- Martz, H. F., & Picard, R. R. (1998). On Comparing PRA Results with Operating Experience. *Reliability Engineering and System Safety*, 187-199.
- Mosleh, A. (1991). Common Cause Failures: An Analysis Methodology and Examples. *Reliability Engineering and System Safety*, 34, 249-292.
- Mosleh, A., & Apostolakis, G. (1984). Models for the Use of Expert Opinions. In R. A. Waller, & V. T. Covello (Ed.), *Low-Probability/High-Consequence Risk Analysis: Issues, Methods, and Case Studies*. Plenum.
- Siu, N. O., & Kelly, D. L. (1998). Bayesian Parameter Estimation in Probabilistic Risk Assessment. *Reliability Engineering and System Safety*, 89-116.
- Spiegelhalter, & al., e. (2002). Bayesian Measures of Model Complexity and Fit. *Journal of the Royal Statistical Society B*, 64, 583-640.

European Commission

EUR 23954 EN – Joint Research Centre – Institute for Energy

Title: Guidelines for Analysis of Data Related to Aging of Nuclear Power Plant Components and Systems. Overview of NPPs component reliability data collection with regards to time-dependent reliability analysis applications. EC JRC Network on Use of Probabilistic Safety Assessments (PSA) for Evaluation of Aging Effects to the Safety of Energy Facilities. Task 4.

Author(s): A.Rodionov, D.Kelly, J.U.Klugel

Luxembourg: Office for Official Publications of the European Communities

2008 – 129 pp. – 21 x 29.7 cm

EUR – Scientific and Technical Research series – ISSN 1018-5593

Abstract

This guideline is intended to provide practical methods for practitioners to use in analyzing component and system reliability data, with a focus on detection and modeling of ageing. The emphasis is on frequentist and Bayesian approaches, implemented with MS EXCEL and the open-source software package WinBUGS. The methods described in this document can be implemented with other software packages.

The mission of the JRC is to provide customer-driven scientific and technical support for the conception, development, implementation and monitoring of EU policies. As a service of the European Commission, the JRC functions as a reference centre of science and technology for the Union. Close to the policy-making process, it serves the common interest of the Member States, while being independent of special interests, whether private or national.

